Budyn, N. S., Croxford, A. J., Bevan, R. L. T., Zhang, J., & Wilcox, P. D. (2021). Characterisation of small embedded two-dimensional defects using multi-view Total Focusing Method imaging algorithm. *NDT and E International*, 119, [102413].
<https://doi.org/10.1016/j.ndteint.2021.102413>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.ndteint.2021.102413](https://doi.org/10.1016/j.ndteint.2021.102413)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://doi.org/10.1016/j.ndteint.2021.102413>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Characterisation of small embedded two-dimensional defects using multi-view Total Focusing Method imaging algorithm

Nicolas Budyn, Anthony J. Croxford, Rhodri L. T. Bevan,
Jie Zhang, Paul D. Wilcox

25th October 2020

Abstract

Defect characterisation is an essential step of nondestructive testing which aims to determine the shape and the size of a defect once detected. The imaging algorithm multi-view Total Focusing Method (TFM) has been successfully used to accurately size defects of size greater than several times the wavelength. However, characterising smaller defects remains difficult due to the diffraction limit. This paper introduces a probabilistic characterisation technique where the unknown defect is compared against a database of reference scatterers obtained with an ultrasonic model. The measured TFM amplitudes act as a defect signature. A Bayesian framework is used to retrieve the reference defects which most credibly explain the measurements, and to calculate uncertainties.

1 Introduction

Defect characterisation is an essential step of nondestructive testing which aims to determine the shape and the size of a defect once detected. In ultrasonic testing, the adoption of array transducers led to the development of numerous techniques, as highlighted by Felice and Fan in a recent review [1]. Larger defects (size higher than a few wavelengths) are typically sized by analysing the tip diffraction or the specular reflection echoes in an A-scan or ultrasonic image. The Multi-view Total Focusing Method (TFM) [2] has been experimentally demonstrated to be especially reliable and accurate for such sizing [3, chapter 5][4]. The characterisation can be furthermore improved by fusing the TFM views [5]. However, the characterisation of smaller defects (size lower than a few wavelengths) remains difficult due to the diffraction limit, which affects beamforming algorithms. The time-reversal MUSIC imaging algorithm [6, 7] can achieve a higher resolution (super resolution), at the price of degraded performance in the presence of noise [8]. Another recent and promising technique is to use a deep learning architecture on TFM images to achieve super resolution [9]. A different route for characterising small defects is to focus on measuring and exploiting the far-field scattered wave field, via the so-called scattering matrix. A common method is to measure the scattering matrix, or a proxy for it, and to find the idealised defect (typically, a perfect crack or an elliptical void) which has

Figure 1: Illustration of the experimental set-up, and of the ultrasonic paths in three TFM views: (a) T-T, (b) LT-T (or T-TL), (c) LL-TL (or LT-LL). The leg in water, always longitudinal, is ignored in this nomenclature.

a scattering matrix that most closely matches that of the unknown defect. These techniques differ mainly by the feature used for the comparison and the class of idealised defect. One technique relies on the maximum amplitude of the scattering matrix to characterise crack-like defects: the angle at which it occurs give the orientation, the length is obtained by comparison against a finite-element model [10]. For the same purpose, other features have also been used such as the width at half maximum of the scattering coefficients of pulse-echo traces [2, 11], the sum of differences of the measured and modelled normalised scattering matrices [12], correlation coefficients and structural similarity index [13], and classifiers in the principal component space [14, 15]. These techniques work best for defects close to the ultrasonic array, so that they are probed over a large range of angles to extract a portion of the scattering matrix. However, the geometry of many practical applications precludes this approach. In the configuration where the defect is far away from the ultrasonic array and insonified over a narrow range of angles, Zhang et al. observed that volumetric and crack-like defects can be distinguished by comparing their image amplitudes in the multi-view TFM algorithm [16]. Instead of using a continuous region in the scattering matrix for one mode combination as the defect signature, their approach uses the maximum responses in multi-view TFM images; effectively these correspond to values taken from multiple smaller regions in the scattering matrices for different mode combinations. The current paper adopts this latter approach, and proposes a quantitative framework to compare these amplitudes against reference scatterers to ultimately determine which ones are the most similar to the unknown defect. Following previous work [14, 17], two types of reference scatterers are considered: the ellipse as an idealised volumetric defect (void, inclusion), and a straight line discontinuity as an idealised crack. The latter is henceforward referred to as crack for brevity. The inference is done using a Bayesian framework.

2 Context

The Total Focusing Method forms an image by delaying and summing the ultrasonic time-traces from every transmitter-receiver pair of the array, termed the Full Matrix Capture (FMC) data [18]. As illustrated in fig. 1, by considering different bulk wave modes propagating in the sample, namely longitudinal (L) and transverse (T) waves, with or without a reflection against the back surface, multiple images can be formed from the same FMC dataset, an approach termed multi-view TFM [2]. Figure 2 shows the TFM views T-T and LT-T of three idealised defects in metal samples: a crack, a side-drilled hole (SDH) and a rectangular

Figure 2: TFM images of three different defects in two different views. The defect shapes are drawn to scale. (a) and (d): 0.5λ long crack, where λ denotes the longitudinal wavelength in the sample. (b) and (e): side-drilled hole of diameter 0.4λ . (c) and (f): 1.2λ -long vertical notch. (a),(b),(c): view T–T. (d),(e),(f): view LT–T. Each image is plotted on a 20 dB dynamic range, where 0 dB corresponds to its maximum amplitude.

notch. These three datasets were the object of a previous publication [19], where the exact set-up can be found; a brief description is recalled here for convenience. The crack dataset results from a 2D finite-element-method simulation of a copper-nickel block, using the Pogo solver [20]. The crack is 0.5λ -long, λ denoting the longitudinal wavelength in the sample, pointing at 20° from the vertical (north-west), and was probed with a 110-element 5-MHz array. The SDH dataset was obtained experimentally from an aluminium sample, inspected with a 64-element 2.5-MHz array. The SDH has a diameter of 0.4λ . Finally, the notch dataset was also obtained experimentally from an aluminium sample, with the same probe as for the SDH. The notch is rectangular and with the long dimension aligned vertically, with a height of 1.2λ and a width of 0.4λ . The ultrasonic data were acquired with a Peak NDT (Derby, U.K.) MicroPulse FMC acquisition device. The multi-TFM algorithm is applied on the full-waveform FMC data. Because linear imaging algorithms such as multi-view TFM are diffraction-limited, the minimum resolution in the image is approximately the wavelength. Therefore, the shapes of the defects shown in fig. 2 are not resolved: the shapes of the indications in the images relate more to the point-spread function rather than the actual defect shapes, which makes their characterisation difficult. Note that the tip diffractions of the notch are just visible in fig. 2c, but this could easily be missed in the presence of stronger structural noise. To attempt to determine the actual shape and size of these defects, a characterisation model is introduced in the following section.

3 Characterisation model

3.1 Overview

The aim is to characterise an unknown defect from its response in multi-view TFM images. To do so, it is compared against reference scatterers. The peak TFM image amplitudes at or close to the defect location in all exploitable views act as a defect signature. A fast semi-analytical ultrasonic model is used to predict the TFM amplitudes of the reference scatterers. The mismatch between the measured response and the closest reference scatterer is treated stochastically with probabilistic modelling, and more specifically using a Bayesian framework. Under this approach, both the vector of observations, denoted x , and the vector which parametrises the reference scatterer, denoted ξ , are treated as random vectors. Two probability distributions are necessary for the inference. The first one describes the plausible TFM amplitudes given that the true defect is a known reference scatterer (forward problem); formally, it is the probability distribution of the observation vector x conditional on the parameter vector ξ . The second one describes the knowledge, or the lack of knowledge, on the parameter vector ξ prior to the observation; it is called the *prior* distribution. Bayes' theorem solves the inverse problem by providing the probability distribution of the parameter vector ξ given the observation x , called the *posterior* distribution. The inference is made by exploiting the posterior distribution.

The reader is referred to classic books on Bayesian analysis for further information, such as (alphabetically) Gelman [21], Kruschke [22], Marin and Robert [23], Robert [24]; this last reference notably contains a philosophical defence of this framework in chapter 11. Bayesian analysis relies heavily on conditional probability. Throughout this paper, for brevity and as usually done in the Bayesian literature, the conditional density of a random variable Y given an observation $X = x$ is denoted $f(y | x)$. $f(x, y)$ denotes the density of the joint distribution

Figure 3: Parametrisation of (a) the ellipse and (b) the crack.

of (X, Y) . $f(x)$ denotes the density of X unconditionally on all other random variables. $f(x)$ will also denote the probability mass function, $P(X = x)$, when X is a discrete random variable.

3.2 Defect signature

The vector of the peak TFM image amplitudes (absolute value of the complex image amplitude) close to or at the defect location in each view acts as a signature of the defect. It is denoted $x = (x_k)_{k=1\dots n}$, where n is the number of views. This vector is used to compare the unknown defect against the reference defect; in machine learning terminology, this is the feature of the model. The case of interest here is relatively small defects, so the defect image depends primarily on the point spread function. The actual shape of the defect is unresolved due to the diffraction limit, which justifies using a single amplitude per image.

3.3 Reference scatterers

The unknown defect is compared against reference scatterers, parametrised by the vector ξ . It is essential that these scatterers cover a wide range of shape and orientation to achieve a good characterisation. Two distinct sets are used: ellipses and cracks. In both sets, the scatterers are two-dimensional (infinite length in the third dimension) and smooth. Figure 3a shows the parametrisation of the ellipse, $\xi = (l, r, \varphi)$: l is the major length, $r \in]0, 1]$ is the aspect ratio, which is the ratio of the minor length to the major length, and φ is the angle. Figure 3b shows the parametrisation of the crack, $\xi = (l, \varphi)$: l is the length, and φ the angle.

The TFM amplitudes of the reference scatterers are calculated with a fast semi-analytical 2D ultrasonic model described in an earlier publication [19], termed the sensitivity model, derived from a ray-based approach [25]. This model uses a single-frequency approximation for computational efficiency. The scattering functions are obtained with a finite-element method [26] for the ellipse, and a semi-analytical model for the crack [27]. The model requires knowing the defect location, which is typically obtained using the experimental TFM images. The ultrasonic model is run on a discrete set of reference defects, with evenly spaced parameters. The limits of the parameters are as follows. The angle φ ranges between 0° and 180° with a step of 2.5° , which covers the whole range of possible angles due to the line symmetry along the defect main axis. The length l ranges between 0.1λ and 2λ with a step of 0.1λ . There

is little incentive to consider larger defects as in such a scenario image-based sizing can be applied [28]. For the ellipse, the aspect ratio r ranges between 0.1 (thin ellipse) and 1 (circle) with a step of 0.1; thinner ellipses were not considered as their sharp geometry degrades the quality of the finite-element mesh. This leads to 14 600 ellipses and 1460 cracks to evaluate with the ultrasonic model. The reference scatterer database is upsampled using cubic spline interpolation to predict the TFM amplitudes of defects which are not evaluated with the ultrasonic model, saving computational time. The interpolation error is negligible compared to the ultrasonic model error: the former is bounded by the difference of amplitudes between two model evaluations, which is on average 0.8 dB here, whereas the latter is typically around 3 dB as discussed in the next section. After upsampling, the databases are populated with 2025 000 ellipses and 13 500 cracks.

Here, it is assumed that little is known on the defect before the characterisation, so the chosen prior distribution of each of the parameters l , r and φ , denoted $f(l)$, $f(r)$, $f(\varphi)$, is a uniform distribution with the limits specified in the previous paragraph. Because these parameters are independent, the prior on ξ is $f(\xi) = f(l)f(r)f(\varphi)$ for the ellipse and $f(\xi) = f(l)f(r)$ for the crack. The prior distribution can be changed to reflect further knowledge on the expected defects. For example, if it is known that the defects in a specific component are growing preferentially in a given orientation, $f(\varphi)$ could be changed to a normal distribution centred at this angle.

3.4 Likelihood

Suppose that the true defect is a known scatterer ξ . The objective is to determine what defect signatures are likely to be observed; in other words, to determine the likelihood function for the observation vector x conditionally on ξ , denoted $f(x | \xi)$. The ultrasonic model described in [19] gives a starting point; the defect signature generated with this model is denoted $h(\xi)$ and is in dB. However, there are multiple sources of uncertainties that need to be accounted for. As for any physics model, the ultrasonic model is a simplification of the reality, based on various assumptions and approximations: linear behaviour of the transducer, homogeneity and isotropy of the material, two-dimensional treatment, plane wave approximation, etc. As the reality is more complex, there is a model error. Also, the unknown defect is likely not a perfect ellipse or crack; the reference scatterers only approximate the real defect. Modelling the uncertainties deterministically is difficult if not impossible; they are therefore treated here as a single additive stochastic error term ε in dB scale (multiplicative error in linear scale), that is

$$x = h(\xi) + \varepsilon. \quad (1)$$

The random noise, typically coming from the electronics, and the structural noise, caused by the material microstructure, both usually modelled with an additive term in linear scale, are not covered in this model, as they are both small compared to the multiplicative noise. Figure 4 shows the empirical distribution of the error ε for the three datasets introduced in section 2 grouped together. All values but one are scattered around 0. The density of the main cluster is approximatively bell-shaped, and is slightly skewed towards positive errors, that is the ultrasonic model underestimates the results. There is a single outlier at 21 dB from the SDH dataset, caused by a combination of a poor signal-to-noise ratio in this view and the low accuracy of the sensitivity model. It is not judicious to use this empirical distribution as the *true* distribution of ε as it may not generalise to other datasets (overfitting); a simpler (less

Figure 4: Errors in dB between the sensitivity model and the observations for the three datasets (54 data points), shown as a histogram (light grey), an empirical density obtained with kernel density estimation (block line), and a rug plot (black rectangles). The dashed black line is the $t(\mu = 0 \text{ dB}, \sigma = 2 \text{ dB}, \nu = 2)$ distribution, representative of prior assumptions on the model error.

accurate but more generalisable) distribution is needed. Because the density is approximatively bell-shaped, ε could be assumed to be normally distributed. However, such a model would be very sensitive to outlying values, which are likely to arise with real data. A more robust regression can be achieved by using as an alternative the Student's t-distribution [21, §17.5]. The Student's t-distribution, $t(\mu, \sigma, \nu)$, parametrised by the location μ , the scale parameter $\sigma > 0$ and the shape parameter $\nu > 0$, is bell-shaped, symmetric with respect to μ , and is a generalisation of the normal distribution $\text{Norm}(\mu, \sigma)$, which corresponds to $\nu = +\infty$. The t-distribution has a heavier tail (the smaller ν , the heavier the tail), so it better fits data with outliers. The location parameter is assumed to be $\mu = 0$, which neglects the slight systematic error in the ultrasonic model to prevent overfitting. Assuming furthermore that (i) the model accuracy is independent of the view, that is the error terms are identically distributed, and (ii) the error terms in different views are independent, then the observation vector conditional on the model parameters has for density

$$f(x \mid \xi, \sigma, \nu) = \prod_{k=1}^n t(x_k - h_k(\xi), \sigma, \nu) = \prod_{k=1}^n \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{(x_k - h_k(\xi))^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}, \quad (2)$$

where Γ is the gamma function (extension of the factorial function), and k is the view index.

The parameters σ and ν are not of immediate interest for the characterisation (in statistical terminology, they are described as nuisance parameters) but they must be chosen carefully as they affect the result. They both quantify the confidence in the ultrasonic model. σ is the typical error in dB between the ultrasonic model and the result. A very large σ leads to a flat density $f(x \mid \xi, \sigma, \nu)$, independent of the observation; the characterisation result would be wholly controlled by the prior distribution on ξ , that is in this case that all reference scatterers

are equally credible. A zero σ leads to a Dirac distribution: only a single scatterer is credible, all others are not; this point estimation would ignore the uncertainties, so it would give a sub-optimal result with excessive confidence. From fig. 4, σ is estimated at a glance to be around 2 or 3 dB. To account for the uncertainty on this parameter, and to avoid the characterisation result depending strongly on an arbitrarily chosen value, σ is left as a stochastic variable with as prior distribution the normal distribution $f(\sigma) = \text{Norm}(2 \text{ dB}, 0.5 \text{ dB})$, truncated to censor the negative values (which happens with a small probability of 3×10^{-5}). The parameter ν controls the tail of the t-distribution, and hence the influence of outliers. A large ν leads to a high sensitivity to outliers; a small ν leads to a more robust but more indecisive characterisation. Gelman recommends a value of a few units [21, §17.5]. As for σ , ν could be treated stochastically, at the expense of an increased computational cost. Therefore, a fixed value of $\nu = 2$ was adopted as it appears to be sufficient to produce acceptable results. A t-distribution representative of these prior assumptions on the model error is shown in fig. 4; it is close to the empirical density.

3.5 Posterior distribution

The likelihood $f(x | \xi, \sigma, \nu)$ (eq. (2)), or $f(x | \xi, \sigma)$ as ν is fixed, relates to the forward problem, that is information on plausible observation x given a reference scatterer ξ . The quantity of interest corresponds however to the inverse problem, that is information on a plausible reference scatterer ξ given an observation x ; formally, $f(\xi | x)$. The inversion is done with Bayes' theorem:

$$f(\xi, \sigma | x) = \frac{f(x | \xi, \sigma)f(\xi, \sigma)}{f(x)}. \quad (3)$$

$f(\xi, \sigma | x)$ is called the *posterior* distribution. The posterior distribution can be seen as the prior distribution, $f(\xi, \sigma)$, updated with the information from the observation via the forward model, $f(x | \xi, \sigma)$. As the parameters ξ and σ are independent, the prior distribution is $f(\xi, \sigma) = f(\xi)f(\sigma)$; $f(\xi)$ was defined in section 3.3, and $f(\sigma)$ in section 3.4. $f(x)$ is formally the marginal distribution of the observation; in practice, it acts as a normalisation constant: recalling that the integral of a probability density function is 1, in particular for $f(\xi, \sigma | x)$:

$$\int \int_{-\infty}^{+\infty} f(\xi, \sigma | x) d\xi d\sigma = 1, \quad (4)$$

then with eq. (3)

$$f(x) = \int \int_{-\infty}^{+\infty} f(x | \xi, \sigma)f(\xi, \sigma) d\xi d\sigma. \quad (5)$$

Note that the notations $\int d\xi = \iiint dl dr d\varphi$ for the ellipse and $\int d\xi = \iint dl d\varphi$ for the crack are used. The posterior distribution without the nuisance parameter σ is obtained by integrating it out:

$$f(\xi | x) = \int_{-\infty}^{+\infty} \frac{f(x | \xi, \sigma)f(\xi, \sigma)}{f(x)} d\sigma. \quad (6)$$

This distribution contains all the possible information on the defect size and orientation recovered by the characterisation model. However, it may be difficult to interpret, so it is useful to summarise it into more tangible quantities, which is the object of the next two subsections.

3.6 Characterisation methodology

The unknown defect can be characterised by deriving from the posterior distribution a single reference scatterer which represents it the “best”, in some sense that needs to be defined. Different point estimates exist, depending on the definition of “best”. A first option is to maximise the posterior density, which gives the maximum a posteriori (MAP) [24, §4.1.2]

$$\xi_{\text{MAP}} = \arg \max_{\xi} f(\xi | x). \quad (7)$$

The MAP is a popular choice in machine learning due to its simplicity [29]. A known limit is its dependence on the (arbitrary) parametrisation of the model [30]: for example, parametrising the defect shape by its major and minor lengths instead of its major and its aspect ratio could lead to a different MAP. An alternative to the MAP is to determine the scatterer which minimises the estimation error of the size and of the orientation. Minimising the quadratic error leads to the posterior mean [24, chapter 2]

$$\bar{\xi} = E(\xi | x) = \int \xi f(\xi | x) d\xi. \quad (8)$$

Minimising the absolute error leads to the posterior median (ibid). These three estimators, MAP, posterior mean and median, are reasonable; however, anticipating the results, they may lead to rather different estimations, which makes the final assessment difficult. The underlying reason is that the characterisation results may be non-unique: multiple scatterers may have the same defect signature. Also, point estimators do not convey confidence information, which is crucial in this application: is the estimated size precise up to 0.01λ ? 0.1λ ? 1λ ?

To overcome these two issues, the uncertainties on the characterisation can be quantified with Highest Posterior Density (HPD) credible sets; for example, “there is a 99% chance that the defect size is between 0.4λ and 0.7λ ”. Formally, for a given probability threshold, α (here, 99%), C_α is a HPD credible set if by definition there exists a density threshold k_α so that [24]

$$C_\alpha = \{\xi : f(\xi | x) \geq k_\alpha\}, \quad (9)$$

and additionally

$$P(C_\alpha | x) = \int_{C_\alpha} f(\xi | x) d\xi = \alpha. \quad (10)$$

A HPD credible set is therefore constructed by selecting in priority the regions of highest density (first equation), until a total posterior probability of α is reached (second equation). Other α -credible sets exist; however, the HPD sets have the smallest volume amongst them (ibid). A criticism is that HPD sets depend on the arbitrary model parametrisation, which is undesirable [30]; their simplicity makes them however a valuable tool.

3.7 Classification of the defect

Two classes of defects are considered, leading to two characterisation models which are run independently: M_1 for ellipses, and M_2 for cracks. They lead to different results: $\xi_1 | x$ under M_1 , $\xi_2 | x$ under M_2 . A solution to determine which model is the most suited is to treat M_1 and M_2 as stochastic variables, and to calculate their posterior probability $f(M_i | x)$ [24, §5.2.2][21, §7.4]. Using Bayes’ theorem, the ratio of posterior probabilities is

$$K = \frac{f(M_1 | x)}{f(M_2 | x)} = \frac{f(M_1)}{f(M_2)} B_{12}, \quad (11)$$

where B_{12} , called the Bayes factor, is defined as

$$B_{12} = \frac{f(x \mid M_1)}{f(x \mid M_2)}. \quad (12)$$

Assuming the two models are as reasonable, their prior probabilities are equal $f(M_1) = f(M_2)$, so K and B_{12} are equal. B_{12} is often used as a more objective proxy for K as it does not depend on the prior on M_1 and M_2 . For $i \in \{1, 2\}$, the likelihood term is

$$f(x \mid M_i) = \int f(x, \xi_i \mid M_i) d\xi_i = \int f(x \mid \xi_i, M_i) f(\xi_i \mid M_i) d\xi_i, \quad (13)$$

where $f(x \mid \xi_i, M_i)$ and $f(\xi_i \mid M_i)$ are respectively the likelihood function and the prior distribution defined in section 3, with a change of notation to make the model M_i explicit.

The Bayes factor quantifies how much the observation supports M_2 versus M_1 . B_{12} is greater than 1 if the data supports M_1 more than M_2 ; B_{12} is less than 1 if the data support M_2 more than M_1 . For example, $B_{12} = 10$ means that it is 10 times more likely the true defect is (or scatters like) an ellipse (M_1) than a crack (M_2). The choice of comparing M_1 against M_2 , or M_2 against M_1 , is arbitrary as the corresponding Bayes factors are related: $B_{12} = 1/B_{21}$. The following scale in dB ($20 \log_{10}(\cdot)$) is usually used [31]: the evidence for M_1 against M_2 is *not worth more than a bare mention* if B_{12} is between 0 and 10 dB, *substantial* between 10 and 20 dB, *strong* between 20 and 40 dB, and *decisive* above 40 dB. Conversely, the scale of evidence for M_2 against M_1 corresponds to opposite dB values: not worth more than a bare mention if B_{12} is between -10 and 0 dB, etc.

The Bayes factor automatically penalises unnecessarily complex models: if the two models fit the observations equally well, it favours the simpler model, which can be seen as an implementation of Ockham's razor [31, 32]. The underlying reason is the probability density tends to spread more evenly as the number of model parameters increase, whereas the density of a simpler model remains sharper provided the model is adequate. Complex models always fit the data better than simpler ones because of their increased number of parameters, however this is not necessarily desirable as they also fit better arbitrary data, including meaningless data, which may lead to an erroneous interpretation of their results. Approximating a crack by an infinitely thin ellipse, the ellipse model can be seen as a generalisation of the crack model; so the ellipse model should theoretically always fit the data at least as well as the crack model. For example, a Bayes factor favouring the crack model, $B_{12} < 0$ dB, indicates that the marginally better fit provided by the extra aspect ratio parameter is not worth the increased complexity.

3.8 Numerical implementation

The Bayesian analysis is implemented as follows. For each of the two defect classes, the database of reference signatures is first formed as described in section 3.3, then the full posterior density $f(\xi, \sigma \mid x)$ is evaluated using eq. (3). Then, numerical integration (rectangle rule, which is sufficiently accurate here) is used to calculate the posterior density $f(\xi \mid x)$ (eq. (6)), the posterior mean and median, and the Bayes factor. Posterior samples, used only for visualisation, are obtained using an algorithm based on grid approximation [21, p. 76]. The calculation of the HPD is done by determining the threshold k_α with numerical optimisation, $P(C_\alpha \mid x)$ being obtained again with numerical integration. For details, the code source

is available as supporting data. The described implementation ensures that the parameter space is well sampled, and yet requires little tuning. However, it is only tractable because the number of parameters is limited (at most 4, for the ellipse model). For models with more parameters, Markov Chain Monte Carlo (MCMC), a class of stochastic algorithms widely used to sample the posterior distribution, would be a suitable alternative. From the authors' experience, MCMC tends to require more tuning to ensure that the whole parameter space is sampled, making it more difficult to use; however, this does highlight how the approach could be employed in more complex cases.

4 Results

Having described the characterisation framework, its performance is demonstrated in this section on the SDH, notch and crack datasets described in section 2. The Bayesian analysis is done by drawing posterior samples, and calculating point estimates, HPD credible sets, and Bayes factors to classify the defect.

4.1 SDH

	Major length (λ)	Aspect ratio	Angle ($^\circ$)
Reference	0.39	1.00	—
MAP	0.63	0.68	108.00
Median	0.43	0.86	93.99
Mean	0.46	0.87	80.25
Standard deviation	0.12	0.11	48.17
Lowest bound of HPD	0.23	0.63	0.00
Highest bound of HPD	0.68	1.00	178.00

Table 1: Characterisation results for the SDH using ellipses.

The characterisation results for the SDH using ellipses are presented in table 1, in fig. 5 as marginal densities, and fig. 6 as point estimates and samples. There are two main clusters, visible in the posterior samples (fig. 6b) and the 2D marginal distribution (fig. 5d): the first one, near the posterior mean and median, is close to the true defect boundaries; the second one, near the MAP, corresponds to a larger and more oblong defect. The uncertainty on the angle is high (only 8 degrees are not in the 99% HPD), which is not surprising as the true defect is axisymmetric. The characterised aspect ratio is between 0.6 and 1.0, which includes the true value (1.0). The characterised major length is between 0.2λ and 0.7λ (except a 0.04λ -wide region), which includes the true value (0.4λ).

Thin ellipses are unlikely, which supports the hypothesis that the true defect is volumetric. To confirm this, the characterisation model using cracks is run. The resulting Bayes factor (eq. (12)) is $B_{12} = 139.9\text{dB}$, which corresponds to decisive evidence for the ellipse model against the crack model.

Figure 5: Probability density of the marginals of the posterior distribution for the characterisation of a SDH with ellipses. The 99% HPD regions are shown in light grey in (a), (b), (c). The dash lines and the red cross correspond to the true defect. Point estimates are shown in (d).

Figure 6: Characterisation of the SDH using ellipses. (a) Point estimates. (b) 100 posterior samples.

	Length (λ)	Aspect ratio	Angle ($^\circ$)
Reference	0.50	0.10	110.00
MAP	0.48	0.10	110.00
Median	0.48	0.11	107.96
Mean	0.48	0.13	108.95
Standard deviation	0.03	0.03	2.00
Lowest bound of HPD	0.40	0.10	103.96
Highest bound of HPD	0.54	0.19	113.83

Table 2: Characterisation results for the crack using ellipses. The reference is here the thinnest considered ellipse matching the true crack geometry.

4.2 Crack

The characterisation results for the crack using ellipses are presented in table 2, in fig. 7 as marginal densities, and fig. 8 as point estimates and samples. The posterior distribution has a single mode, which includes the true value. A precision of 0.14λ for the major length, and 9.9° for the angle, is achieved. The aspect ratio marginal peaks at the lower bound (0.1), which suggests that the defect has an aspect ratio of at most 0.1. An aspect ratio below 0.1 seems credible, but a proper investigation would require to run the model with thinner ellipses, which is difficult due to meshing issue in the finite-element simulation.

As the defect appears thin, the possibility that it is planar instead of volumetric is investigated by comparing it against a database of cracks in a second analysis. The results are presented in table 3. Its marginals are shown in fig. 9, and the samples in fig. 10. The results are comparable with the ellipse model (without the aspect ratio): a single mode near the true values. The crack model slightly underestimates the major length; the MAP is 10% under the true value. A Bayes factor of $B_{12} = -32.8\text{dB}$ is obtained, which is a strong evidence for

	Length (λ)	Angle ($^\circ$)
Reference	0.50	110.00
MAP	0.44	110.00
Median	0.43	108.97
Mean	0.43	109.92
Standard deviation	0.02	1.52
Lowest bound of HPD	0.39	105.19
Highest bound of HPD	0.49	113.95

Table 3: Characterisation results for the crack using cracks.

the crack model. This does not guarantee the true defect is a crack; strictly speaking, this Bayes factor only suggests the aspect ratio variable of the ellipse model is an unnecessary complication to explain the observed TFM amplitudes. It does not exclude the possibility that the true defect may be an ellipse thinner than investigated in the ellipse model. However, engineering prudence dictates interpreting the results as the worst case scenario, i.e. that the true defect is a crack.

4.3 Notch

	Length (λ)	Aspect ratio	Angle ($^\circ$)
Reference	1.17	0.33	90.00
MAP	1.09	0.38	100.00
Median	1.08	0.51	92.11
Mean	0.91	0.52	78.62
Standard deviation	0.64	0.10	62.45
Lowest bound of HPD	0.22	0.30	0.00
Highest bound of HPD	2.00	0.84	178.00

Table 4: Characterisation results for the notch using ellipses. The reference is here an ellipse with the same major length and aspect ratio as the notch.

The characterisation results for the notch using ellipses are presented in table 4, in fig. 11 as marginal densities, and fig. 12 as point estimates and samples. The true defect is a vertical (90°) rectangular notch of dimensions $1.2 \times 0.4\lambda$ (0.33 aspect ratio). The posterior distribution presents multiple clusters and large uncertainties. A first cluster corresponds to a small ($\approx 0.3\lambda$) horizontal (0°) ellipse, which may be explained by the scattering of the upper and lower part of the notch. The other credible ellipses match more closely the true major length, aspect ratio and orientation of the notch. In particular, the MAP is an ellipse of major length 1.1λ , which is a 7% relative error compared to the notch. The HPD of the major length ranges from 0.2λ to 2.0λ (the total width of the plausible intervals is 1.1λ after removing the gaps), which is close to the full range studied. Also, its density increases near the upper bound (2.0λ), which was chosen as the validity limit of the ultrasonic model (point-like assumption). This is problematic as it indicates the characterisation may depend on the arbitrary chosen

Figure 7: Probability density of the marginals of the posterior distribution for the characterisation of a crack with ellipses. The 99% HPD regions are shown in light grey in (a), (b), (c). The dash lines and the red cross correspond to an ellipse with similar size and orientation as the true defect, and the minimum investigated aspect ratio.

Figure 8: Characterisation of the crack using ellipses. (a) Point estimates. (b) 100 posterior samples.

upper bound, so the sizing results are questionable. On the positive side, the presence of such peak near at the upper bound provides a warning sign that the characterisation technique may be failing. It could show a fundamental limitation of this characterisation approach to size larger defects, and/or it could be caused by the degradation of the accuracy of the ultrasonic forward model, which is only valid for small defects. Despite these limitations, the characterisation model gives a reasonable idea of the orientation and aspect ratio of the defect, without providing undue certainty. The Bayes factor $B_{12} = 35.5$ dB confirms the data strongly supports the ellipse model over the crack model, so the defect is volumetric.

5 Conclusion

A Bayesian model was introduced to characterise a defect by comparing its signature against reference ellipses and cracks. The experimental validation was done for a crack, a SDH, and a rectangular notch. The characterisation is reasonably good, although the precision is variable: a precision of 0.14λ was obtained in the best case, 1.1λ in the worst case. The inherent uncertainty on the length, aspect ratio and angle of the defect is visualised and quantified with the posterior samples, the marginal densities and the HPD regions. This is especially important for the SDH and the notch, where several clusters of solutions exist. Therefore, this approach gives a higher degree of confidence in the characterisation. In comparison, typically used point estimates convey less information and may be deceiving. The Bayes factor between the ellipse model and the crack model was used to classify a defect as volumetric or crack-like. The crack defect was however well approximated by a thin ellipse.

The main limit is that this characterisation technique was only tested on embedded smooth 2D defects. It could be extended to more realistic cases (3D rough defects), under the condition that a suitable ultrasonic model is available. A second limit is that the assumption that the measured TFM amplitude is the defect response plus-minus a noise term (eq. (1)) breaks down

Figure 9: Probability density of the marginals of the posterior distribution for the characterisation of a crack using cracks. The 99% HPD regions are shown in light grey in (a) and (b). The dash lines and the red cross correspond to the true defect.

Figure 10: Characterisation of the crack using cracks. (a) Point estimates. (b) 100 posterior samples.

if the image signal-to-noise ratio (SNR) is low: if the structural noise is larger than the signal, the measured value is in reality the noise floor. A workaround is to exclude TFM images with low SNR from the characterisation, as done in previous work from the current authors [33]. Future work could focus on including the structural noise level in the uncertainty model, and on the validation of this technique on more cases. Also, it is at this stage unclear whether accounting for the angular variation of the scattering of the defect within individual views would further improve the characterisation; this could be investigated in the future.

6 Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under the grants EP/N015924/1 and EP/L015587/1, the U.K. Research Centre in Non-Destructive Testing (RCNDE), BAE Systems, EDF, Hitachi and Wood. Portions of this work were presented at the 46th Annual Review of Progress in Quantitative Nondestructive Evaluation in 2019, paper number 6965.

7 Data access statement

Supporting data will be available for download from the Research Data Repository of University of Bristol based on paper acceptance.

References

1. Felice, M. V. & Fan, Z. Sizing of Flaws Using Ultrasonic Bulk Wave Testing: A Review. *Ultrasonics* **88**, 26–42. ISSN: 0041-624X (1st Aug. 2018).

Figure 11: Probability density of the marginals of the posterior distribution for the characterisation of a notch with ellipses. The 99% HPD regions are shown in light grey in (a), (b), (c). The dash lines and the red cross correspond to an ellipse with the same major length and aspect ratio as the notch. Point estimates are shown in (d).

Figure 12: Characterisation of the notch using ellipses. (a) Point estimates. (b) 100 posterior samples.

2. Zhang, J., Drinkwater, B. W., Wilcox, P. D. & Hunter, A. J. Defect Detection Using Ultrasonic Arrays: The Multi-Mode Total Focusing Method. *NDT & E International* **43**, 123–133. ISSN: 0963-8695 (Mar. 2010).
3. Sy, K. *Étude et Développement de Méthodes de Caractérisation de Défauts Basées Sur Les Reconstructions Ultrasonores TFM* thesis (Paris Saclay, 15th Feb. 2018).
4. Felice, M. V., Velichko, A. & Wilcox, P. D. Accurate Depth Measurement of Small Surface-Breaking Cracks Using an Ultrasonic Array Post-Processing Technique. *NDT & E International* **68**, 105–112. ISSN: 0963-8695 (Supplement C 1st Dec. 2014).
5. Bevan, R. L., Budyn, N., Zhang, J., Croxford, A. J., Kitazawa, S. & Wilcox, P. D. Data Fusion of Multi-View Ultrasonic Imaging for Characterisation of Large Defects. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **67**, 2387–2401. ISSN: 1525-8955 (2020).
6. Gruber, F. K., Marengo, E. A. & Devaney, A. J. Time-Reversal Imaging with Multiple Signal Classification Considering Multiple Scattering between the Targets. *The Journal of the Acoustical Society of America* **115**, 3042–3047. ISSN: 0001-4966 (1st June 2004).
7. Marengo, E. A., Gruber, F. K. & Simonetti, F. Time-Reversal MUSIC Imaging of Extended Targets. *IEEE Transactions on Image Processing* **16**, 1967–1984. ISSN: 1941-0042 (Aug. 2007).
8. Fan, C., Caleap, M., Pan, M. & Drinkwater, B. W. A Comparison between Ultrasonic Array Beamforming and Super Resolution Imaging Algorithms for Non-Destructive Evaluation. *Ultrasonics* **54**, 1842–1850. ISSN: 0041-624X (Sept. 2014).
9. Song, H. & Yang, Y. Super-Resolution Visualization of Subwavelength Defects via Deep Learning-Enhanced Ultrasonic Beamforming: A Proof-of-Principle Study. *NDT & E International* **116**, 102344. ISSN: 0963-8695 (1st Dec. 2020).

10. Zhang, J., Drinkwater, B. & Wilcox, P. Defect Characterization Using an Ultrasonic Array to Measure the Scattering Coefficient Matrix. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **55**, 2254–2265. ISSN: 0885-3010 (Oct. 2008).
11. Velichko, A. & Wilcox, P. D. Reversible Back-Propagation Imaging Algorithm for Post-processing of Ultrasonic Array Data. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **56**, 2492–2503. ISSN: 0885-3010 (Nov. 2009).
12. Tant, K. M. M., Mulholland, A. J. & Gachagan, A. A Model-Based Approach to Crack Sizing with Ultrasonic Arrays. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **62**, 915–926. ISSN: 0885-3010 (May 2015).
13. Bai, L., Velichko, A. & Drinkwater, B. W. Ultrasonic Characterization of Crack-like Defects Using Scattering Matrix Similarity Metrics. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **62**, 545–559. ISSN: 0885-3010 (Mar. 2015).
14. Bai, L., Velichko, A. & Drinkwater, B. W. Characterization of Defects Using Ultrasonic Arrays: A Dynamic Classifier Approach. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **62**, 2146–2160. ISSN: 0885-3010 (Dec. 2015).
15. Velichko, A., Bai, L. & Drinkwater, B. W. Ultrasonic Defect Characterization Using Parametric-Manifold Mapping. *Proc. R. Soc. A* **473**, 20170056. ISSN: 1364-5021, 1471-2946 (1st June 2017).
16. Zhang, J., Barber, T., Nixon, A. & Wilcox, P. Investigation into Distinguishing between Small Volumetric and Crack-like Defects Using Multi-View Total Focusing Method Images. *AIP Conference Proceedings* **1806**, 040003. ISSN: 0094-243X (16th Feb. 2017).
17. Moreau, L., Hunter, A. J., Velichko, A. & Wilcox, P. D. 3-D Reconstruction of Sub-Wavelength Scatterers from the Measurement of Scattered Fields in Elastic Waveguides. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **61**, 1864–1879 (Nov. 2014).
18. Holmes, C., Drinkwater, B. W. & Wilcox, P. D. Post-Processing of the Full Matrix of Ultrasonic Transmit–Receive Array Data for Non-Destructive Evaluation. *NDT & E International* **38**, 701–711. ISSN: 0963-8695 (Dec. 2005).
19. Budyn, N., Bevan, R. L. T., Zhang, J., Croxford, A. J. & Wilcox, P. D. A Model for Multiview Ultrasonic Array Inspection of Small Two-Dimensional Defects. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* **66**, 1129–1139. ISSN: 1525-8955. pmid: [30990182](#) (2019).
20. Huthwaite, P. Accelerated Finite Element Elastodynamic Simulations Using the GPU. *Journal of Computational Physics* **257**, 687–707. ISSN: 0021-9991 (15th Jan. 2014).
21. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. *Bayesian Data Analysis* Third. 677 pp. ISBN: 978-1-4398-4095-5. Google Books: [ZXL6AQAAQBAJ](#) (CRC Press, 1st Nov. 2013).
22. Kruschke, J. K. *Doing Bayesian Data Analysis : A Tutorial with R and BUGS* ISBN: 978-0-12-381485-2 (Academic Press, Amsterdam ; 2011).
23. Marin, J.-M. & Robert, C. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics* 265 pp. ISBN: 978-0-387-38979-0. Google Books: [Lx74oEz5pmsC](#) (Springer Science & Business Media, 6th Feb. 2007).

24. Robert, C. *The Bayesian Choice* 620 pp. ISBN: 978-0-387-71598-8. Google Books: [6oQ4s8Pq9pYC](#) (Springer Science & Business Media, 27th Aug. 2007).
25. Schmerr, L. W. *Fundamentals of Ultrasonic Nondestructive Evaluation: A Modeling Approach* 2nd ed. 2016 edition. 778 pp. ISBN: 978-3-319-30461-8 (Springer, New York, NY, 18th May 2016).
26. Velichko, A. & Wilcox, P. D. A Generalized Approach for Efficient Finite Element Modeling of Elastodynamic Scattering in Two and Three Dimensions. *The Journal of the Acoustical Society of America* **128**, 1004–1014. ISSN: 0001-4966 (1st Sept. 2010).
27. Glushkov, E., Glushkova, N., Ekhlakov, A. & Shapar, E. An Analytically Based Computer Model for Surface Measurements in Ultrasonic Crack Detection. *Wave Motion* **43**, 458–473. ISSN: 0165-2125 (1st June 2006).
28. Zhang, J., Drinkwater, B. W. & Wilcox, P. D. The Use of Ultrasonic Arrays to Characterize Crack-Like Defects. *Journal of Nondestructive Evaluation* **29**, 222–232. ISSN: 1573-4862 (1st Dec. 2010).
29. Bishop, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)* ISBN: 0-387-31073-8 (Springer-Verlag, Berlin, Heidelberg, 2006).
30. Druilhet, P. & Marin, J.-M. Invariant HPD Credible Sets and MAP Estimators. *Bayesian Analysis* **2**, 681–691. ISSN: 1936-0975, 1931-6690 (Dec. 2007).
31. Kass, R. E. & Raftery, A. E. Bayes Factors. *Journal of the American Statistical Association* **90**, 773–795. ISSN: 0162-1459. JSTOR: [2291091](#) (1995).
32. Jefferys, W. H. & Berger, J. O. Ockham’s Razor and Bayesian Analysis. *American Scientist* **80**, 64–72. ISSN: 0003-0996. JSTOR: [29774559](#) (1992).
33. Bevan, R. L. T., Zhang, J., Budyn, N., Croxford, A. J. & Wilcox, P. D. Experimental Quantification of Noise in Linear Ultrasonic Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **66**, 79–90. ISSN: 0885-3010 (2018).