



**Cite this article:** Velichko A, Bai L, Drinkwater BW. 2017 Ultrasonic defect characterization using parametric-manifold mapping. *Proc. R. Soc. A* **473**: 20170056. <http://dx.doi.org/10.1098/rspa.2017.0056>

Received: 26 January 2017

Accepted: 5 May 2017

**Subject Areas:**

mechanical engineering

**Keywords:**

ultrasonic non-destructive testing, defect characterization, parametric manifold

**Author for correspondence:**

A. Velichko

e-mail: [a.velichko@bristol.ac.uk](mailto:a.velichko@bristol.ac.uk)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3787379>.

# Ultrasonic defect characterization using parametric-manifold mapping

A. Velichko, L. Bai and B. W. Drinkwater

Department of Mechanical Engineering, University of Bristol, Queens Building, University Walk, Bristol BS8 1TR, UK

LB, 0000-0002-0482-1514

The aim of ultrasonic non-destructive evaluation includes the detection and characterization of defects, and an understanding of the nature of defects is essential for the assessment of structural integrity in safety critical systems. In general, the defect characterization challenge involves an estimation of defect parameters from measured data. In this paper, we explore the extent to which defects can be characterized by their ultrasonic scattering behaviour. Given a number of ultrasonic measurements, we show that characterization information can be extracted by projecting the measurement onto a parametric manifold in principal component space. We show that this manifold represents the entirety of the characterization information available from far-field harmonic ultrasound. We seek to understand the nature of this information and hence provide definitive statements on the defect characterization performance that is, in principle, extractable from typical measurement scenarios. In experiments, the characterization problem of surface-breaking cracks and the more general problem of elliptical voids are studied, and a good agreement is achieved between the actual parameter values and the characterization results. The nature of the parametric manifold enables us to explain and quantify why some defects are relatively easy to characterize, whereas others are inherently challenging.

## 1. Introduction

The aim of ultrasonic non-destructive evaluation and structural health monitoring includes the detection and characterization of defects [1–5], and an understanding

of the nature of defects is essential for the assessment of structural integrity in safety critical systems [6]. There are several types of defects that can potentially occur in a structure, whose effects need to be treated differently. Cracks are commonly regarded as ‘the defects of most concern’ [7] because they can lead to rapid growth and hence failure of a structure [8]. For this reason, characterization of cracks has been studied extensively in the literature, including measurements based on single element transducers [9,10] and transducer arrays [11,12]. The proposed approaches include measuring the scattered amplitude [10], the time-of-flight diffraction technique [13], image-based characterization [11,14] and characterization using the scattering matrix [11,12], which gives a far-field scattering amplitude for every combination of incident and scattered directions.

Once a defect has been non-destructively detected, the next requirement is to characterize it and hence to discover its physical nature. In the absence of any characterization information, the worst case scenario is assumed, which is often to classify the detected defect as an unfavourably (w.r.t. the loading) oriented surface-breaking crack. Here, the worst case defect is one that will most rapidly result in a failure of the structure. Improved defect characterization information allows the worst case scenario to be replaced by a more accurate representation of reality. Of course, in some cases the characterization may confirm the worst case scenario, but in others the detected defects can be relatively benign, for example rounded pores or inclusions introduced in the manufacturing process. Whatever the characterization result, the outcome of more accurate characterization is a better prediction of the remaining life of a structure.

There are two main approaches to the defect characterization problem seen in the literature. The first attempts to reconstruct the defect geometry without any preliminary assumptions about the defect. For example, an iterative method can be applied, in which a defect geometry is iteratively updated and a forward scattering model is used to calculate a scattering matrix until a match with the measured data is reached [15]. In some cases a model-based inversion, which makes use of an approximate analytical expression for the forward scattering problem, can be applied. This approach is widely used in guided waves tomography for the reconstruction of the thickness map of corrosion damage [16]. Alternatively, semi-analytical inversion schemes have been developed [17]. However, a regularization procedure is required in general inversion approaches [18] in order to deal with the ill-posedness of the studied problem, and issues including numerical stability and convergence need to be addressed [19]. This has led researchers to consider more targeted approaches, and the second approach makes use of the fact that there is only a limited number of possible defect types that can occur in practice. Each defect type can then be efficiently described by a limited set of parameters. In this case, the characterization problem becomes one of estimating this smaller set of defect parameters from the measured scattering data. Typically, the method of solution adopted is to form a large database from many forward simulations and then compare experimental data with that in the database, i.e. the database is searched and the closest match is used to characterize the defect. In these database search methods, use can be made of the vast body of the literature on classification algorithms and their applications [20–24]. Neural networks and the support vector machine are examples of the widely used classification approaches. Their use can be found in a range of applications, including radar target recognition [25], underwater target classification [26], classification of electroencephalogram signals [27] and in bioinformatics (e.g. gene selection for cancer classification [28]). Using this classification algorithm approach to the characterization problem, good results have been achieved on simulated data and in idealized experiments containing machined notches [12] and volumetric elliptical voids [29]. In addition, the effect of coherent grain noise on these classification schemes was explored [30].

However, there is a philosophical problem with the use of the above classification approach—the classification algorithms are ‘black-boxes’ completely defined by the training data. In addition, the classification approach relies on subjective choices such as the defect class definition and the number of defect classes, which limits the generalizability of the approach, especially when three or more parameters need to be determined from the inversion process. In this paper, a new defect characterization method is proposed to address the above issues. The crucial observation is that

the defect database can be represented by a parametric manifold in the measurement domain, and can be approximated to any resolution with a finite number of training samples using interpolation schemes. As such, any characterization procedure is fully determined by the shape of the parametric manifold, which also determines the achievable characterization accuracy.

The proposed approach has some important practical benefits. Firstly, the characterization is performed in the principal component domain. In this case, only a few largest principal components can be taken, which resolves the issues related to high dimensional characterization spaces. Secondly, the manifold representation of the parametric defect space provides a much more intuitive/insightful geometric understanding and helps to ‘visualize’ the defect characterization problem. The proposed method naturally handles any measurement scheme (e.g. arrays, multiple probes and scanning) and allows us to consider different defect types. Thirdly, in practice, the available measurement information is always limited, and, in this case, an estimation of the characterization uncertainty is critical. The proposed approach naturally introduces the method for the characterization uncertainty estimation, so the final characterization result is represented by the probability density distribution in the defect parameter space.

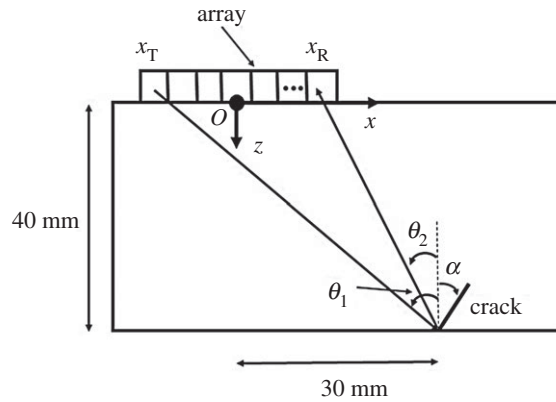
Here, we focus our efforts on the use of an ultrasonic array to characterize two different defect types: surface-breaking cracks and elliptical voids. All defects are assumed to be two dimensional and a one-dimensional linear ultrasonic array is used to perform relevant measurements. However, it should be stressed that the characterization procedure proposed in the paper is independent of the dimensionality of defects and potentially can be directly applied to characterize three-dimensional defects. As stated above, cracks are detrimental to structural integrity, and the detection and characterization of surface-breaking cracks is of particular industrial interest. Volumetric voids (pores) are also important because they can be potential initiation sites of cracks [31], and ellipses provide a reasonable simplification of a wide range of volumetric voids. Surface-breaking cracks can be described by two parameters: size and orientation angle. For elliptical voids, besides size and orientation angle, a third parameter—aspect ratio—is needed to define the shape. Hence, the characterization of elliptical voids naturally extracts the aspect ratio, which is a measure of sharpness and could be useful in fatigue life predictions.

## 2. Defect characterization problem

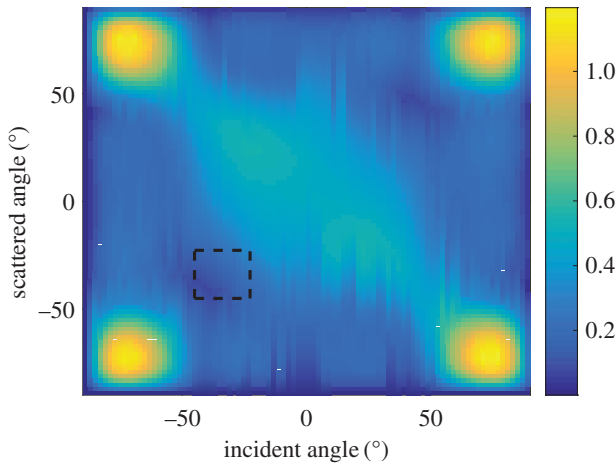
Figure 1 shows a typical ultrasonic array inspection configuration for surface-breaking cracks located on the back face of a flat plate. The incident and scattering angles and the orientation angle of surface-breaking cracks are all defined with respect to the back surface normal, and are positive if measured clockwise. In the example shown here, the material is aluminium (Young’s modulus, 69 GPa; Poisson’s ratio, 0.334; density, 2700 kg m<sup>-3</sup>), and the thickness of the test specimen is 40 mm. A 2.5 MHz, 64-element array is used as an example throughout this paper, and the element pitch of the array is 0.50 mm (i.e. 0.2 $\lambda$  at the centre frequency). In figure 1, the array is moved away from the crack by 30 mm, because otherwise the crack would be obscured by the high-intensity reflection from the back wall in the image.

Firstly, defects with idealized geometries are considered, so the cracks are assumed to be smooth. In this case, each surface-breaking crack is fully characterized by two parameters—length and orientation angle. In reality, all cracks have some level of roughness, which is usually described by RMS height,  $\sigma_{\text{rough}}$ , and correlation length,  $\lambda_{\text{rough}}$  [32]. However, if the roughness level is not too high, so that the RMS height is much smaller than the ultrasonic wavelength, then the geometry of the crack can still be approximately described by its length and orientation angle. The defect characterization problem can then be formulated as estimating the crack parameters from the ultrasonic array measurements.

From the defect characterization point of view, the information that can be extracted from the ultrasonic array data is the scattering behaviour of the defect. The scattering information can be represented in the form of a scattering matrix [11]. More specifically, for a unit amplitude plane incident wave propagating in the direction  $\theta_{\text{in}}$ , the scattered wave in the far field of the defect at



**Figure 1.** Array measurement configuration in simulation and experiments (for surface-breaking cracks).



**Figure 2.** The scattering matrix of a  $1\lambda, 0^\circ$  surface-breaking crack. The dashed box represents the angular range extractable from the measurement configuration shown in figure 1. (Online version in colour.)

the direction  $\theta_{sc}$  is given by

$$u_{sc} = \sqrt{\frac{\lambda}{r}} S(\theta_{in}, \theta_{sc}, \omega) e^{ikr}, \quad (2.1)$$

where  $\lambda$  is the ultrasonic wavelength,  $k = 2\pi/\lambda$  is the wavenumber,  $\omega$  is the angular frequency,  $r$  is the distance from the nominal defect centre and  $S$  is the scattering matrix. Note that the proposed characterization method does not depend on the particular testing procedure used to perform measurements. To illustrate the performance of the characterization method simple direct contact measurements were used. In this case, the scattering information extracted from the ultrasonic array data corresponds to the longitudinal waves, and shear waves and mode conversion effects are not considered. However, the proposed approach is generally applicable to any experimental configuration and wave mode type.

The scattering matrices of surface-breaking cracks were simulated using an efficient finite-element model [33], and figure 2 shows the scattering matrix of a  $1\lambda, 0^\circ$  surface-breaking smooth crack. Although the scattering matrix is shown for the range of  $[-90^\circ, 90^\circ]$ , the maxima around the incident and scattered angles of  $\pm 90^\circ$  are not extractable for the example shown in figure 1 because of the finite size of the array aperture. The dashed box in figure 2 represents the

angular range extractable from the measurement configuration shown in figure 1. In general, the scattering matrix is a complex-valued function; however, its phase depends on the position of the nominal centre of the defect and, as this is not known *a priori*, it is difficult to extract the phase from the measurements. Therefore, in this paper only the amplitude of the scattering matrix is considered. However, it is noted that further information is contained in the phase part of the scattering matrix and future work could look at how to make use of this reliably.

### 3. Defect manifold

#### (a) Principal component analysis

The defect characterization problem consists of estimating the defect parameters from the extracted components of the scattering matrix. However, before the characterization procedure is developed, we explore the general structure of the set of the scattering matrices for a certain defect type.

Let us consider the set of all possible scattering matrices for a particular measurement configuration

$$S = \{S(\theta_{\text{in}}, \theta_{\text{sc}}, \omega), \theta_{\text{in},1} \leq \theta_{\text{in}} \leq \theta_{\text{in},2}, \theta_{\text{sc},1} \leq \theta_{\text{sc}} \leq \theta_{\text{sc},2}, \omega_1 \leq \omega \leq \omega_2\}, \quad (3.1)$$

which will be referred to as the  $s$ -space.

For any defect characterization procedure, some metric is required to quantify the distance between a measurement and the entire  $s$ -space. The distance (or similarity) between two scattering matrices can be measured in many different ways. For example, the  $L_2$  (or Euclidean) norm can be used as

$$\left. \begin{aligned} \|S_1 - S_2\| &= \sqrt{(S_1 - S_2, S_1 - S_2)_{L_2}}, \\ (S_1, S_2)_{L_2} &= \int S_1 S_2 \, d\theta_{\text{in}} \, d\theta_{\text{sc}} \, d\omega. \end{aligned} \right\} \quad (3.2)$$

In practice, the scattering matrix can be measured at a finite number of incident and scattered angles and frequencies only. For simplicity, it is assumed that the measurement angles,  $\theta_{\text{in},n}$ ,  $\theta_{\text{sc},m}$ , and frequencies,  $\omega_k$ , are equally spaced and there are  $N_{\text{in}}$  incident angles,  $N_{\text{sc}}$  scattered angles and  $N_f$  frequencies. In this case, each scattering matrix  $S$  can be considered as a vector  $\mathbf{s} = \{s_1, \dots, s_{N_s}\}^T$ ,  $s_l = S(\theta_{\text{in},n}, \theta_{\text{sc},m}, \omega_k)$  in an  $N_s = N_{\text{in}} \times N_{\text{sc}} \times N_f$ -dimensional space.

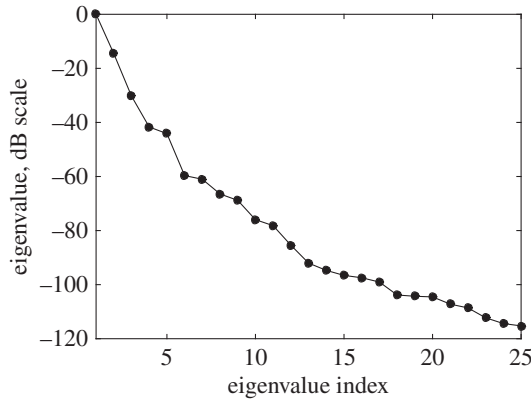
For the typical ultrasonic array inspection  $N_{\text{in}}, N_{\text{sc}}, N_f \sim 10$ , therefore, the dimension of the  $s$ -space is  $N_s \sim 10^3$ . The analysis of data in such high dimensional space is difficult. However, in many cases each scattering matrix can be efficiently described by much fewer components. One possibility to reveal this reduction in dimensionality is to apply principal component analysis (PCA) [34,35] to the scattering matrices. PCA takes into account the variability of data in different directions, and chooses a new coordinate system, so each point in the manifold is represented with fewer coordinates, which are referred to as the principal components hereafter.

The class of defects that can be parametrically described by  $N_p$  parameters  $\mathbf{p} = \{p_1, \dots, p_{N_p}\}$ ,  $p_{1n} \leq p_n \leq p_{2n}$ ,  $n = 1, \dots, N_p$ , is now considered. The corresponding set of scattering matrices is defined as  $S_p = S(\theta_{\text{in}}, \theta_{\text{sc}}, \omega; \mathbf{p})$ . The set of scattering matrices in the defect class can be represented by the  $N_s \times M$  matrix  $\mathbf{S}_p = \{s_p(\mathbf{p}_1) - \bar{s}_p, \dots, s_p(\mathbf{p}_M) - \bar{s}_p\} \in \mathbb{R}^{N_s \times M}$ , where  $M$  is the total number of scattering matrices in the defect class, vector  $\mathbf{s}_p$  corresponds to the defect scattering matrix  $S_p$  and  $\bar{s}_p$  is the vector corresponding to the mean scattering matrix of the defect class. The new coordinate system is obtained by singular value decomposition of the covariance matrix of  $\mathbf{S}_p$

$$\mathbf{R} = \mathbf{V} \mathbf{D} \mathbf{V}^T, \quad (3.3)$$

where  $\mathbf{R}$  is the  $N_s \times N_s$  covariance matrix of  $\mathbf{S}_p$  and can be obtained as

$$\mathbf{R} = \frac{1}{M-1} \mathbf{S}_p \mathbf{S}_p^T. \quad (3.4)$$



**Figure 3.** Eigenvalues of the covariance matrix (25 largest eigenvalues) for surface-breaking cracks.

The diagonal matrix  $\mathbf{D}$  contains the eigenvalues of  $\mathbf{R}$ . The coordinate axes of the new coordinate system are now given by the column vectors of  $\mathbf{V}$ , so any scattering matrix,  $\mathbf{s}$ , in the new coordinate system is represented by the vector  $\mathbf{s}^{(\text{pc})}$  as

$$\mathbf{s}^{(\text{pc})} = \mathbf{V}^T(\mathbf{s} - \bar{\mathbf{s}}_p). \quad (3.5)$$

After the application of PCA, the defect class is effectively embedded in a lower dimensional space (which we call principal component space, or pc-space), since the first  $N_{\text{pc}}$  coordinates account for most of the variation of the set  $S_p$ . Normally, we have  $N_{\text{pc}} \ll N_s$ , and the value of  $N_{\text{pc}}$  can be determined, for example, by setting some threshold as

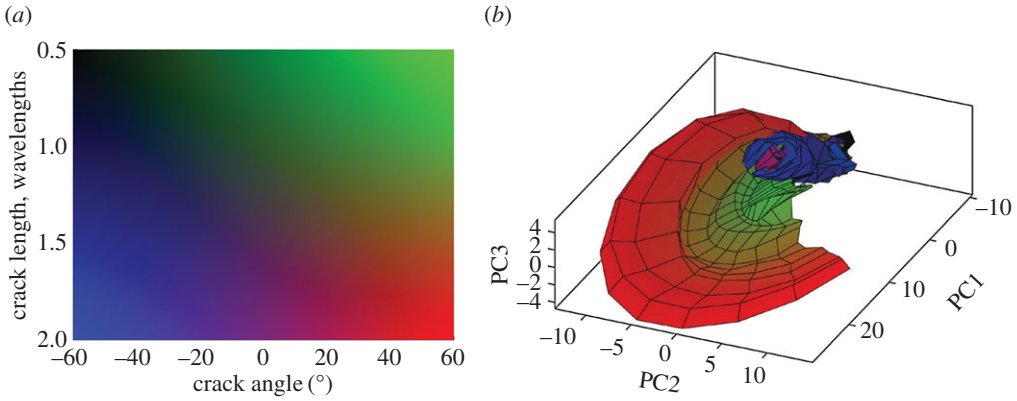
$$\frac{d_n}{\max d_n} \geq d_0, \quad n = 1, \dots, N_{\text{pc}}, \quad (3.6)$$

where  $d_n$  are the eigenvalues of covariance matrix  $\mathbf{R}$  and  $d_0$  is the threshold. Therefore, only principal components with eigenvalues greater than the threshold are retained.

## (b) Structure of defect manifold

The important observation is that the set of defect class scattering matrices represents an  $N_p$ -dimensional manifold in the pc-space. This manifold will be referred to as the defect manifold or  $d$ -manifold. The structure of the  $d$ -manifold provides a fundamental insight into the defect characterization problem and determines the achievable characterization accuracy.

For the example configuration considered in §2 (figure 1), the defect parameter space is defined as  $\mathbf{p} = \{\theta_{\text{crack}}, a_{\text{crack}}\}$ ,  $-60^\circ \leq \theta_{\text{crack}} \leq 60^\circ$ ,  $0.5\lambda \leq a_{\text{crack}} \leq 2\lambda$ , where  $\theta_{\text{crack}}$  is the crack orientation and  $a_{\text{crack}}$  is the crack length. The scattering matrices were calculated on a uniform grid in the parameter space with the parameter steps  $\Delta\theta_{\text{crack}} = 2^\circ$ ,  $\Delta a_{\text{crack}} = 0.1\lambda$ . Figure 3 shows the first 25 largest eigenvalues of the covariance matrix  $\mathbf{R}$ . The threshold of  $d_0 = 0.005$  was taken, and in this case  $N_{\text{pc}} = 5$ . Figure 4a,b shows the parameter  $p$ -space and the shape of the  $d$ -manifold in three-dimensional pc-space, respectively. Each point on the defect manifold corresponds to some particular point in the defect parameter space. To visualize this mapping, a colour map in the defect parameter space was used, and the same colour map was then used to plot the  $d$ -manifold. In other words, the same colours in the defect parameter space and the defect manifold correspond to one particular set of defect parameters. It can be seen (by observing the red region in figure 4) that, for crack parameters  $\lambda \leq a_{\text{crack}} \leq 2\lambda$ ,  $40^\circ \leq \theta_{\text{crack}} \leq 60^\circ$ , this surface is well resolved, suggesting that in this region there is enough information for unique defect characterization. Note that, at these crack orientation angles, the specular reflection from the crack can be detected by the array, and, therefore, the measured part of the scattering matrix



**Figure 4.** Surface-breaking cracks: (a) parameter space; (b)  $d$ -manifold in three-dimensional pc-space.

contains its maximum peak. For other parameters of the crack the shape of the  $d$ -manifold is more complex, which can be explained by the following factors. Firstly, the corresponding regions of the  $d$ -manifold are defined by more than three principal components and cannot be visualized in the three-dimensional pc-space. Secondly, in this case, the specular reflection from the crack is not detected by the array, and the measured part of scattering matrix is less sensitive to the crack parameters. Therefore, the points  $\mathbf{s}^{(\text{pc})}$ , which represent the scattering matrices in pc-space, are located close to each other and hence defect characterization is more challenging.

We can now explore the global structure of the  $d$ -manifold. The distance between two scattering matrices  $S_{p1} = S_p(\mathbf{p}_1)$ ,  $S_{p2} = S_p(\mathbf{p}_2)$  can be measured in pc-space using the  $L_2$  norm defined by (3.2) as

$$\|S_{p1} - S_{p2}\|_{\text{pc}} = \sqrt{(\mathbf{s}_{p1}^{(\text{pc})} - \mathbf{s}_{p2}^{(\text{pc})}, \mathbf{s}_{p1}^{(\text{pc})} - \mathbf{s}_{p2}^{(\text{pc})})}. \quad (3.7)$$

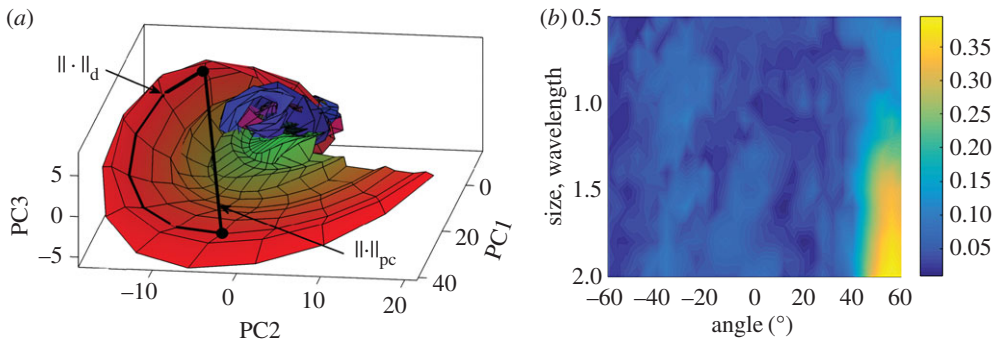
Alternatively, the distance metric in pc-space leads to a corresponding distance metric on the  $d$ -manifold. Then, an alternative distance measure between two scattering matrices  $S_{p1}$ ,  $S_{p2}$  can be taken as the length of the geodesic line between  $\mathbf{s}_{p1}^{(\text{pc})}$  and  $\mathbf{s}_{p2}^{(\text{pc})}$  on the  $d$ -manifold, or the minimum length of all possible paths between  $\mathbf{s}_{p1}^{(\text{pc})}$  and  $\mathbf{s}_{p2}^{(\text{pc})}$  on the  $d$ -manifold,

$$\|S_{p1} - S_{p2}\|_d = \min_{\gamma} \int_{\mathbf{s}_{p1}^{(\text{pc})}}^{\mathbf{s}_{p2}^{(\text{pc})}} |d\mathbf{s}_p^{(\text{pc})}|. \quad (3.8)$$

The physical meaning of the metric (3.8) is illustrated in figure 5a. The  $d$ -manifold distance metric is given by the length of the geodesic line (the shortest distance) connecting two points on the  $d$ -manifold. It can be seen that the manifold distance (3.8) between any two scattering matrices is always greater than the Euclidean distance (3.7) between the same scattering matrices in pc-space,  $\|\cdot\|_d \geq \|\cdot\|_{\text{pc}}$ . Then, the geometry of the  $d$ -manifold can be characterized by the following dimensionless parameter  $I_d$ , which is referred to as the  $d$ -index:

$$I_d(\mathbf{p}) = \min_{\mathbf{q}} \frac{\|S_p(\mathbf{p}) - S_p(\mathbf{q})\|_{\text{pc}}}{\|S_p(\mathbf{p}) - S_p(\mathbf{q})\|_d}, \quad (3.9)$$

where  $\mathbf{p}$ ,  $\mathbf{q}$  are parameter vectors from the defect parameter space. The  $d$ -index is always less than 1,  $0 \leq I_d \leq 1$ , and characterizes the stability of the characterization result. Qualitatively, a small value of the  $d$ -index at the point  $\mathbf{p}$  means that there exists another parameter vector  $\mathbf{q}$  which is not located in the vicinity of  $\mathbf{p}$ , but the scattering matrices  $S_p(\mathbf{p})$  and  $S_p(\mathbf{q})$  are very similar to each



**Figure 5.** (a) Schematic of distance metric in the pc-space,  $\|\cdot\|_{pc}$ , and on the  $d$ -manifold,  $\|\cdot\|_d$ ; (b) map of the  $d$ -index for the surface-breaking cracks. (Online version in colour.)

other. Therefore, at such parameter points the defect characterization will be sensitive to noise and unique defect characterization could be difficult.

Figure 5b shows the  $d$ -index of the defect manifold as a function of surface-breaking crack parameters (orientation angle and crack length). The calculations of (3.8) were performed using Dijkstra's algorithm [36]. It can be seen that in the region  $\lambda \leq a_{\text{crack}} \leq 2\lambda$ ,  $40^\circ \leq \theta_{\text{crack}} \leq 60^\circ$  of the parameter space the  $d$ -index is high, suggesting that the characterization result will be stable with respect to noise. However, characterization in all other points will be more sensitive to noise and, therefore, more uncertain. Note that this result agrees with the three-dimensional shape of the  $d$ -manifold shown in figure 4b, but also allows us to quantitatively characterize the structure of the  $d$ -manifold in higher dimensional spaces.

## 4. Defect characterization method

Note that each defect class consists of a database of defects with idealized geometries, for example straight cracks with different lengths and orientations. An experimentally measured scattering matrix always contains some noise compared with the scattering matrices in the defect database. This noise can be a random experimental noise, but also coherent noise which is caused by the differences between real defect geometries and those of the defect in the database, for example surface roughness of the crack. There are also a number of other sources of coherent noise apart from the defect roughness. For example, noise can arise from interference from other defects and structural features located near the defect of interest. This noise is the result of imperfect focusing due to the finite array aperture and the diffraction limit. Another possible reason for coherent noise is the limitations of the two-dimensional model of the ultrasonic array and defect scattering. In any case, this means that, in practice, it will be impossible to achieve an exact match between an experimentally extracted scattering matrix and the scattering matrices from some idealized defect class.

Based on the analysis performed above, it can be seen that any experimental scattering matrix can be represented by a vector  $\mathbf{s}_{\text{exp}}^{(pc)}$  in pc-space. However, in general, this vector is not necessarily located on the  $d$ -manifold of the defect class because of the combined effects of random and coherent experimental noise. This can be expressed as

$$\mathbf{s}_{\text{exp}}^{(pc)} = \mathbf{s}_p^{(pc)} + \mathbf{n}_p^{(pc)}, \quad (4.1)$$

where the vector  $\mathbf{s}_p^{(pc)}$  corresponds to the scattering matrix of the defect with the parameter  $\mathbf{p}$  and  $\mathbf{n}_p^{(pc)}$  represents a noise perturbation of this scattering matrix in pc-space.

It is assumed that the noise is described by some statistical model, so the noise vector  $\mathbf{n}_p^{(pc)}$  in expression (4.1) represents one particular realization of the noise. Then, the defect characterization



problem can be formulated in the following way: given the occurrence of the measured scattering matrix  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$ , what is the probability that this scattering matrix can be represented by the defect parameter  $\mathbf{p}$ ? Note that the defect parameter  $\mathbf{p}$  represents a continuous variable. Therefore, the probability is given by  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})d\mathbf{p}$ , where  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})$  is the conditional probability density function of the defect parameter  $\mathbf{p}$  and  $d\mathbf{p} = \Delta p_1 \times \dots \times \Delta p_{N_p}$  is a volume element in parameter space.

Alternatively, this question can be formulated in a different form: given the defect parameter  $\mathbf{p}$ , what is the probability that, for some particular noise realization, the perturbed scattering matrix  $\mathbf{s}_p^{(\text{pc})} + \mathbf{n}_p^{(\text{pc})}$  will match the experimental measurements  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$ ? Similar to the previous case, this probability is described by  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p})ds^{(\text{pc})}$ , where  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p})$  is the conditional probability density function of the vector  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$  in pc-space and  $ds^{(\text{pc})}$  is a volume element in pc-space.

It is easy to show that the probability density functions  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})$  and  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p})$  are directly related to each other. According to Bayes' theorem, the posterior probability  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})$  is given by

$$\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})}) = \rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p}) \frac{\rho_p(\mathbf{p})}{\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})})}, \quad (4.2)$$

where  $\rho_p(\mathbf{p})$  is the marginal probability density distribution of the defect parameters  $\mathbf{p}$  and  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})})$  is the marginal probability distribution of the measured scattering matrix. In the rest of this paper, it is assumed for simplicity that all defect parameters from the defect class and all possible measured scattering matrices are equally probable. In other words, probability distributions  $\rho_p(\mathbf{p})$  and  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})})$  are uniform. In this case

$$\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})}) = C \rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p}), \quad (4.3)$$

where the normalization constant  $C$  is given by integration of the conditional probability over the parameter space

$$C = \left( \int \rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p}) d\mathbf{p} \right)^{-1}. \quad (4.4)$$

Note that in a given industrial inspection some additional information might be available about possible defect types and defect parameters. This information can be naturally incorporated into the characterization procedure by using non-uniform marginal probability density distributions of the defect parameters. Once the probability density function is estimated, the characterization result is given by the defect parameters  $\mathbf{p}_c$  with the highest probability  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})$ .

However, practically it is more convenient to calculate the probability density function  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p})$  at each defect parameter point  $\mathbf{p}$  first, and then calculate the function  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})$  using expressions (4.3) and (4.4). The conditional probability density function  $\rho_S(\mathbf{s}^{(\text{pc})} | \mathbf{p})$  as a function of the vector  $\mathbf{s}^{(\text{pc})}$  describes the distribution of noise in pc-space. This function can be estimated from an experiment, or calculated using some assumed noise model. In any case, the defect characterization result and the corresponding confidence level depend on the choice of the noise model. Therefore, the knowledge of  $\rho_S(\mathbf{s}^{(\text{pc})} | \mathbf{p})$  is critical for defect characterization.

Generally, the maximum probability point  $\mathbf{p}_c$  has to be estimated numerically. However, one important case when this point can be easily found is for isotropic noise distribution in pc-space. This means that the probability density function  $\rho_S(\mathbf{s}^{(\text{pc})} | \mathbf{p})$  depends on the distance between the experimental (noisy) scattering matrix  $\mathbf{s}^{(\text{pc})}$  and the scattering matrix  $\mathbf{s}_p^{(\text{pc})}$  of the defect with parameter  $\mathbf{p}$  only. If in addition the probability density function  $\rho_S(\mathbf{s}^{(\text{pc})} | \mathbf{p})$  is independent of the parameter point  $\mathbf{p}$ , then

$$\rho_S(\mathbf{s}^{(\text{pc})} | \mathbf{p}) = h(|\mathbf{s}^{(\text{pc})} - \mathbf{s}_p^{(\text{pc})}|). \quad (4.5)$$

The most probable defect parameter  $\mathbf{p}_c$  corresponds to the point  $\mathbf{s}_{p_c}^{(pc)}$  on the  $d$ -manifold. If the function  $h(r)$  monotonically decays as  $r \rightarrow \infty$ , then the point  $\mathbf{s}_{p_c}^{(pc)}$  corresponds to the nearest point or simply the projection of the testing point  $\mathbf{s}_{\text{exp}}^{(pc)}$  onto the  $d$ -manifold. It is possible to derive a semi-analytic expression for the projection point on the  $d$ -manifold and the correspond result is given in appendix Aa. Finally, the characterization parameter vector  $\mathbf{p}_c$  can be found by mapping the point  $\mathbf{s}_{p_c}^{(pc)}$  back into defect parameter space.

## 5. General model of coherent noise

It has been noted earlier that the conditional probability density function  $\rho_S(\mathbf{s}^{(pc)} | \mathbf{p})$  describes the distribution of noise in pc-space. This function defines the defect characterization result and, therefore, plays a central part in the defect characterization procedure. Practically, the noise distribution can be estimated from experimental testing under realistic measurement conditions. However, this approach requires many samples with real defects in order to measure the statistics of the noise. Alternatively, the function  $\rho_S(\mathbf{s}^{(pc)} | \mathbf{p})$  can be simulated using some assumed noise model. The difficulty is that, in general, there are many different factors which contribute to the noise. Detailed analysis of the different noise models is beyond the scope of this paper and will be performed in a separate publication. Here, we just note that usually random noise can be suppressed by averaging and also the PCA acts as an additional filter [12]. Therefore, in practice, the coherent noise makes the main contribution to the total noise level. In this section, a general coherent noise model is proposed.

Figure 6a shows the experimentally measured absolute value of the scattering matrix for a surface-breaking crack of  $1.13\lambda$  length and  $45^\circ$  orientation. The experimental set-up is the same as described in §2 (figure 1). The difference between the measured scattering matrix and the simulated scattering matrix for the crack with the same parameters (noise part of the scattering matrix) is shown in figure 6b. It can be seen that the noise is coherent, different components of the noise scattering matrix are not independent and are seen to be correlated with each other. Moreover, in this particular case the structure of the noise scattering matrix is very similar to the structure of the scattering matrix itself. Note that here we do not specifically comment on the source of the noise and attempt to develop a general model.

The noise part of the scattering matrix can be considered as a rough two-dimensional surface and, hence, various models of surface roughness can be applied to model the coherent noise. In this case, the amplitude of each component of the noise scattering matrix is described by some probability density function, and the correlation between different components is described by a correlation function. In this paper, a Gaussian model of correlated noise is chosen.

Note that generally the noise affects both the amplitude and phase of the scattering matrix, or its real and imaginary parts. The experimental complex-valued scattering matrix  $S_{\text{exp}}(\theta_{\text{in}}, \theta_{\text{sc}})$  can be represented in the form

$$S_{\text{exp}} = S_p + S_{\text{coh},1} + iS_{\text{coh},2}, \quad (5.1)$$

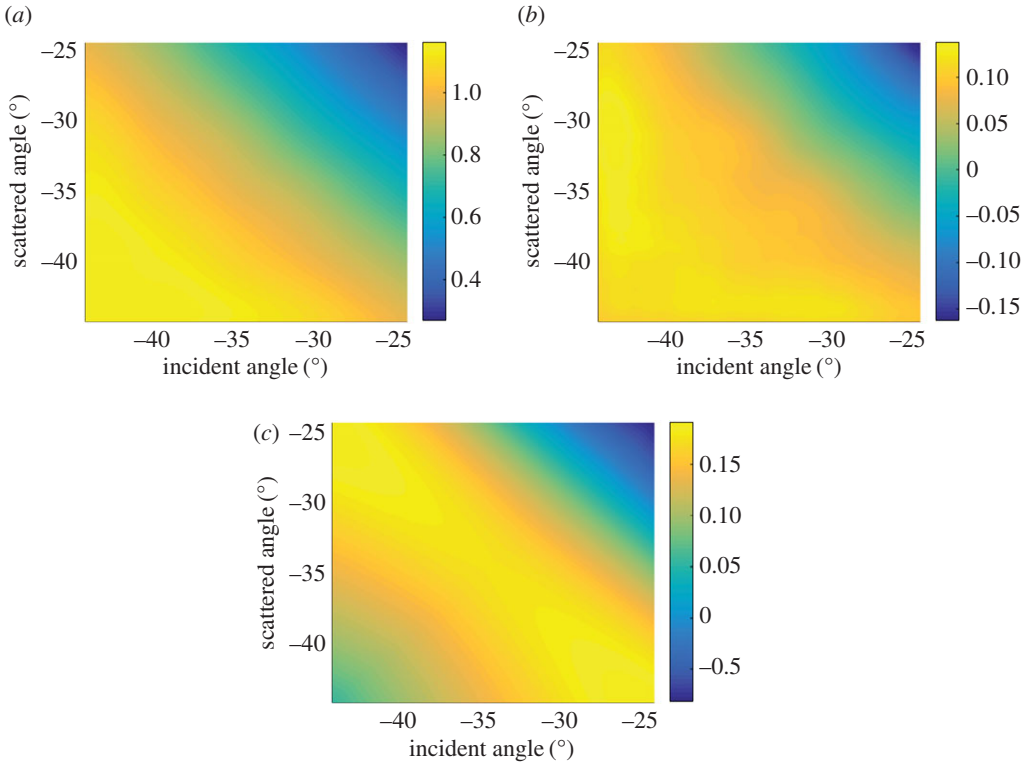
where  $S_p$  is the noiseless scattering matrix. The real functions  $S_{\text{coh},n}$ ,  $n = 1, 2$ , are now statistically described by the same probability density function  $\rho_{\text{coh}}(S)$ . For simplicity, it is assumed that the mean value of each component of the noise scattering matrix is zero. This also implies that the average value of the noise scattering matrix is zero, so

$$\langle S_{\text{coh},n} \rangle = 0, \quad n = 1, 2, \quad (5.2)$$

where the angled brackets denote averaging over incident and scattered angles.

The correlation function represents the extent to which the noise at one point  $(\theta_{\text{in}}, \theta_{\text{sc}})$  determines the noise at some point  $(\zeta_1, \zeta_2)$  away, and is defined as

$$C(\zeta_1, \zeta_2) = \frac{1}{\sigma_{\text{coh}}^2} \langle S_{\text{coh},n}(\theta_{\text{in}}, \theta_{\text{sc}}) S_{\text{coh},n}(\theta_{\text{in}} + \zeta_1, \theta_{\text{sc}} + \zeta_2) \rangle. \quad (5.3)$$



**Figure 6.** Scattering matrix of a surface-breaking crack with parameters  $a_{\text{crack}} = 1.13\lambda$ ,  $\theta_{\text{crack}} = 45^\circ$ : (a) measured scattering matrix; (b) noise part of the scattering matrix; (c) one realization of Gaussian coherent noise with parameters  $\sigma_{\text{coh}} = 0.1$ ,  $\lambda_{1\text{coh}} = 30^\circ$ ,  $\lambda_{2\text{coh}} = 10^\circ$  and  $\phi_{\text{coh}} = -45^\circ$ . (Online version in colour.)

To simplify analytical manipulations, both the noise amplitude distribution and correlation functions are regarded as Gaussian. In this case, the probability density function  $\rho_{\text{coh}} \sim \mathcal{N}(0, \sigma_{\text{coh}}^2)$ , where  $\sigma_{\text{coh}}$  is the standard deviation of noise amplitude. The two-dimensional correlation function  $C(\xi_1, \xi_2)$  can be written in the form

$$C(\xi_1, \xi_2) = \exp\left(-\frac{\tilde{\xi}_1^2}{\lambda_{1\text{coh}}^2} - \frac{\tilde{\xi}_2^2}{\lambda_{2\text{coh}}^2}\right) \quad (5.4)$$

and

$$\begin{pmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \end{pmatrix} = \mathbf{R} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} \cos \phi_{\text{coh}} & \sin \phi_{\text{coh}} \\ -\sin \phi_{\text{coh}} & \cos \phi_{\text{coh}} \end{pmatrix}. \quad (5.5)$$

Here, the isolines of the correlation function represent ellipses with the axes  $\lambda_{1\text{coh}}$ ,  $\lambda_{2\text{coh}}$ , rotated by the angle  $\phi_{\text{coh}}$  with respect to the coordinate system  $(\xi_1, \xi_2)$ .

A particular realization of a coherent noise can be simulated by the convolution of the correlation function, with uncorrelated unit variance and zero mean Gaussian white noise  $n_g(\theta_{\text{in}}, \theta_{\text{sc}})$

$$S_{\text{coh},n}(\theta_{\text{in}}, \theta_{\text{sc}}) = (C(\theta_{\text{in}}, \theta_{\text{sc}}) \otimes n_g(\theta_{\text{in}}, \theta_{\text{sc}}) - \mu') \frac{\sigma_{\text{coh}}}{\sigma'}, \quad (5.6)$$

where  $\otimes$  is the convolution operator, and  $\mu'$  and  $\sigma'$  are, respectively, the mean and standard deviations of the convolution  $C \otimes n_g$  over incident and scattered angles [37].

In this paper, only an amplitude of the scattering matrix,  $|S_{\text{exp}}|$ , is considered. If the noise amplitude is small,  $\sigma_{\text{coh}} \ll |S_p|$ , then the amplitude of the scattering matrix can be approximately written as

$$|S_{\text{exp}}| = |S_p| + S_{\text{coh}}, \quad S_{\text{coh}} = \frac{\Re[S_p]}{|S_p|} S_{1\text{coh}} + \frac{\Im[S_p]}{|S_p|} S_{2\text{coh}}. \quad (5.7)$$

Because the sum of two independent normally distributed variables also has a normal distribution, then from (5.7) it follows that  $S_{\text{coh}}$  is also described by the model (5.6).

Figure 6c shows one realization of the coherent noise for the parameters  $\sigma_{\text{coh}} = 0.1$ ,  $\lambda_{1\text{coh}} = 30^\circ$ ,  $\lambda_{2\text{coh}} = 10^\circ$  and  $\phi_{\text{coh}} = -45^\circ$ . It can be seen that the structure of the simulated noise scattering matrix is very similar to the experimental noise scattering matrix shown in figure 6b.

In pc-space, each scattering matrix is represented by an  $N_s$ -dimensional vector, so expression (5.7) can be written in a form similar to expression (4.1)

$$\mathbf{s}_{\text{exp}}^{(\text{pc})} = \mathbf{s}_p^{(\text{pc})} + \mathbf{n}_{\text{coh}}^{(\text{pc})}, \quad (5.8)$$

where the noise vector  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$  corresponds to the noise scattering matrix  $S_{\text{coh}}$ . Below, it is assumed for simplicity that the noise is independent of the defect parameter point and described by the same parameters  $\sigma_{\text{coh}}$ ,  $\lambda_{1\text{coh}}$ ,  $\lambda_{2\text{coh}}$  and  $\phi_{\text{coh}}$  for all points on the  $d$ -manifold.

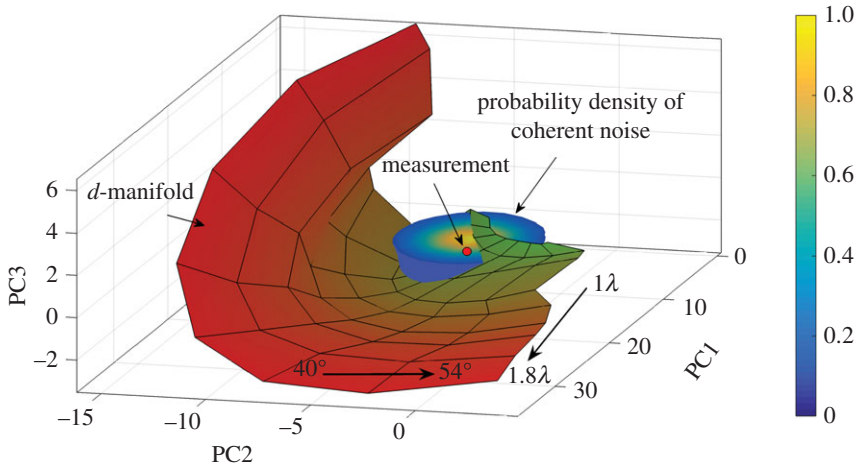
The defect characterization method requires the knowledge of the probability density function  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p})$ . Using expression (5.8) and taking into account that the noise does not depend on the defect parameter  $\mathbf{p}$ , the function  $\rho_S$  can be written as

$$\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p}) = \rho_S(\mathbf{s}_p^{(\text{pc})} + \mathbf{n}_{\text{coh}}^{(\text{pc})} | \mathbf{p}) = \rho_n(\mathbf{n}_{\text{coh}}^{(\text{pc})}), \quad (5.9)$$

where  $\rho_n$  is the probability density function of the noise vector  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$ . An important advantage of the proposed coherent noise model is that the function  $\rho_n(\mathbf{n}_{\text{coh}}^{(\text{pc})})$  can be written in an explicit form and the details are given in appendix Ab. Moreover, as shown in appendix Ab, the characterization result (defect parameters corresponding to the highest probability) can be found by the projection of the measurement point on the  $d$ -manifold in the normalized noise pc-space, which represents the rotated and scaled defect class pc-space. Or, alternatively, this projection can be performed in defect class pc-space, but using a different metric, which corresponds to the distance in the normalized noise pc-space.

To illustrate the distribution of noise in pc-space the  $d$ -manifold for surface-breaking cracks is considered. The experimental set-up is assumed to be the same as in §2 (figure 1). The noise probability density function was calculated using expression (A 10) with noise parameters  $\sigma_{\text{coh}} = 0.1$ ,  $\lambda_{\text{coh},1} = 30^\circ$ ,  $\lambda_{\text{coh},2} = 10^\circ$  and  $\phi_{\text{coh}} = -45^\circ$ . Figure 7 shows the noise distribution around the point on the  $d$ -manifold corresponding to the crack parameters  $a_{\text{crack}} = 1.13\lambda$  and  $\theta_{\text{crack}} = 45^\circ$ . It can be seen that the noise distribution in the defect class pc-space has the shape of an ellipsoid which is rotated relative to the defect class principal components coordinate system. Then, the position of the measurement point  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$  inside this ellipsoid defines the probability density function  $\rho_S(\mathbf{s}_{\text{exp}}^{(\text{pc})} | \mathbf{p})$ ,  $\mathbf{p} = \{a_{\text{crack}}, \theta_{\text{crack}}\}$ , and hence the probability that the measurement corresponds to a particular set of defect parameters.

Note that the Bayesian statistical approach described in §4 is similar to the method previously used in [15] for characterization of corrosion defects in plates. However, the method proposed in this paper has several major differences. Firstly, the characterization is performed in pc-space, which provides a simple geometrical interpretation of the characterization process and also allows us to filter random experimental noise from the measurements [12]. Secondly, the fact that the characterization uncertainty defined by the noise model is highlighted and a general model of the coherent noise is developed.



**Figure 7.** Probability density function (normalized to its maximum value) of Gaussian coherent noise  $\rho_S(\mathbf{s}^{(pc)} | \mathbf{p})$  in the pc-space of surface-breaking cracks at the crack's parameter point  $a_{\text{crack}} = 1.13\lambda$ ,  $\theta_{\text{crack}} = 45^\circ$ . Noise parameters are  $\sigma_{\text{coh}} = 0.1$ ,  $\lambda_{\text{coh},1} = 30^\circ$ ,  $\lambda_{\text{coh},2} = 10^\circ$  and  $\phi_{\text{coh}} = -45^\circ$ . Colourmap of the manifold is given by figure 4a; colourbar is related to the probability density function.

## 6. Experiments

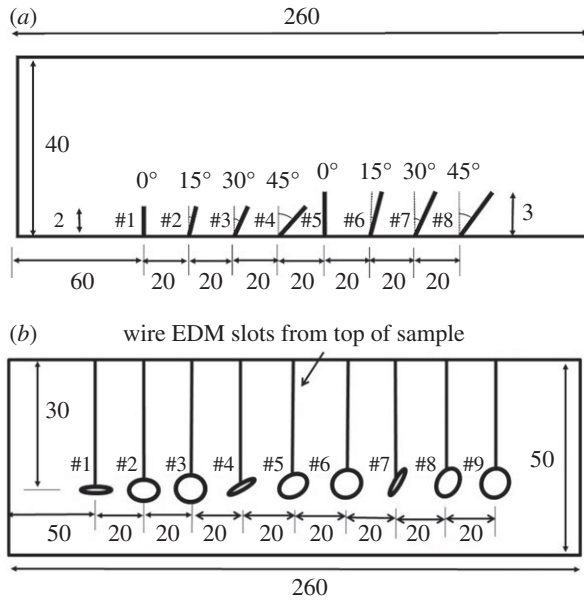
### (a) Experimental characterization procedure

Experimentally, the performance of the characterization method is studied on two different defect types: surface-breaking cracks and elliptical voids (see figure 8 for the sample geometries). As discussed earlier, both defects are of particular industrial interest, and they have different numbers of parameters of interest (two for surface-breaking cracks and three for elliptical voids). The true defect parameters are given in tables 1 and 2. In both cases, the  $s$ -spaces include the scattering matrices of defects with sizes between  $0.5\lambda$  and  $2\lambda$ , as the scattering matrix is most informative within this size range [11]. For defects larger than  $2\lambda$ , their characterization should be possible directly from high-resolution ultrasonic images, such as the ones obtained with the total focusing method (TFM) [38]. The considered angle range of surface-breaking cracks is from  $-60^\circ$  to  $60^\circ$ . For elliptical voids, aspect ratios between 0.1 and 0.9 and ellipse orientation angles between  $-90^\circ$  and  $90^\circ$  are considered. Note that the orientation angle of elliptical voids is measured with respect to the array direction, and the angle of a horizontal ellipse is  $0^\circ$ . Therefore, orientation angle of the crack,  $\theta_{\text{crack}}$ , is related to the orientation angle of the void,  $\theta_{\text{void}}$ , by

$$\theta_{\text{crack}} + \theta_{\text{void}} = 90^\circ. \quad (6.1)$$

In this paper, the sub-array imaging technique [39] is used to extract the scattering matrices from experimental data. In this case, the amplitude of the scattering matrix also needs to be normalized by comparing with a reference scatterer. Here, the back wall of the specimen was used as a reference scatterer. The corresponding array data can be simulated using the hybrid model introduced in [40]. To extract the absolute amplitude of the scattering matrix correctly, the experimental array data are multiplied by a constant gain, which makes the maximum TFM image amplitude of the reference scatterer for experimental and simulated data equal.

Note that the sub-array imaging technique introduces a specific coherent noise to the extracted scattering matrices. In particular, the value of the scattering matrix at any incident and scattered angles is represented by an averaging of the array data from several consecutive array elements (i.e. sub-array). This results in a 'smoothing' effect relative to the true scattering matrix [39].



**Figure 8.** (Dimensions are in millimeters.) Experimental sample geometries containing: (a) electrical discharge machining (EDM) notches; (b) elliptical voids.

**Table 1.** Parameters of the machined notches shown in figure 8a.

defect	parameters	
	size ( $\lambda$ )	angle ( $^\circ$ )
1	0.80	0
2	0.83	15
3	0.92	30
4	1.13	45
5	1.20	0
6	1.24	15
7	1.39	30
8	1.70	45

Since this is a deterministic effect, it can be resolved by constructing a ‘sub-array version’ of the scattering matrix in  $s$ -space. The sub-array scattering matrix,  $S_{sa}$ , can be obtained by

$$S_{sa}(\theta_{in,k}^a, \theta_{sc,l}^a) = \frac{1}{N_{sa}^2} \sum_{i \in a_k, j \in a_l} S(\theta_{in,i}, \theta_{sc,j}), \quad (6.2)$$

where  $a_k$  denotes the  $k$ th sub-array aperture with the corresponding incident,  $\theta_{in,k}^a$ , and scattered,  $\theta_{sc,k}^a$ , angles. The value of the scattering matrix  $S(\theta_{in,i}, \theta_{sc,j})$  corresponds to the  $i$ th transmitter element and the  $j$ th receiver element in the full array aperture.

To calculate the probability density function  $\rho_p(\mathbf{p} | \mathbf{s}_{exp}^{(pc)})$  of defect parameters the Gaussian coherent noise model developed in §5 was used. The noise parameters  $\sigma_{coh}$ ,  $\lambda_{coh,1}$ ,  $\lambda_{coh,2}$  and  $\phi_{coh}$  were defined from comparison of the measured scattering matrices with modelled scattering

**Table 2.** Parameters of the elliptical voids shown in figure 8*b*. AR represents the aspect ratio.

defect	parameters		
	AR	size ( $\lambda$ )	angle ( $^\circ$ )
1	0.32	1.00	0
2	0.50	1.00	0
3	0.80	1.00	0
4	0.32	1.00	30
5	0.50	1.00	30
6	0.80	1.00	30
7	0.32	1.00	60
8	0.50	1.00	60
9	0.80	1.00	60

matrices of the defects with ideal geometries using the maximum-likelihood estimation method. In this case, expression (A 8) for the noise probability density function in the defect class pc-space was used. For the case of surface-breaking cracks, the maximum of the likelihood function is achieved for the noise parameters  $\sigma_{\text{coh}} = 0.1$ ,  $\lambda_{\text{coh},1} = 30^\circ$ ,  $\lambda_{\text{coh},2} = 10^\circ$ ,  $\phi_{\text{coh}} = 45^\circ$ . For the case of elliptical voids, the noise parameters were estimated as  $\sigma_{\text{coh}} = 0.1$ ,  $\lambda_{\text{coh},1} = 20^\circ$ ,  $\lambda_{\text{coh},2} = 10^\circ$ ,  $\phi_{\text{coh}} = 45^\circ$ .

For each defect the probability density map corresponding to each defect class (surface-breaking cracks and elliptical voids) was calculated using expressions (4.3), (4.4), (5.9) and (A 8). Note that if the  $d$ -manifold for the defect class is defined by some sampling scheme, then any point in the parameter space can be mapped into the principal components pc-space. Practically, this means that, although the sampling of the  $d$ -manifold is finite, the probability density map in the parameter space can be calculated with any resolution.

It should be stressed that the developed defect characterization approach does not give just a single answer. The method provides quantitative information about possible defect parameters and, more importantly, about the confidence level of the characterization result. The main output of the characterization procedure is the probability density map of the defect parameters,  $\rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})})$ . The characterization result is then given by the most probable parameters,  $\mathbf{p}_c$ , where the probability density function has the maximum value.

The structure of the probability function in the defect parameter space can be characterized by the square root of the second moment of the probability distribution about the characterization point  $\mathbf{p}_c$

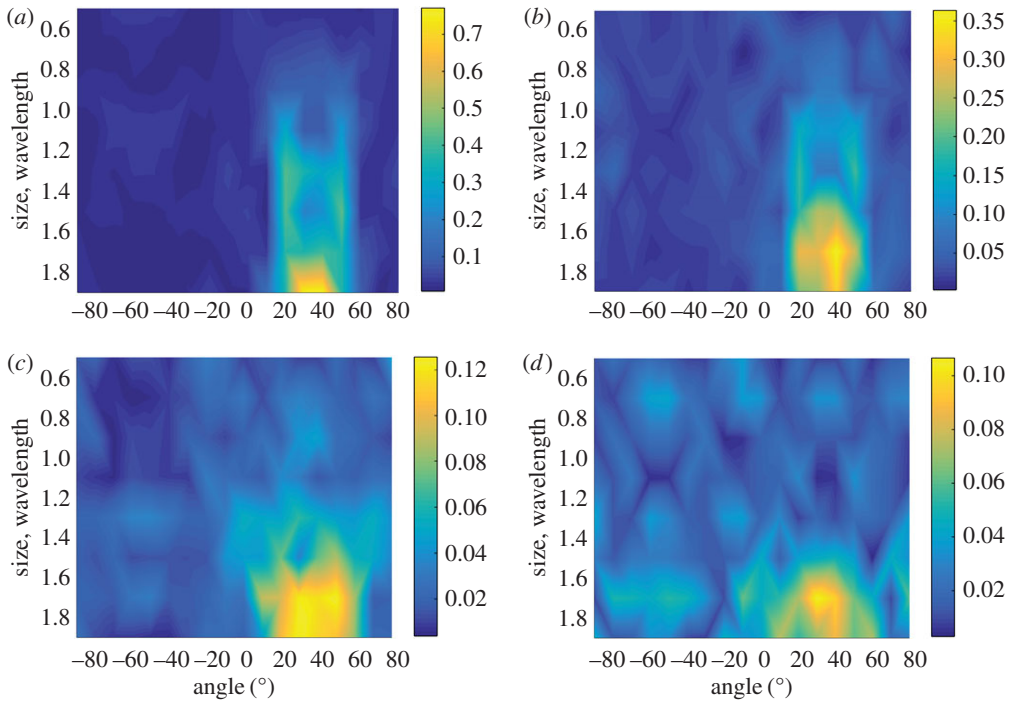
$$\sigma_i^{(p)} = \sqrt{\int (p_i - p_{c,i})^2 \rho_p(\mathbf{p} | \mathbf{s}_{\text{exp}}^{(\text{pc})}) dp}, \quad (6.3)$$

where the index  $i$  denotes the  $i$ th component of the vectors  $\mathbf{p}$  and  $\mathbf{p}_c$ , and the integral is calculated over all parameter space. The physical meaning of the value  $\sigma_i^{(p)}$  is a standard deviation of the  $i$ th defect parameter about the characterization point.

The relative noise amplitude, or the relative error of the characterization result, can be defined as

$$\tilde{a}_{\text{noise}} = \frac{|\mathbf{s}_{\text{exp}}^{(\text{pc})} - \mathbf{s}_c^{(\text{pc})}|}{|\mathbf{s}_c^{(\text{pc})}|}, \quad (6.4)$$

where the vector  $\mathbf{s}_c^{(\text{pc})}$  corresponds to the characterization point on the  $d$ -manifold in the pc-space. Note that the relative noise amplitude can also be used to compare the characterization results in two defect classes.



**Figure 9.** Map of the  $d$ -index for the elliptical voids corresponding to the  $-45^\circ \leq \theta_{in}, \theta_{sc} \leq -23^\circ$  part of the scattering matrix. Each figure represents a slice at constant aspect ratio equal to (a) 0.1, (b) 0.3, (c) 0.6, (d) 0.9. (Online version in colour.)

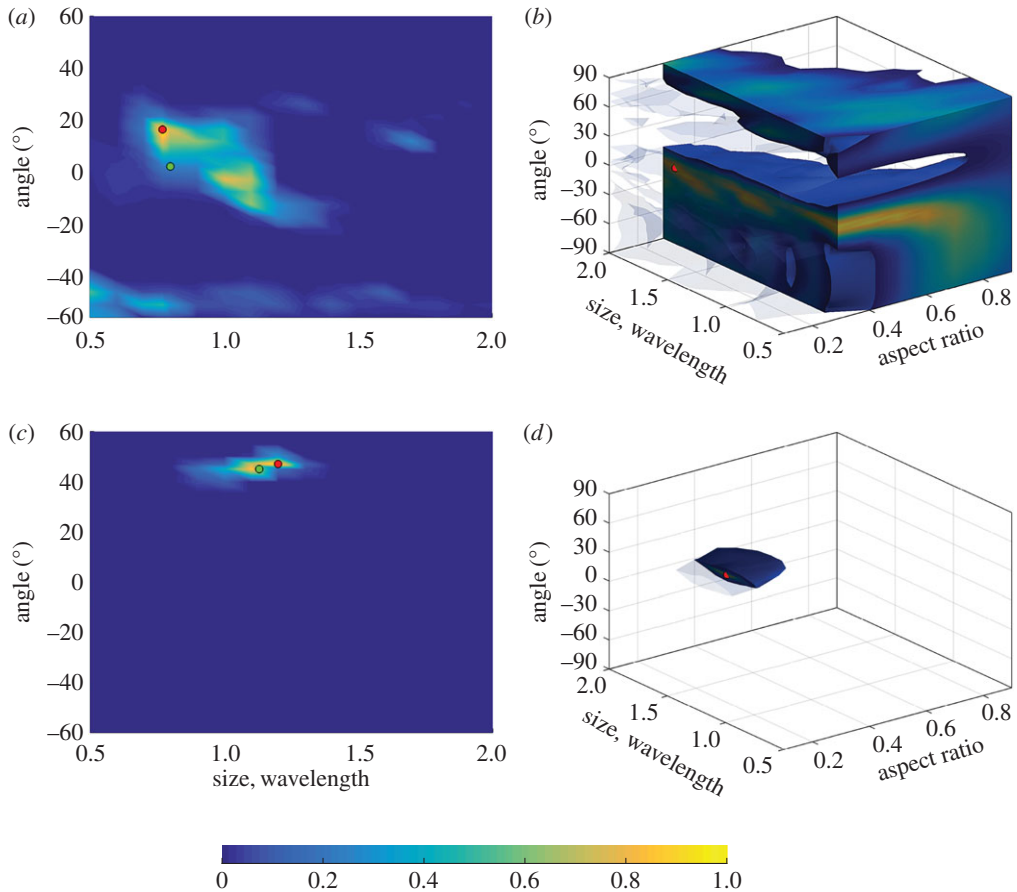
Another important characteristic of the uncertainty of the characterization result is the  $d$ -index of the defect class at the characterization point. The  $d$ -index is defined by expression (3.9) and describes the sensitivity of the characterization result to the noise. For each defect class, the  $d$ -index characterizes the structure of the defect class  $d$ -manifold and depends on the amount of information in the measured part of the scattering matrix. Application of these parameters for interpretation of the characterization results is illustrated in the next section.

## (b) Results for surface-breaking cracks

The aluminium test specimen shown in figure 8*a* contains eight EDM notches. The vertical depth of defects 1–4 is 2 mm, and defects 5–8 have the same vertical depth of 3 mm. The actual size of the defects can be calculated as  $d/\cos(\alpha)$ , where  $d$  and  $\alpha$  represent the vertical depth and the orientation angle, respectively. As a result, the studied defects have sizes between  $0.8\lambda$  (defect 1) and  $1.7\lambda$  (defect 8), and angles between  $0^\circ$  and  $45^\circ$ . The array measurement configuration is shown in figure 1. Note that the array is moved away from the crack by 30 mm, because otherwise the crack would be obscured by the high-intensity reflection from the back wall in the image. Therefore, the scattering matrix for each crack was measured for the  $-45^\circ \leq \theta_{in}, \theta_{sc} \leq -23^\circ$  angular interval.

The  $d$ -indexes for the surface-breaking cracks and elliptical voids corresponding to the measurement angular interval are shown in figures 5*b* and 9, respectively. Taking into account the relationship (6.1) between crack and void orientation angles, it can be seen that, in both cases, the most stable parameter regions correspond to the similar defect geometries:  $1\lambda \leq a_{crack} \leq 2\lambda$ ,  $40^\circ \leq \theta_{crack} \leq 60^\circ$  for the cracks and  $1\lambda \leq a_{void} \leq 2\lambda$ ,  $30^\circ \leq \theta_{void} \leq 50^\circ$  for narrow elliptical voids with aspect ratio between 0.1 and 0.3. Note that for these defect orientation angles the measured part of the scattering matrix contains the specular reflection and, therefore, it seems reasonable that this is the most informative part from the defect characterization point of view.





**Figure 10.** Probability maps of the machined notches shown in figure 8a for defects 1 (a,b) and 4 (c,d), characterized in the surface-breaking crack  $s$ -space (a,c) and in the void  $s$ -space (b,d). The red dots indicate the maximum probability point; the green dots are the true parameter values. The probabilities shown here are normalized to the maximum value.

The characterization results of the machined notches in the surface-breaking crack defect class are summarized in table 3. The corresponding probability density functions of defect parameters are illustrated in figure 10 for defects 1 and 4.

Firstly, it is noted that, for all defects, the correct defect parameters are located in the range of the probability density function (see figure 10 as an example). Therefore, true parameters represent a possible characterization point in the parameter space, although this point does not always correspond to the highest probability.

From table 3, it can be seen that the standard deviations of defect parameters about the characterization point are well correlated with the values of the  $d$ -index. This confirms that the confidence level of the characterization results is defined by the global structure of the  $d$ -manifold. According to the discussion on the  $d$ -index, defects 4 and 8 are favourably oriented, i.e. have high  $d$ -index values. Consequently, the characterization results show that these two defects are correctly characterized with high confidence (figure 10c for defect 4). For the defects 1–3, 5 and 6, the  $d$ -indexes are low. This is in agreement with probability maps (figure 10a for defect 1), which show that, for these defects, there exist multiple regions of high probability in the defect parameter spaces and, therefore, the characterization uncertainty is large.

Now, the characterization of crack defects in the elliptical void defect class is considered. The corresponding results are given in table 4. The probability density in this case is a function of three parameters and is shown in figure 10b,d for defects 1 and 4. The characterization

**Table 3.** Characterization results of the machined notches shown in figure 8a in the surface-breaking crack defect class.  $\sigma_l^{(p)}$  and  $\sigma_\theta^{(p)}$  correspond to the standard deviation of defect size and angle about the characterization point.

characterization results, cracks in crack database						
defect	size ( $\lambda$ )	angle ( $^\circ$ )	$d$ -index	$\sigma_l^{(p)}$ ( $\lambda$ )	$\sigma_\theta^{(p)}$ ( $^\circ$ )	$\tilde{a}_{\text{noise}}$
1	0.77	14.88	0.024	0.32	33.92	0.033
2	1.10	−15.00	0.035	0.28	28.75	0.080
3	0.97	27.87	0.059	0.35	46.96	0.078
4	1.21	47.03	0.150	0.14	2.45	0.063
5	1.31	15.00	0.033	0.27	19.30	0.480
6	1.79	10.00	0.035	0.82	21.44	0.250
7	1.44	30.33	0.040	0.04	1.40	0.090
8	1.70	46.71	0.230	0.06	0.11	0.140

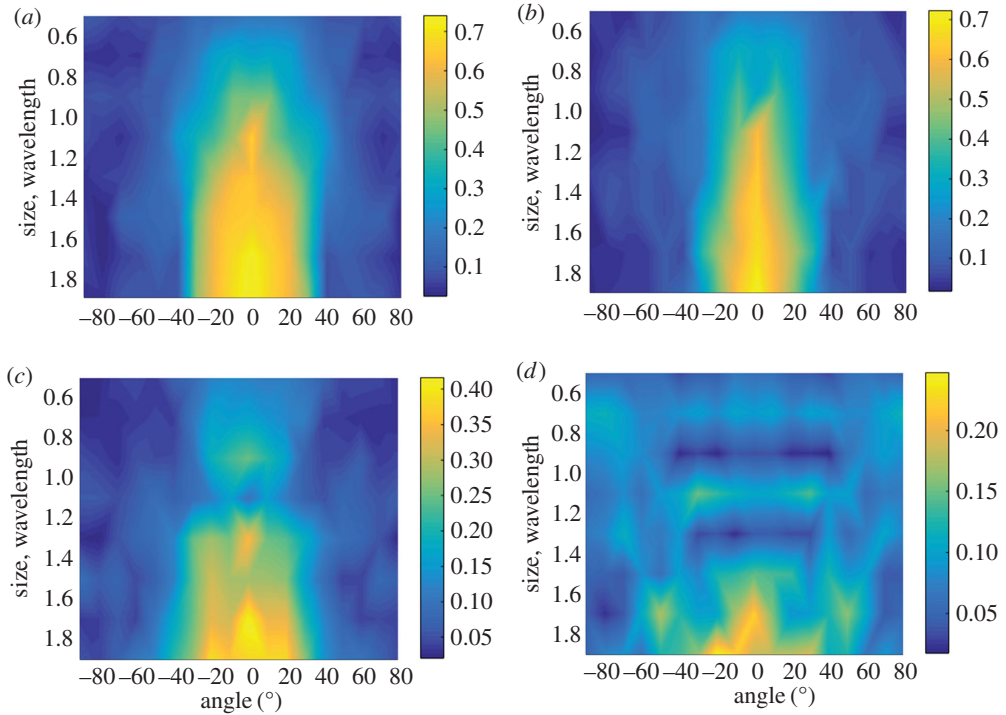
**Table 4.** Characterization results of the machined notches shown in figure 8a in the void defect class. AR represents the aspect ratio;  $\sigma_{\text{AR}}^{(p)}$ ,  $\sigma_l^{(p)}$  and  $\sigma_\theta^{(p)}$  correspond to the standard deviation of the defect aspect ratio, size and angle about the characterization point, respectively.

characterization results, cracks in void database								
defect	AR	size ( $\lambda$ )	angle ( $^\circ$ )	$d$ -index	$\sigma_{\text{AR}}^{(p)}$	$\sigma_l^{(p)}$ ( $\lambda$ )	$\sigma_\theta^{(p)}$ ( $^\circ$ )	$\tilde{a}_{\text{noise}}$
1	0.29	1.89	−11.72	0.013	0.260	0.88	45.66	0
2	0.25	1.10	5.68	0.010	0.260	0.43	56.29	0
3	0.21	1.25	56.96	0.140	0.210	0.39	45.58	0
4	0.19	1.21	42.68	0.180	0.070	0.13	1.76	0.03
5	0.11	1.85	53.46	0.230	0.460	0.72	66.00	0.39
6	0.16	1.77	55.83	0.210	0.490	0.73	79.47	0.40
7	0.27	1.08	7.29	0.014	0.260	0.43	50.34	0.44
8	0.10	1.86	43.13	0.610	0.009	0.07	0.14	0.09

results for different defect classes can be compared using the relative noise amplitudes of the characterization results,  $\tilde{a}_{\text{noise}}$ , and the characterization defect class is defined as the class with the minimum value of  $\tilde{a}_{\text{noise}}$ . Note that  $\tilde{a}_{\text{noise}} = 0$  for defects 1–3. Geometrically, it means that the corresponding experimental point  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$  is exactly located on the  $d$ -manifold.

From table 4, it can be seen that only defects 6 and 7 are classified as surface-breaking cracks. According to the minimum relative noise amplitude criterion, all other defects should be classified as voids. However, it can be seen that, in the elliptical void class, all surface-breaking cracks are characterized as narrow ellipses with small aspect ratio (e.g. 0.2). Also, the results for defects 1–3, 5 and 6 have very large uncertainty in both defect classes, confirming that, for these defects, the amount of information in the measured part of the scattering matrix is not enough for unique defect characterization.

On the other hand, the characterization results of defects 4 and 8 in the void class have similar sizes and orientation angles to the true crack parameters. So, in this case, the characterization result in the void defect class brings confidence, confirming that the defects are crack-like and the method is providing correct defect parameters. Physically, this means that, for these



**Figure 11.** Map of the  $d$ -index for the elliptical voids corresponding to the  $-30^\circ \leq \theta_{in}, \theta_{sc} \leq 30^\circ$  part of the scattering matrix. Each figure represents a slice at constant aspect ratio equal to (a) 0.1, (b) 0.3, (c) 0.6, (d) 0.9. (Online version in colour.)

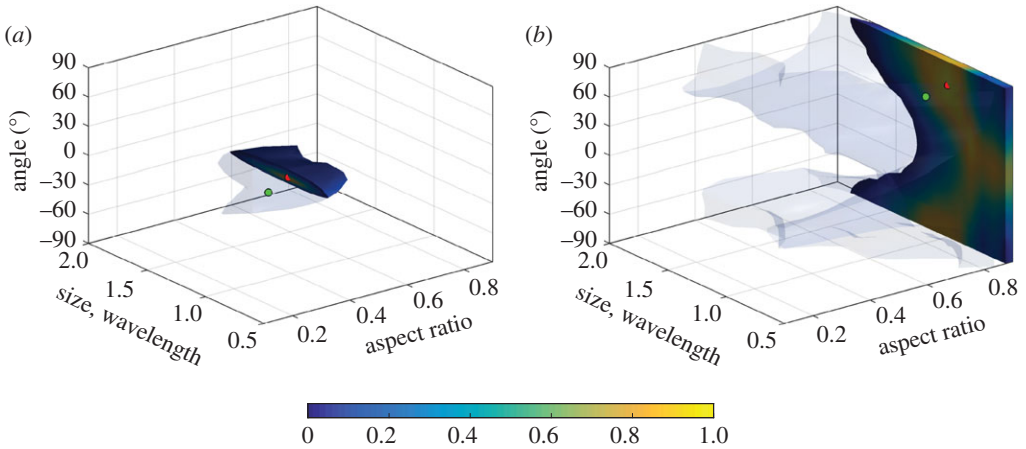
defects, the influence of the back wall on the scattering behaviour in the considered angular measurement range is small. Consequently, the measured parts of the scattering matrices for surface-breaking cracks and isolated cracks are very similar to each other. Note that, in practice, a more sophisticated model of defect class prior probability  $\rho_p(\mathbf{p})$  can be adopted based on specific measurement scenarios. For example, if a defect is found near the surface, then it is more likely to be a surface-breaking crack than a void.

### (c) Results for elliptical voids

The array measurements of the specimen containing elliptical voids were performed by positioning the array exactly above each defect. For this experimental configuration the measured part of the scattering matrix corresponds to the angles  $-30^\circ \leq \theta_{in}, \theta_{sc} \leq 30^\circ$ . As in the previous section, all defects were characterized in the two defect classes: elliptical voids and surface-breaking cracks. However, for all elliptical defects the relative noise amplitude was smaller for the characterization results corresponding to the void defect class. Therefore, all defects were characterized as voids, and, below, the results only for the void defect class are discussed.

The  $d$ -index for the elliptical void defect class is shown in figure 11. From this figure, it follows that the defect parameters can be estimated with relatively high confidence for ellipses with aspect ratios between 0.1 and 0.6, sizes between  $1\lambda$  and  $2\lambda$ , and orientation angles between  $-30^\circ$  and  $30^\circ$ . Note that, for these defect orientation angles, the measured part of the scattering matrix contains specular reflection.

The characterization results of the elliptical voids are summarized in table 5 and the probability density functions for defects 1 and 9 are shown in figure 12. Note that, similar to the surface-breaking crack characterization, all true defect parameters are located in the range of the corresponding probability density functions in the defect parameter space. It is seen that the confidence level of the characterization is correlated with the values of the  $d$ -index. For example,



**Figure 12.** Probability maps of the elliptical voids 1 (a) and 9 (b) (figure 8b) characterized in the void  $s$ -space. The slice through the three-dimensional probability map at the constant aspect ratio, corresponding to the maximum probability point, is shown. The red dots indicate the maximum probability point and the green dots are the true parameter values. The probabilities shown here are normalized to the maximum value.

**Table 5.** Characterization results of the elliptical voids shown in figure 8b in the void defect class. AR represents the aspect ratio;  $\sigma_{AR}^{(p)}$ ,  $\sigma_l^{(p)}$  and  $\sigma_\theta^{(p)}$  correspond to the standard deviation of the defect aspect ratio, size and angle about the characterization point, respectively.

characterization results, voids in void database								
defect	AR	size ( $\lambda$ )	angle ( $^\circ$ )	$d$ -index	$\sigma_{AR}^{(p)}$	$\sigma_l^{(p)}$ ( $\lambda$ )	$\sigma_\theta^{(p)}$ ( $^\circ$ )	$\tilde{a}_{noise}$
1	0.43	1.09	1.56	0.53	0.08	0.20	2.75	0.03
2	0.51	1.11	4.00	0.38	0.17	0.38	13.18	0.06
3	0.77	1.17	8.00	0.12	0.10	0.40	41.23	0.02
4	0.34	1.80	36.81	0.16	0.12	0.64	8.26	0.21
5	0.54	1.05	24.83	0.12	0.15	0.47	21.83	0.15
6	0.78	0.97	14.38	0.06	0.10	0.51	46.94	0.02
7	0.42	0.95	60.65	0.09	0.13	0.46	29.17	0.02
8	0.61	0.72	51.47	0.04	0.15	0.56	30.74	0.06
9	0.87	0.99	66.00	0.08	0.17	0.44	42.27	0.25

the characterization result of defect 1 has the highest  $d$ -index, and this defect is correctly characterized with high confidence. The corresponding probability density function in figure 12a is focused around the true parameter point in the defect parameter space.

On the other hand, the value of the  $d$ -index for defect 9 is very low, which results in the high characterization uncertainty seen in figure 12b. Note that the scattering matrix of voids with an aspect ratio close to 1 is insensitive to defect orientation angle, so low values of the  $d$ -index in this case correspond to uncertainty with respect to the orientation angle. It is also seen that, although the defect size is correctly characterized, its uncertainty is also large. However, the aspect ratio is estimated correctly and the corresponding standard deviation is small.

## 7. Conclusion

A new defect characterization approach which can potentially be applied to any defect databases (on condition that the defects are parameterized with several continuous variables) has been proposed in this paper. In this method, the defect database has the form of a parametric manifold, and can be approximated to any resolution from a finite number of samples. The characterization problem can be formulated as calculating the posterior probability  $\rho_p(\mathbf{p} | S_{\text{exp}})$  for every possible parameter value  $\mathbf{p}$  given the occurrence of the measured scattering matrix  $S_{\text{exp}}$ . The result is dependent on the specific noise model which is described by a probability density function  $\rho_S(S_{\text{exp}} | \mathbf{p})$  of the measured scattering matrix given the occurrence of the defect with parameters  $\mathbf{p}$ . A general Gaussian coherent noise model is proposed, and the analytical expression for the corresponding probability density function is derived.

The defect characterization data provided by the developed method can be represented in different forms and described by different parameters. For example, the probability distribution in the defect parameter space can be quantified by the standard deviation of defect parameters about the characterization point (see expression (6.3)). The sensitivity of the characterization result to noise can be described by the value of the  $d$ -index, which captures the structure of the  $d$ -manifold. To assess the general confidence of the characterization result and compare results obtained in the different defect classes, the relative noise amplitude (expression (6.4)) can be used, which describes how closely the experimental  $S$ -matrix matches the characterization result. Moreover, the developed defect characterization method allows us to define other quantitative characterization parameters, which can be useful for specific practical applications.

Experimentally, the proposed approach is used to characterize surface-breaking defects and elliptical voids. Note that one of the advantages of the proposed method is its ability to quantify the defect characterization uncertainty. In general, the characterization results are accurate, except for the cases where the defects are unfavourably oriented, which results in a lack of useful information in the array data. It is also concluded that the characterization results of unfavourably oriented defects tend to have higher uncertainty. Furthermore, reliable estimation of the noise parameters are required for accurate quantification of the characterization uncertainty.

**Data accessibility.** The datasets supporting this article have been uploaded as part of the electronic supplementary material.

**Authors' contributions.** A.V. and B.W.D. conceived the study. L.B. and A.V. developed the theory. L.B. led the data analysis and performed the experimental work with contributions from all authors. All authors wrote the paper.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was funded by the EPSRC under grant no. EP/L022125/1.

## A. Appendix

### (a) Sampling of defect manifold

In this section, the problem of practical implementation of the characterization procedure proposed in §4 is considered. In general, the defect  $d$ -manifold cannot be described analytically and, therefore, has to be defined by some sampling. Firstly, the parameter space is sampled and then triangulated using, for example, Delaunay triangulation [41,42]. Then, for any parameter vector  $\mathbf{p}$ , the position of the point  $S_p(\mathbf{p})$  on the  $d$ -manifold can be approximately found by using an interpolation scheme and, hence, the  $d$ -manifold can be approximated with any resolution. Inversely, any point on the  $d$ -manifold can be approximately mapped back into the parameter space. Therefore, the structure of the  $d$ -manifold is completely defined by its sampling points.

In pc-space, each  $N_p$ -dimensional triangle is given by  $N_p + 1$  sampling points  $\mathbf{s}_n^{(\text{pc})}$ ,  $n = 0, \dots, N_p$  and defines a 'facet' of the  $d$ -manifold. The local coordinate system in each 'facet' is

described by the  $N_p$  vectors  $\mathbf{w}_n = \mathbf{s}_n^{(\text{pc})} - \mathbf{s}_0^{(\text{pc})}$ ,  $n = 1, \dots, N_p$ . Then, any scattering matrix,  $\mathbf{s}^{(\text{pc})}$ , as a vector in pc-space can be represented in the form

$$\mathbf{s}^{(\text{pc})} = \mathbf{s}_0^{(\text{pc})} + \mathbf{s}_{\parallel}^{(\text{pc})} + \mathbf{s}_{\perp}^{(\text{pc})}, \quad \mathbf{s}_{\parallel}^{(\text{pc})} = \sum_{n=1}^{N_p} b_n \mathbf{w}_n. \quad (\text{A } 1)$$

Here, the vector  $\mathbf{s}_{\parallel}^{(\text{pc})}$  represents the projection of the vector  $\mathbf{s}^{(\text{pc})}$  onto the facet subspace and the vector  $\mathbf{s}_{\perp}^{(\text{pc})}$  is orthogonal to the facet subspace. Therefore, the distance from the point  $\mathbf{s}^{(\text{pc})}$  to the facet subspace is equal to  $|\mathbf{s}_{\perp}^{(\text{pc})}| = |\mathbf{s}^{(\text{pc})} - \mathbf{s}_{\parallel}^{(\text{pc})} - \mathbf{s}_0^{(\text{pc})}|$ . The coefficients  $b_n$  are the coordinates of the vector  $\mathbf{s}_{\parallel}^{(\text{pc})}$  in the facet local coordinate system. Based on (A 1), it can be shown that

$$\mathbf{b} = \mathbf{B}(\mathbf{s}^{(\text{pc})} - \mathbf{s}_0^{(\text{pc})}), \quad \mathbf{B} = \mathbf{G}^{-1} \mathbf{W}^T, \quad (\text{A } 2)$$

where  $\mathbf{b} = \{b_1, \dots, b_{N_p}\}^T$ , and the matrices  $\mathbf{G}$ ,  $\mathbf{V}$  are given by  $G_{ij} = (\mathbf{w}_i, \mathbf{w}_j)$ ,  $W_{ij} = w_{ij}$ .

In §4, it has been shown that if the noise is isotropic in pc-space, then the characterization point is given by the projection of the measured scattering matrix  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$  onto the  $d$ -manifold. Expression (A 1) shows that the problem of finding the projection is equivalent to the following optimization problem:

$$\min_{\text{all facets}} |\mathbf{s}_{\text{exp}}^{(\text{pc})} - \mathbf{s}_{\text{exp},\parallel}^{(\text{pc})} - \mathbf{s}_0^{(\text{pc})}|, \quad b_n \geq 0, \quad \sum_{n=1}^{N_p} b_n \leq 1, \quad (\text{A } 3)$$

where  $\mathbf{s}_{\text{exp}}^{(\text{pc})}$  is the vector in the pc-space corresponding to the scattering matrix  $S_{\text{exp}}$ , the coefficients  $b_n$  are calculated from (A 2) and the minimum is taken over all facets. In some cases, it is possible that the problem (A 3) does not have a solution. Then, the point on the  $d$ -manifold corresponding to the shortest distance is located on the boundary of some facet. In this case, the minimization procedure (A 3) can be iteratively applied to the boundary of each facet.

Once the projection point is found, it can be mapped back into the parameter space. If the projection facet is defined by the points  $\mathbf{s}_n^{(\text{pc})}$ ,  $n = 0, \dots, N_p$ , which correspond to the parameter vectors  $\mathbf{p}_n$ ,  $\mathbf{s}_n^{(\text{pc})} = \mathbf{s}_p^{(\text{pc})}(\mathbf{p}_n)$ , then the characterization parameter vector,  $\mathbf{p}_c$ , is given by

$$\mathbf{p}_c = \mathbf{p}_0 + \sum_{n=1}^{N_p} b_n (\mathbf{p}_n - \mathbf{p}_0). \quad (\text{A } 4)$$

## (b) Probability density function of coherent noise

In this section, it is shown that the probability density function of the Gaussian coherent noise,  $\rho_n(\mathbf{n}_{\text{coh}}^{(\text{pc})})$ , can be written in the analytical form. The first step is to apply the PCA to the set of noise vectors  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$ . The noise vector in the new noise pc-space will be denoted as  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$ , where the  $\text{pc}_n$  index specifically indicates that the noise pc-space is different from the defect class pc-space. Mathematically, the transformation of  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$  into  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  is the same as described in §3a (equations (3.3)–(3.5)). Note that the mean of  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$  is zero, therefore the transformation of  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$  into  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  is equivalent to the rotation of the coordinate system in the defect database pc-space and is given by

$$\mathbf{n}_{\text{coh}}^{(\text{pc}_n)} = \mathbf{V}_{\text{coh}}^T \mathbf{n}_{\text{coh}}^{(\text{pc})}. \quad (\text{A } 5)$$

The matrix  $\mathbf{V}_{\text{coh}}$  is obtained from the singular value decomposition of the covariance matrix  $\mathbf{R}_{\text{coh}}^{(\text{pc})}$  of vectors  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$

$$\mathbf{R}_{\text{coh}}^{(\text{pc})} = \mathbf{V}_{\text{coh}} \mathbf{D}_{\text{coh}} \mathbf{V}_{\text{coh}}^T, \quad (\text{A } 6)$$

where the diagonal matrix  $\mathbf{D}_{\text{coh}} = \text{diag}\{d_{\text{coh},1}, \dots, d_{\text{coh},N_s}\}$  contains the eigenvalues of  $\mathbf{R}_{\text{coh}}^{(\text{pc})}$ .

Because all operations are linear, then from the noise model (5.6) it follows that each component of the noise vector  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  is represented by the linear combination of independent normally distributed random variables. Moreover, it is easy to check that the covariance matrix  $\mathbf{R}_{\text{coh}}^{(\text{pc}_n)}$  of vectors  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  has diagonal form

$$\mathbf{R}_{\text{coh}}^{(\text{pc}_n)} = \frac{1}{N_{\text{coh}} - 1} \mathbf{D}_{\text{coh}}, \quad (\text{A } 7)$$

where  $N_{\text{coh}}$  is the number of noise vectors  $\mathbf{n}_{\text{coh}}^{(\text{pc})}$ . This expression shows that the components of the noise vector  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  now represent independent normally distributed random variables,  $n_{\text{coh},i}^{(\text{pc}_n)} \sim \mathcal{N}(0, d_{\text{coh},i})$ , with the variances equal to the eigenvalues  $d_{\text{coh},i}$ ,  $i = 1 \dots N_s$ .

The noise probability density function  $\rho_n(\mathbf{n}_{\text{coh}}^{(\text{pc})})$  in the defect class pc-space represents the multivariate normal distribution and, using relationship (A 5), can be written as

$$\rho_n(\mathbf{n}_{\text{coh}}^{(\text{pc})}) = \frac{1}{\sqrt{\det(2\pi \mathbf{D}_{\text{coh}})}} \exp \left[ -\frac{1}{2} \mathbf{n}_{\text{coh}}^{(\text{pc})T} \mathbf{V}_{\text{coh}} \mathbf{D}_{\text{coh}}^{-1} \mathbf{V}_{\text{coh}}^T \mathbf{n}_{\text{coh}}^{(\text{pc})} \right]. \quad (\text{A } 8)$$

If the noise pc-space is normalized with respect to the standard deviations  $\sqrt{d_{\text{coh},i}}$

$$\tilde{\mathbf{n}}_{\text{coh}}^{(\text{pc}_n)} = \sqrt{\mathbf{D}_{\text{coh}}^{-1}} \mathbf{n}_{\text{coh}}^{(\text{pc}_n)}, \quad (\text{A } 9)$$

then each component  $\tilde{n}_{\text{coh},i}^{(\text{pc}_n)}$  of the vector  $\tilde{\mathbf{n}}_{\text{coh}}^{(\text{pc}_n)}$  in the normalized noise pc-space has the standard normal distribution  $\tilde{n}_{\text{coh},i}^{(\text{pc}_n)} \sim \mathcal{N}(0, 1)$ . Therefore, the noise probability density function  $\rho_n(\tilde{\mathbf{n}}_{\text{coh}}^{(\text{pc}_n)})$  in the normalized noise pc-space is given by

$$\rho_n(\tilde{\mathbf{n}}_{\text{coh}}^{(\text{pc}_n)}) = \frac{1}{(2\pi)^{N_s/2}} \exp \left[ -\frac{1}{2} \|\tilde{\mathbf{n}}_{\text{coh}}^{(\text{pc}_n)}\|^2 \right]. \quad (\text{A } 10)$$

Expression (A 10) provides an efficient way to calculate the defect characterization parameters. According to the discussion in §4, the characterization result corresponds to the point on the  $d$ -manifold with the highest probability. Formula (A 10) shows that the noise distribution in the normalized noise pc-space is isotropic and the maximum probability point corresponds to the smallest distance from the measurement point to the  $d$ -manifold in this space. Therefore, the characterization result can be calculated as a projection of the measurement point onto the  $d$ -manifold in the normalized noise pc-space. In this case, the method described in appendix Aa can be used; however, all vectors have to be transformed into the normalized noise pc-space using expressions (A 5) and (A 9) first. Alternatively, expression (A 8) shows that this projection can also be performed in the defect class pc-space, but in this case a different metric, corresponding to the distance in the normalized noise pc-space, must be used.

Finally, it is noted that, similarly to the defect  $d$ -manifold, the noise distribution  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  in the noise pc-space is described by the relatively small number of the first largest components, so the noise variances  $d_{\text{coh},i} \rightarrow 0$  as  $i \rightarrow \infty$ . Therefore, in order to avoid large errors in calculation of distances in the normalized noise pc-space, vectors  $\mathbf{n}_{\text{coh}}^{(\text{pc}_n)}$  have to be truncated using some threshold before applying scaling (A 9).

## References

1. Achenbach JD. 2000 Quantitative nondestructive evaluation. *Int. J. Solids Struct.* **37**, 13–27. (doi:10.1016/S0020-7683(99)00074-8)
2. Worden K, Farrar CR, Manson G, Park G. 2007 The fundamental axioms of structural health monitoring. *Proc. R. Soc. A* **463**, 1639–1664. (doi:10.1098/rspa.2007.1834)
3. Bray DE, Stanley RK. 1997 *Non-destructive evaluation, a tool in design manufacturing and service*, revised edn. Boca Raton, FL: CRC Press.
4. Kundu T. 2004 *Ultrasonic nondestructive evaluation: engineering and biological material characterization*. Boca Raton, FL: CRC Press.

5. Farrar CR, Worden K. 2007 An introduction to structural health monitoring. *Phil. Trans. R. Soc. A* **364**, 303–315. (doi:10.1098/rsta.2006.1928)
6. Cunningham LJ, Mulholland AJ, Tant KMM, Gachagan A, Harvey G, Bird C. 2016 The detection of flaws in austenitic welds using the decomposition of the time-reversal operator. *Proc. R. Soc. A* **472**, 20150500. (doi:10.1098/rspa.2015.0500)
7. Connolly GD, Lowe MJS, Temple JAG, Rokhlin SI. 2009 The application of Fermat's principle for imaging anisotropic and inhomogeneous media with application to austenitic steel weld inspection. *Proc. R. Soc. A* **465**, 3401–3423. (doi:10.1098/rspa.2009.0272)
8. Callister WD. 2006 *Materials science and engineering: an introduction*, 7th edn. New York, NY: Wiley.
9. Doyle PA, Scala CM. 1978 Crack depth measurement by ultrasonics: a review. *NDT E Int.* **16**, 164–170. (doi:10.1016/0041-624X(78)90072-0)
10. Blitz J, Simpson G. 1996 *Ultrasonic methods of non-destructive testing*. London, UK: Chapman & Hall.
11. Zhang J, Drinkwater BW, Wilcox PD. 2010 The use of ultrasonic arrays to characterise crack-like defects. *J. Nondestruct. Eval.* **29**, 222–232. (doi:10.1007/s10921-010-0080-6)
12. Bai L, Velichko A, Drinkwater BW. 2015 Ultrasonic characterization of crack-like defects using scattering matrix similarity metrics. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **62**, 545–559. (doi:10.1109/TUFFC.2014.006848)
13. Baby S, Balasubramanian T, Pardikar RJ, Palaniappan M, Subbaratnam R. 2003 Time-of-flight diffraction (TOFD) technique for accurate sizing of surface-breaking cracks. *Insight* **45**, 426–430. (doi:10.1109/TUFFC.2014.006848)
14. Felice MV, Velichko A, Wilcox PD. 2014 Accurate depth measurement of small surface-breaking cracks using an ultrasonic array post-processing technique. *NDT E Int.* **68**, 105–112. (doi:10.1016/j.ndteint.2014.08.004)
15. Moreau L, Hunter A, Velichko A, Wilcox P. 2014 3-D reconstruction of sub-wavelength scatterers from the measurement of scattered fields in elastic waveguides. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **61**, 1864–1879. (doi:10.1109/TUFFC.2014.006619)
16. Huithwaite P. 2016 Guided wave tomography with an improved scattering model. *Proc. R. Soc. A* **472**, 20160643. (doi:10.1098/rspa.2016.0643)
17. Colton D, Coyle J, Monk P. 2000 Recent developments in inverse acoustic scattering theory. *SIAM Rev.* **42**, 369–414. (doi:10.1137/S0036144500367337)
18. Colton D, Haddar H, Piana M. 2003 The linear sampling method in inverse electromagnetic scattering theory. *Inverse Probl.* **19**, S105–S137. (doi:10.1088/0266-5611/19/6/057)
19. Potthast R. 2006 A survey on sampling and probe methods for inverse problems. *Inverse Probl.* **22**, R1–R47. (doi:10.1088/0266-5611/22/2/R01)
20. Hastie T, Tibshirani R, Friedman J. 2009 *The elements of statistical learning*, 2nd edn. New York, NY: Springer-Verlag.
21. Theodoridis S, Koutroumbas K. 2006 *Pattern recognition*, 3rd edn. Orlando, FL: Academic Press.
22. Sambath S, Nagaraj P, Selvakumar N. 2011 Automatic defect classification in ultrasonic NDT using artificial intelligence. *J. Nondestruct. Eval.* **30**, 20–28. (doi:10.1007/s10921-010-0086-0)
23. Romdhani S, Torr P, Schölkopf B, Blake A. 2004 Efficient face detection by a cascaded support-vector machine expansion. *Proc. R. Soc. A* **460**, 3283–3297. (doi:10.1098/rspa.2004.1333)
24. Masnata A, Sunseri M. 1996 Neural network classification of flaws detected by ultrasonic means. *NDT E Int.* **29**, 87–93. (doi:10.1016/0963-8695(95)00053-4)
25. Zhao Q, Bao Z. 1996 Radar target recognition using a radial basis function neural network. *Neural Net.* **9**, 709–720. (doi:10.1016/0893-6080(96)00088-3)
26. Fischell EM, Schmidt H. 2015 Classification of underwater targets from autonomous underwater vehicle sampled bistatic acoustic scattered fields. *J. Acoust. Soc. Am.* **138**, 3773–3784. (doi:10.1121/1.4938017)
27. Bajaj V, Pachori RB. 2013 Automatic classification of sleep stages based on the time-frequency image of EEG signals. *Comput. Methods Prog. Bio.* **112**, 320–328. (doi:10.1016/j.cmpb.2013.07.006)
28. Guyon I, Weston J, Barnhill S, Vapnik V. 2002 Gene selection for cancer classification using support vector machines. *Mach. Learn.* **46**, 389–422. (doi:10.1023/A:1012487302797)



29. Bai L, Velichko A, Drinkwater BW. 2015 Characterization of defects using ultrasonic arrays: a dynamic classifier approach. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **62**, 2146–2160. (doi:10.1109/TUFFC.2015.007334)
30. Malkin R, Caleap M, Drinkwater BW. 2016 A numerical database for ultrasonic defect characterisation using array data: robustness and accuracy. *NDT E Int.* **83**, 94–103. (doi:10.1016/j.ndteint.2016.06.006)
31. Gerard DA, Koss DA. 1990 Porosity and crack initiation during low cycle fatigue. *Mat. Sci. Eng. A-Struct.* **129**, 77–85. (doi:10.1016/0921-5093(90)90346-5)
32. Shi F, Choi W, Lowe MJS, Skelton EA, Craster RV. 2015 The validity of Kirchhoff theory for scattering of elastic waves from rough surfaces. *Proc. R. Soc. A* **471**, 20140977. (doi:10.1098/rspa.2014.0977)
33. Velichko A, Wilcox PD. 2012 Efficient finite element modelling of elastodynamic scattering with non-reflecting boundary conditions. *AIP Conf. Proc.* **1430**, 142–149. (doi:10.1063/1.4716224)
34. Jolliffe IT. 2002 *Principal component analysis*, 2nd edn. New York, NY: Springer.
35. Cross EJ, Manson G, Worden K, Pierce SG. 2012 Features for damage detection with insensitivity to environmental and operational variations. *Proc. R. Soc. A* **468**, 4098–4122. (doi:10.1098/rspa.2012.0031)
36. Dijkstra E. 1959 A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269–271. (doi:10.1007/BF01386390)
37. Whitehouse DJ. 2000 Surface characterization and roughness measurement in engineering. *Photomech. Top. Appl. Phys.* **77**, 413–461. (doi:10.1007/3-540-48800-6\_12)
38. Holmes C, Drinkwater BW, Wilcox PD. 2005 Post-processing of the full matrix of ultrasonic transmit-receive array data for non-destructive evaluation. *NDT E Int.* **38**, 701–711. (doi:10.1016/j.ndteint.2005.04.002)
39. Zhang J, Drinkwater BW, Wilcox PD. 2008 Defect characterization using an ultrasonic array to measure the scattering coefficient matrix. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **55**, 2254–2265. (doi:10.1109/TUFFC.924)
40. Zhang J, Drinkwater BW, Wilcox PD. 2013 Comparison of ultrasonic array imaging algorithms for nondestructive evaluation. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **60**, 1732–1745. (doi:10.1109/TUFFC.2013.2754)
41. Farin G. 1997 *Curves and surfaces for computer aided geometric design*, 4th edn. San Diego, CA: Academic.
42. Amidror I. 2002 Scattered data interpolation methods for electronic imaging systems: a survey. *J. Electron. Imaging* **11**, 157–176. (doi:10.1117/1.1455013)