Jensen-Shannon Divergence for Non-Destructive Incipient Crack Detection and Estimation

Xiaoxia Zhang, Claude Delpha, Demba Diallo

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 28, 2020, accepted June 18, 2020, date of current version July 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3004658

Jensen-Shannon Divergence for Non-Destructive Incipient Crack Detection and Estimation

XIAOXIA ZHANG[1], CLAUDE DELPHA[1], (Senior Member, IEEE), AND DEMBA DIALLO[2], (Senior Member, IEEE)

[1]Université Paris Saclay, CNRS, CentraleSupelec, Laboratoire des Signaux et Systèmes, 91190 Gif Sur Yvette, France
[2]CentraleSupelec, CNRS, Group of Electrical Engineering of Paris, Université Paris Saclay, 91192 Gif Sur Yvette, France

Corresponding author: Claude Delpha (claude.delpha@l2s.centralesupelec.fr)

This work was supported in part by the iCODE Institute, Research Project of the IDEX Paris-Saclay, and in part by the Hadamard Mathematics LabEx (LMH) through the Programme des Investissements d'Avenira under Grant ANR-11-LABX-0056-LMH.

ABSTRACT Nowadays industrial process needs more and more accurate nondestructive procedures for material crack detection and diagnosis. Early detection in that cases are very challenging issues: more the cracks are incipient and higher is the difficulty for its detection and estimation. Indeed, these incipient cracks which cause non-obvious changes in sensor measurements needs to be properly detected and estimated. For conductive materials measurement based on impedance maps obtained from Eddy Current Testing (ECT) are used but the presence of environmental noise can mask the crack information and induce missed detection and false estimation. In this paper, we highlight the limitation of classical techniques and address this problem using a methodology based on wavelet transform and Jensen-Shannon divergence in the framework of Noisy Independent Component Analysis (NICA). For our work, the impedance maps are considered as a mixture information. Then, source signals containing the fault features are obtained by the application of the Independent Component Analysis regarding the noise. A wavelet decomposition is then used and operates as a noise reduction operation. Jensen-Shannon (JSD) divergence is then proposed for the crack detection. Thanks to a theoretical derivation, the fault severity estimation is obtained. The performances are evaluated and the superiority validated regarding other techniques already used in the literature. The performances limits are evaluated for noise varying environments and the optimal diagnosis is obtained for several incipient cracks.

INDEX TERMS Non destructive crack diagnosis, detection, estimation, independent component analysis, wavelet denoising, Jensen-Shannon divergence, Kullback-Leibler divergence, CUSUM.

I. INTRODUCTION

In nowadays industrial process, safety is one of the major concern [1]. As an example, for industry transportation (planes, trains, vehicles, ...), chemical industry (chemical tanks, pipelines, ...) or electrical industry (nuclear power plants, renewable energy plants, ...) abnormal behaviors can lead to tragic phenomena as crashes, fires, or other more serious accidents. These abnormal operations can be due to electrical or mechanical failures that have to be predicted or at least detected at their early stage to be monitored. Among the mechanical failures, material cracks in the systems' structure are very dangerous for the process robustness. Regular inspection of the system is necessary to allow early detection of incipient cracks and then monitoring without stopping the system [2], [3]. Thus, Non-destructive testing (NDT) are mandatory. According to the American Society for Testing and Materials (ASTM), NDT is defined as: "*the development and application of technical methods to examine materials or components in ways that do not impair future usefulness and serviceability in order to detect, locate, measure and evaluate discontinuities, defects and other imperfections; assess integrity, properties and composition; and measure geometrical characteristic*" [4]–[6]. To proceed to this evaluation, model-based and data-driven methods which depend on different kind of information (electrical signals, acoustic emissions, electromagnetic fields, radiology information, multiple data types ...) can be considered [7], [8]. Model-based methods have to face with high processing complexity, long computational time, accurate modeling and

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Tucci.

design, ...These sophisticated methods are difficult to be obtained for all the system operating conditions. In fault detection and diagnosis (FDD) process, data-driven methods have shown their efficiencies in the extraction and analysis of fault features. Signal analysis-based methods which depend on the priori information of signals can be attributed to the data-driven methods. With these data-driven methods, the fault detection turns into a pattern recognition problem, which is often based on statistical and neural analysis approaches. In the last decade, Multivariate Analysis (MVA)-based and machine learning-based methods have received significantly increasing attention. As an example, for the fault detection in high-dimension data space, the Principal Component Analysis (PCA)-based techniques, Partial Least Squares (PLS)-based techniques, Fisher Discriminant Analysis (FDA)-based techniques and so on have been proposed and widely applied. In recent years, some new dimension reduction techniques-based methods have been developed, for example, the probability-relevant PCA (PRPCA)-based method has been proposed and applied in the incipient fault detection for the high-speed trains [9], probability density estimation and Bayesian causal analysis based method has been proposed for fault detection and fault backtracking [10], and a novel dual robustness projection latent structure method based on the $\mathcal{L}_1$ norm has been proposed to improve the robustness of the PLS method and validated in Tennessee Eastman process [11]. The machine learning-based methods such as Support Vector Machine (SVM), expert system, deep learning, and Artificial Neural Network (ANN) also show their potentials for fault detection and fault diagnosis. Moreover, some hybrid methods have also been developed. For example, a novel fault detection method based on Wavelet Packet Decomposition (WPD) coupled with a bilayer Convolutional Neural Network (biCNN) denoted as WPD-biCNN has been proposed in fault detection for industrial system and its fault detection ability has been shown compared to Back Propagation Neural Network (BPNN), SVM, CNN, WPD-SVM, biCNN [12]. In other applications, the method combining Kernel Principal Component Analysis (KPCA) and K-Nearest Neighbors (KNN) has been applied for fault diagnosis of bearings and shown a good accuracy [13].

The most typically used for crack detection in conductive material is based on Eddy current testing [14]. With this technique, the inhomogeneity due to the defect will cause impedance shifts or variations compared to the initial health condition of the material and then the fault can be detected. However, its detection capability can be limited by the fault severity of the cracks and also the internal or external environment. For the serious cracks, who lead to obvious changes in the measured material impedance, the detection is quite easy considering a basic analysis. In the case of smaller (minor) cracks, that can be considered as incipient faults, the evolution of the impedance value is weaker and the impedance measurement variation is largely affected by the external environment, the roughness of the material surface, the internal noise, the sensor noise sensitivity, ... [6], [7]. The weak impedance caused by the existence of minor cracks can be masked by these disturbances. Thus, the fault signature compared to the background noise and disturbance leads to an ECT signal that cannot highlight the presence of the crack.

In [7], J. Harmouche *et al.* have proposed to use of the Kullback-Leibler Divergence (KLD) for detecting the minor cracks in the ECT signal within a nuisance level equivalent to a 20dB signal to noise ratio (SNR). This divergence measure allows to evaluate the differences between the probability density functions of the reference signal and the testing one. High efficiency was obtained for the detection of incipient faults compared to other classical statistical measurements like the statistical moments with order 1 to 4, the Hotelling's test $\mathbf{T}^2$ or the squared prediction error (SPE). This divergence method, based on the relative entropy, has shown its superiority [7], [15]–[17] in the particular case of incipient fault detection. Its main benefit is its ability to consider the whole signal information but its performances are somewhat is limited within the presence of high noise levels.

In [18] M. Basseville *et al.* has introduced the Cumulative sum (CUSUM) as an algorithm based on the likelihood between the samples. Cumulative sum is one widely used control chart for detecting and locating a small shift in a process [19], it outperforms of the Shewhart control chart for detecting the small change in a small signal to noise ratio [20]. Moreover, its efficiency for the incipient fault detection has been proved in [21] for multiple phase electrical systems. In this paper, the detection capability of CUSUM for the incipient material cracks in high perturbation level is studied, and its detection performance is compared with the KLD ones proposed in [7].

Recently in [22], [23] the Jensen-Shannon divergence (JSD) has been proposed for fault diagnosis purpose. JSD is a sensitive technique based on Shannon entropy excess of a couple of distributions with respect to the mixture of their respective entropies [24]. It is widely applied in the anomaly change detection, such as bio-informatics, genome comparison, protein surface comparison, image process [24], [25]. In recent works [22], X. Zhang *et al.* have proposed to use Jensen-Shannon divergence for incipient fault detection and diagnosis (estimation) in high noise levels and its performances in noisy cases for an autoregressive system has been proved. One major benefit in diagnosis using these methods is that not only local modifications in a signal are considered but the global information in the overall signal are taken into account for the diagnosis decision. Other techniques like the Wasserstein distance can also be considered for fault detection. This distance measures the work required to transport the probability mass of a data distribution. It takes into account the metric space but it is mentioned in the literature [26], [27] as insensitive to small wiggles, not a smooth functional and not robust. Then, in the case of incipient fault detection and estimation the Wasserstein distance will not be the most suitable.

In the present work we propose to use Jensen-Shannon Divergence as the feature analysis technique in a particular

framework to do the incipient cracks detection and diagnosis (Fault estimation) with the ECT experimental data particularly in noisy conditions. Then, we compare the obtained performances with the CUSUM and the KLD ones to state with the efficiency of this technique.

For now, the evaluation of the crack diagnosis is obtained in the original ECT signal domain [7]. Nevertheless, as previously mentioned, the ECT impedance signal is modified with the presence of the crack, meaning that the crack induce an additional information in the healthy signal. The faulty ECT signal can then be considered as a mixture of these two information. In this paper we propose to develop this idea that for crack detection, the source signals can be considered as two types of information, the healthy surface information and the crack one. We propose to proceed the fault detection and diagnosis after a blind source separation. Independent Component Analysis (ICA) is a widely used technique for blind source separation (BSS) [28]–[31]. Most of the ICA application is regardless of the noise. However, in real-life applications, the noise information is always remaining as sensor noise or source noise. The ICA model considering noise is named NICA (Noisy Independent Component Analysis) [32]. Based on this separation we can consider a famous technique for noise reduction in the literature as the wavelet decomposition [33]–[35]. This time-scale analysis decomposition method allows to reduce the high-frequency noise information with a common procedure. It can then be applied as a signal conditioning method on the separated noisy source signals obtained from the NICA. With these conditioned features, the Jensen-Shannon divergence is used for incipient fault detection based on the remaining source signals. Thus, the detection capabilities of our proposed method with the influence of the perturbations is discussed and the performances for incipient cracks are compared with the Kullback-Leibler divergence and CUSUM ones.

The final part studied in this paper is the fault severity estimation. This latter is necessary for qualifying the fault level and providing recommendations in the decision making process for the system safety [1], [36]. However, most usual statistic technics fails for estimating the fault severity particularly for incipient ones [36], [37]. We propose in this work, an estimation model depending on the Jensen-Shannon Divergence. This model is derived taking into account our particular feature space. The estimation accuracies are evaluated and the obtained results validates the good estimation performances compared to the approximated real fault severities.

This paper is organized as follows. Section II first describes the problem statement of minor cracks (considered as incipient fault) detection and diagnosis. Then, our proposed detection scheme is described in section IV and the theoretical background is derived. Section V provides the detection and estimation study for our proposed detection scheme. The performances are then studied in section VI. Detection performances for the minor cracks using the proposed method are first studied and compared to those in the literature for high perturbation levels. Secondly, the fault estimation is evaluated. The accuracy of the proposed estimation model is validated proving the good performances of our proposal. Section VII, concludes the paper.

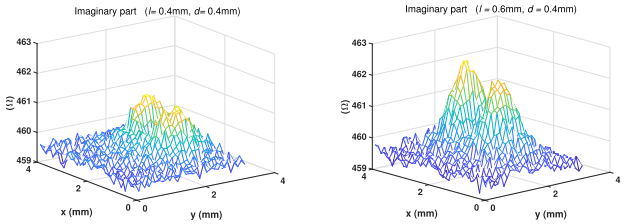


FIGURE 1. Imaginary impedance of the ECT map for large cracks.

II. PROBLEM STATEMENT

The detection and estimation for incipient cracks is a tedious problem [7], [38]. ECT is one of the physical principle applied for detection purpose [2]. It is based on the application of the principle of electromagnetic induction in a conductive material. Using an energized coil, an eddy current flowing the material can be generated. Measuring the instantaneous variations of this current leads to the evaluation of the material impedance. The existence of a surface crack will change the eddy current flow in the material and then induce variations in the impedance. Thus, the evolution of the material impedance with the presence of the crack denoted Z_f can be measured and written as:

$$Z_f = Z_h + \Delta Z = Z_h + \tilde{a} Z_h = R_f + jY_f \tag{1}$$

where R_f and Y_f are the real impedance and imaginary impedance parts of the complex impedance measured in faulty condition on different locations of the material. The remaining ΔZ is the variation on the healthy impedance values Z_h caused by the crack occurrence and $\tilde{a}$ corresponds to its severity. Measured on several positions of the material, these values leads to vectors. As an example, we plot in Figure 1 the evolution of the imaginary part of the impedance for a squared conductive plate containing cracks with $400\mu m$ depth and two different lengths $400\mu m$ and $600\mu m$. These two cracks size are seen as serious cracks: they leads to obvious changes in the measured material impedance. With such an impedance variation, these cracks can then be easily detected with a basic analysis.

Nevertheless, in the case of smaller cracks, the evolution of the impedance value is weaker. Basically, it does not significantly highlights the presence of these minor cracks. The information corresponding to the presence of the crack is small and partially masked by the presence of the environmental nuisance due, for example, to the roughness of the surface of the considered plate. As an example, Figure 2 shows the imaginary part of the impedance for four cracks with length or depth equal to $100\mu m$ or $200\mu m$. The main issue here leads to find a suitable detection method that can efficiently detect the minor cracks (namely incipient cracks).

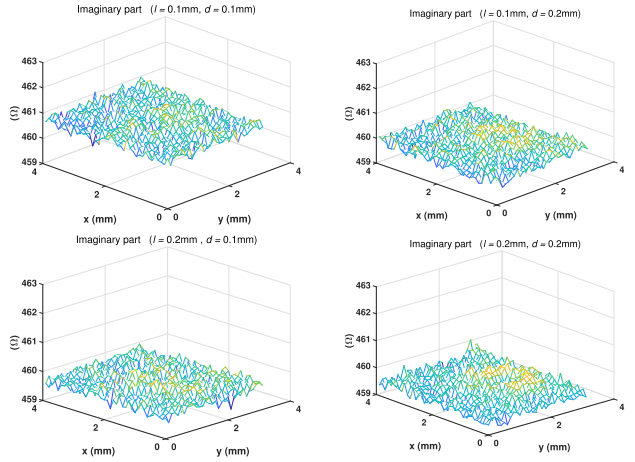


FIGURE 2. Imaginary impedance of the ECT map for incipient cracks.

The tedious task will be to ensure this detection with a good robustness even at a high perturbation levels.

In the previous work, most of the existed literatures give contributions for large cracks diagnosis such as Principal Component Analysis (PCA) [39], Independent Component Analysis (ICA) [40] or Fisher Discriminant Analysis (FDA) [41], etc.... However, for incipient cracks, that can be considered as incipient faults, the diagnosis is more tedious: the weak evolution caused by the incipient cracks can be masked by the disturbances such as external environment, the roughness of the material surface, the internal noise, the sensor noise sensitivity, ... Some hybrid methods for incipient fault detection, based on dimension reduction techniques, are not appropriate in this case for the measured signals from the ECT process.

In [7], the Kullback-Leibler divergence have been applied for incipient cracks detection. Its capabilities have been proved to be more efficient than other current statistical functions such as the mean, the variance, the skewness and the kurtosis when the perturbation level corresponds to a signal to noise ratio (SNR) equivalent to 20dB. Nevertheless, the detection of the incipient cracks ($\leq 200\mu m$) in a more severe noisy environment is not satisfying and has to be improved. The goal of this work is to propose a solution to tackle this problem and show how it is possible to improve the detection of the incipient crack sizes even while the noise widely affects the measurement and increases detection complexity. Furthermore, for these detected cracks, we propose to estimate the crack severity for diagnosis purpose with a good accuracy even for noisy environments.

For our work we will consider experimental data obtained from the same ECT laboratory system described in [7]. The cracks are created in plate conductive material. They consists in rectangular notches of small sizes created using Electro Discharge-Machining (EDM). Each notch has an opening size of $100\mu m$. Different cracks depth (d) and length (l) values are considered for this study. We mainly focus on incipient cracks such as l or d are equal to $200\mu m$ and $100\mu m$ with several combination. The perturbation level will be considered as equivalent to several signal to noise ratio in the range $[0dB, 20dB]$.

III. NOTATIONS AND ACRONYMS

The paper main abbreviations and notations are given in the following tables.

TABLE 1. Notations.

N	Sample size
Z	Observation
R	The real part impedance
Y	The imaginary part impedance
$\hat{}$	Estimated function
U^*	Healthy and noise-free information U
U_h	The information U computed in healthy condition
U_f	The information U computed in faulty condition
$I(.)$	Kullback-Leibler information
$D(.)$	Divergence value
h	Threshold
S	Separated source signal
W	Demixing matrix
V	Additive noise vector
X_f	Faulty source signal before filtering
X_h	Healthy source signal before filtering
F	Fault matrix before filtering
U^{fil}	Information U obtained after filtering
μ	Mean of the distribution
σ	Variance of the distribution
a	Fault severity on the filtered source signal
l	Length of the crack size
d	Depth of the crack size
C_ξ	Crack sizes, $\xi = \{1, 2, 3, 4\}$

TABLE 2. Abbreviations and acronyms.

NICA	Noisy Independent Component Analysis
ICA	Independent Component Analysis
PCA	Principal Component Analysis
FDA	Fisher Discriminant Analysis
BSS	Blind Source Seperation
NDT	Non-Destructive Testing
KLD	Kullback-Leibler divergence
JSD	Jensen-Shannon divergence
DWT	Discrete Wavelet Transform
CUSUM	Cumulative Sum
ECT	Eddy Current Testing
SNR	Signal to Noise Ratio
pdf	Probability Density Function
P_{FA}	Probability of False Alarm
P_{MD}	Probability of Missed Detection
P_D	Probability of Detection

IV. DIAGNOSIS PROCESS AND METHODOLOGY

For our study, the complex impedance information is used for the detection procedure. We propose to use a four steps procedure to evaluate the detection and diagnosis: Modelling, Preprocessing, Feature extraction and Feature analysis.

In this section, we first remind the proposed four steps' tuning proposed in the literature for using the Kullback-Leibler

divergence. Based on this knowledge, we propose the evolutions of some steps in the methodology in order to improve the final detection and diagnosis results for the incipient cracks in the noisy environment.

A. BASIC PROCESS USING KULLBACK-LEIBLER DIVERGENCE

In [7] J. Harmouche et al consider that the four main steps for the crack detection can be:

- Modelling: Data driven. The process knowledge is obtained from the experimental data. In that case, the use of the impedance imaginary part is proposed for its better sensitivity to the presence of the crack.
- Preprocessing: Time series analysis. The impedance map measured for the healthy or faulty conditions are converted into data vectors to be evaluated. In the faulty case several crack size are considered.
- Feature extraction: statistical analysis. The statistical distributions or probability density functions (pdf) of the vectors are evaluated. These pdf are obtained using kernel density estimations [42].
- Feature analysis: divergence study. The Kullback-Leibler divergence between the healthy and faulty distribution is evaluated. Kullback-Leibler divergence is popular for its benefit in measuring minor differences between two probability distribution without any type assumption [43]. For, two continuous probability density functions (pdfs) $p(x)$ and $q(x)$, respectively corresponding to the reference healthy pdf and the testing, the Kullback-Leibler (KL) Information I of $p(x)$ with respect to $q(x)$ is defined as:

$$I(p||q) = \int p(x) log \frac{p(x)}{q(x)} dx \tag{2}$$

The Kullback-Leibler divergence is obtained as the symmetric operation of the KL information [15]:

$$D_{KL}(p, q) = I(p||q) + I(q||p) \tag{3}$$

For this detection process, the decision is obtained comparing the KL value to a threshold settled for a given false alarm probability P_{FA}. One main drawback of this process is the wide sensitivity of the decision to the environmental disturbance: below $SNR = 20$dB the smallest crack sizes are not efficiently detected. Two ways can be considered for improving this process. The first one consists in evaluating the presence of the crack using other statistical tools. The second one consists in doing the crack evaluation in more adapted feature space where the influence of the noise is lowed.

We propose in this work to explore both of the solutions using other specific techniques. The first solution is based on the use of Cumulative Sum (CUSUM) and the second one is focused on the fault evaluation in the NICA feature space using Jensen-Shannon Divergence and Wavelet transformation. These two processes are described in the following.

B. IMPROVED PROCESS USING CUMULATIVE SUM

For this process, we propose to keep unchanged the modeling and the preprocessing steps and focus on the feature extraction and analysis ones. For the feature extraction, the impedance data vectors in healthy and faulty conditions are concatenated and directly studied. For the feature analysis, we propose to use the CUSUM algorithm that is usually efficient in noisy environments. Nevertheless its accuracy for such incipient cracks has not been studied.

The CUSUM is one widely used control chart in detecting and locating the small shift in a process [19]–[21]. The CUSUM which based on the Gaussian distributed process [18], [21] is used in this paper, it is defined as the sum of the sufficient statistics c_i such as:

$$\mathsf{C}_k = \sum_{i=1}^{k} \mathsf{c}_i \tag{4}$$

where $k \in \{1, \ldots, N\}$ is the number of samples size and c_i is considered in the case of mean changes (c_{i_μ}) or variance changes (c_{i_σ}) in the considered signal. As shown in [7], there is not an obvious change in the mean for the incipient crack sizes. So, in this paper, the CUSUM algorithm for detecting change shifts in the signal variance is considered such as:

$$\mathsf{c}_{i_\sigma} = \left(\frac{\sigma_h}{\sigma_f}\right) + \left(\frac{1}{\sigma_h^2} - \frac{1}{\sigma_f^2}\right) \times \frac{(x_i - \mu_h)^2}{2} \tag{5}$$

where μ_h, μ_f, σ_h and σ_f are respectively the mean and the standard deviation value in healthy and faulty conditions. The sample x_i is the current measure for the ith observation of the signal. The CUSUM decision function is then given as:

$$G_k = \left(\mathsf{C}_k - \min_{1 \le j \ge k} (\mathsf{C}_j)\right) \tag{6}$$

In the fault detection process, G_k is compared to a given detection threshold h_c. The selection methodology of the CUSUM threshold h_c is discussed in [44]. It can be computed using the sequential probability ratio test [18] or other probabilistic conditions. In in this paper, this threshold h_c is chosen as the maximum of the healthy CUSUM values. This mean that the probability for falsely claiming a fault is null ($P_{FA} = 0$).

In different noise conditions the CUSUM is computed and the performances evaluated.

C. IMPROVED PROCESS USING JSD IN THE NICA FRAMEWORK

The detection for incipient fault in a noisy environment is difficult. The fault can be seen as an hidden additional information in a healthy signal can be partially or totally masked by this nuisance [45], [46]. For the incipient cracks diagnosis, we propose a new methodology based on the four main steps introduced above. For this proposal, these main steps are particularly tuned using the combination of several signal processing tools with respect to a global fault

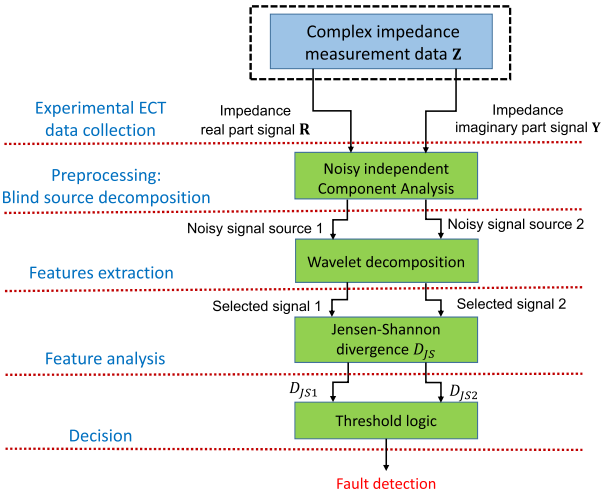


FIGURE 3. Proposed detection process.

detection methodology considering a data-driven approach summarized in Figure 3.

As described in this scheme, our work is based on the complex impedance measurement data (real and imaginary part). At first, the collected impedance data are arranged into vectors to be conditioned and then analysed. At second, Noisy Independent Component Analysis is applied as the preprocessing operation for obtaining the blind source separation based on two measured signals: the real impedance signal and the imaginary one. Then two separated source signals are obtained. Based on these latter, we proceed at third for each one to the features extraction. In this operation, the wavelet decomposition is considered and help to partially lower the noise of the sources. At fourth, the feature analysis using Jensen-Shannon Divergence is applied for both of the selected signals. Finally, the calculated divergence values are compared to the corresponding threshold to decide for the fault detection. Then, the fault can be estimated based on a proposed model.

In the following, we detail the theoretical background for processing these different steps.

1) MODELLING - DATA-DRIVEN

The collected ECT data leads to complex impedance values Z with its real part R and the imaginary one Y. For the plate conductive material, faulty and healthy zones are evaluated to respectively obtain data with and without the presence of the crack information. This impedance information are organised in separate vectors to be evaluated.

2) PREPROCESSING - BLIND SOURCE DECOMPOSITION

As described in the equation 1, the presence of the crack can be viewed as an additional information on the healthy baseline impedance. Based on this knowledge we propose to try to separate in different independent sources the studied signals. Independent component analysis (ICA) is a technique allowing to find a solution for blind source separation (BSS) problems. It assume that the given observations Z can be written as the generic model $Z = AS$, where A is a squared invertible mixing matrix and S is the desired source signals [28]. Using ICA, we can find the optimal demixing matrix denoted W so that the separated source signals $\hat{S}$ can be obtained such as:

$$\hat{S} = WZ \tag{7}$$

where $W = A^{-1}$.

Considering noisy systems, two models can be denoted: one with sensor noise and the other one with sources noise [28], [29]. In our work, the FASTICA algorithm [47], [48] is used for source separation on the two parts of the complex impedance signal R and Y. The optimal demixing matrix is calculated in the case of healthy reference signals and then reused for faulty ones. Thus, we obtain in the faulty case two independent components with slight differences compared to the healthy ones. These both components also contain the remaining noise must be evaluated to analyse the fault information. We need then to extract features with partially reduced high frequency noise without affecting the fault information.

3) FEATURE EXTRACTION - WAVELET TRANSFORM

Based on the obtained two noisy source signals, we propose the use of the wavelet transformation to partially reduce the high frequency noise without affecting the low frequency fault information. The wavelet decomposition is a widely applied transformation for overcoming the limitations in time-frequency resolution of the classical Fourier transform and its extended versions. It has been applied in different kind of applications and has been proved to be efficient for signal noise reduction [32], [34], [35]. The continuous wavelet transform (CWT) of a target signal $\hat{S}(t)$ is defined as:

$$\langle \hat{S}, \psi^{\diamond}_{r,\mathsf{b}} \rangle = \int_{-\infty}^{+\infty} \hat{S}(t)\psi^{\diamond}_{r,\mathbf{b}}(t)dt \qquad r, \mathsf{b} \in \mathbb{R};\ \ r \neq 0 \tag{8}$$

where r is the scaling parameter which is reciprocal of frequency and b indicates the translation along with the time axis, $^{\diamond}$ is the complex conjugate operator. The function $\psi^{\diamond}_{r,\mathsf{b}}$ is defined as the scaled and translated version of the mother wavelet denoted ψ, which represent the detailed high frequency part of the signal. It can be written as:

$$\psi^{\diamond}_{r,\mathsf{b}} = \frac{1}{\sqrt{r}}\psi^{\diamond}\left(\frac{t-\mathsf{b}}{r}\right)dt \tag{9}$$

In the one-dimensional CWT, the signal is analysed by functions obtained while scaling and translating the mother wavelet. Let's denote the scaling parameter $r = r_0^j$ and the translation parameter $\mathsf{b} = \mathsf{k}r_0^j\mathsf{b}_0$, where j, k corresponds to their different decomposition levels. The values r_0 and b_0 respectively initialize the r and b values. They should be as small as possible to ensure the accuracy of the reconstructed signal. In this paper, we set $r_0 = 2$ and $\mathsf{b}_0 = 1$.

For discrete wavelet transform (DWT), a scaling function (ϕ) representing the smooth trend part of the signal (low frequency) is defined according to [49] as:

$$\phi_{j,\mathsf{k}}(t) = 2^{-\frac{j}{2}}\phi(2^{-j}t - \mathsf{k}) \tag{10}$$

The signal $g(t)$ decomposed by DWT can then be written as a combination of the scaling and wavelet functions [49]–[51]:

$$\hat{S}(t) = \sum_{\mathsf{k}=-\infty}^{\infty} L_{\mho}(\mathsf{k})\phi_{\mho,\mathsf{k}}(t) + \sum_{j=1}^{\mho}\sum_{\mathsf{k}=-\infty}^{\infty} D_j(\mathsf{k})\psi_{j,\mathsf{k}}(t) \tag{11}$$

where $L_{\mho}$ is the approximation coefficient at level $\mho$, and D_j is the detail coefficient at level j. These coefficients are:

$$L_{\mho}(\mathsf{k}) = \int_{-\infty}^{\infty} \hat{S}(t)\cdot\phi_{\mho,\mathsf{k}}(t)dt \tag{12}$$

$$D_j(\mathsf{k}) = \int_{-\infty}^{\infty} \hat{S}(t)\cdot\psi_{j,\mathsf{k}}(t)dt \tag{13}$$

$L_{\mho}$ and D_j are respectively the inner product of $\hat{S}(t)$ with the scaling function $\phi(t)$ or the wavelet function $\psi(t)$.

Basically, the approximation coefficient leads to a low-pass operation and the details corresponds to a high-pass one.

In our work the wavelet transform is applied to the noisy source signals in healthy and faulty conditions. The highest level of the low-pass approximations is retained as the noise reduced signals for both of the sources.

4) FEATURE ANALYSIS - JENSEN-SHANNON DIVERGENCE

We propose the application of the Jensen-Shannon divergence (JSD) for analysing the features. JSD is the increment of the Shannon entropy (S_E) for two considered distributions p and q. The computation of JSD can be written as the value D_{JS} such as:

$$D_{JS}(p,q) = S_E\left[\frac{p+q}{2}\right] - \frac{S_E(p)+S_E(q)}{2} \tag{14}$$

JSD is then a symmetric operation based on the Kullback-Leibler Information (I) related to the mean mixture distribution M between p and q. It can be denoted as:

$$D_{JS}(p,q) = \frac{1}{2}I(p||M) + \frac{1}{2}I(q||M) \tag{15}$$

where M is a mixture distribution whereas $M = \frac{1}{2}(p+q)$.

Based on the obtained D_{JS} values we can proceed to the decision leading or not to the fault detection by comparing D_{JS} to a well established threshold value h such as:

- if $D_{JS} > \mathsf{h}$ a crack is detected: the hypothesis $\mathcal{H}_1$ is validated, the system is faulty
- if $D_{JS} < \mathsf{h}$ no crack is detected: the hypothesis $\mathcal{H}_0$ is validated, the system is healthy

This threshold has to be properly tuned according to D_{JS}. To reach the optimized detection performances, h is obtained by minimising the probability of false alarms (P_{FA}) and maximising the probability of detection (P_D).

With the detection knowledge, we can proceed to the crack size estimation. For this purpose, we need to derive the theoretical behavior of the JSD with the presence of the crack and then compute the estimated value of the fault severity. This theoretical study is given in the following section V.

V. CRACK DIAGNOSIS THEORETICAL STUDY

A. DEFINITIONS

Let's consider that the healthy real impedance vector R_h and the imaginary one Y_h are assumed to be generated by adding a numerical white noise to the original measurements:

$$R_h = R_h^* + V_R \quad Y_h = Y_h^* + V_Y \tag{16}$$

where R_h^* and Y_h^* are the healthy and noisy-free measurement vectors, V_R and V_Y are additive noise vectors.

After the applying the NICA on the healthy data, the demixing matrix W with elements $w_{\kappa,\ell}$ and two healthy noisy source signals $\hat{S}_1^{\,h}$ and $\hat{S}_2^{\,h}$ are obtained such as:

$$\begin{bmatrix}\hat{S}_1^{\,h}\\ \hat{S}_2^{\,h}\end{bmatrix} = \begin{bmatrix}w_{11} & w_{12}\\ w_{21} & w_{22}\end{bmatrix}\begin{bmatrix}R_h^* + V_R\\ Y_h^* + V_Y\end{bmatrix} \tag{17}$$

For the faulty input data vectors, we use the same demixing matrix W, and extract the faulty sources signals $\hat{S}_1^{\,f}$ and $\hat{S}_2^{\,f}$:

$$\begin{bmatrix}\hat{S}_1^{\,f}\\ \hat{S}_2^{\,f}\end{bmatrix} = \begin{bmatrix}w_{11} & w_{12}\\ w_{21} & w_{22}\end{bmatrix}\begin{bmatrix}R_f^* + V_R\\ Y_f^* + V_Y\end{bmatrix} \tag{18}$$

Each of the noisy faulty sources which contain most fault features is assumed to be defined as a signal X_f such as $X_f = X_h + F$, where X_h is corresponding to the healthy information and F is the incipient fault ones denoted as $F = [f_{11},\ldots,f_{1i},\ldots,f_{1N}]$ where i is the instant time sample and N is the total number of samples. Considering that the elements of the vector X_h are x_{1i} and that $X_h = X_h + \tilde{a}X_h$, the fault vector can be written as $F = \tilde{a}\times[0,\ldots,x_{1b},\ldots,x_{1N}]$, such as $\tilde{a}$ corresponds to the fault severity and b is the instant time when the fault occurs.

After applying DWT, the noise contained in the faulty and the healthy sources are reduced, then the filtered faulty source signal can be defined as X_f^{fil} whereas $X_f^{fil} = X_h^{fil} + F^{fil}$. So, $X_h^{fil} = [x_{11}^{fil},\ldots,x_{1i}^{fil},\ldots,x_{1N}^{fil}]$ are the healthy filtered sources and $F^{fil} = [f_{11}^{fil},\ldots,f_{1i}^{fil},\ldots,f_{1N}^{fil}]$ the filtered incipient fault information. Moreover, $F^{fil} = a\times[0,\ldots,x_{1b}^{fil},\ldots,x_{1N}^{fil}]$ with a the filtered fault severity.

B. ASSUMPTIONS

Jensen-Shannon divergence is applied as the feature analysis technique to compare the pdf of the filtered source signal in healthy conditions X_h^{fil} and faulty ones X_f^{fil}.

Let's consider that the mean and the variance of X_h^{fil} are respectively μ_h and σ_h^2. As shown in [7], the statistical mean criteria fails for detecting the incipient cracks with a surface area smaller than $0.02mm^2$ with a SNR equal to 20dB. So, for this work, we can assume that the center of the preprocessed

input data after the occurrence of incipient cracks in a high perturbation level is remained unchanged. Thus, we have:

$$\mu_f = \mu_h \tag{19}$$

The variance σ_f^2 of the filtered faulty signal can be written as:

$$\begin{aligned}\sigma_f^2 &= \mathbb{E}\left[\left(X_f^{fil} - mean(X_f^{fil})\right)^2\right] \\ &= \sigma_h^2 + \frac{1}{N}\sum_{i=1}^{i=N}\left(f_{1i}^{fil} - \frac{1}{N}\sum_{i=b}^{i=N} f_{1i}^{fil}\right)^2 \\ &\quad + \frac{2}{N}\sum_{i=1}^{i=N}\left(f_{1i}^{fil} - \frac{1}{N}\sum_{i=b}^{i=N} f_{1i}^{fil}\right)\left(x_{1i}^{fil} - \mu_h\right)\end{aligned} \tag{20}$$

We set that:

$$\frac{1}{N}\sum_{i=b}^{i=N} f_{1i}^{fil} = \frac{1}{N}\sum_{i=b}^{i=N} a \times x_{1i}^{fil} = a \times \mu_1 \tag{21}$$

Based on (21), some simplification in (20) leads to:

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{i=N}\left(f_{1i}^{fil} - \frac{1}{N}\sum_{i=b}^{i=N} f_{1i}^{fil}\right)^2 \\ &= \frac{a^2}{N}\left[(b-1)\mu_1^2 + \sum_{i=b}^{i=N}(x_{1i}^{fil} - \mu_1)^2\right] \\ &= c_1 a^2\end{aligned} \tag{22}$$

where c_1 is denoted as:

$$c_1 = \frac{1}{N}\left[(b-1)\mu_1^2 + \sum_{i=b}^{i=N}(x_{1i}^{fil} - \mu_1)^2\right] \tag{23}$$

For the second part of (20), we have:

$$\begin{aligned}&\frac{2}{N}\sum_{i=1}^{i=N}\left(f_{1i}^{fil} - \frac{1}{N}\sum_{i=b}^{i=N} f_{1i}^{fil}\right)\left(x_{1i}^{fil} - \mu_h\right) \\ &= \frac{2a}{N}\sum_{i=1}^{i=b-1}(-\mu_1)(x_{1i}^{fil} - \mu_h) \\ &\quad + \frac{2a}{N}\sum_{i=b}^{i=N}(x_{1i}^{fil} - \mu_1)(x_{1i}^{fil} - \mu_h) \\ &= 2c_2 a\end{aligned} \tag{24}$$

where c_2 is defined as:

$$\begin{aligned}c_2 = \frac{1}{N}[&\sum_{i=1}^{i=b-1}(-\mu_1)(x_{1i}^{fil} - \mu_h) \\ &+ \sum_{i=b}^{i=N}(x_{1i}^{fil} - \mu_1)(x_{1i}^{fil} - \mu_h)]\end{aligned} \tag{25}$$

Based on (22) and (24) we can rewrite (20) as:

$$\sigma_f^2 = \sigma_h^2 + c_1 a^2 + 2c_2 a \tag{26}$$

It is well known that the midpoint measure for two normal distributions which have the same mean is unimodal [52]. In our case, if we assume that one of the filtered components can be approximate as Gaussian distributed, the comparison of the healthy and the faulty signals can be done using this assumption. In that case the mixture distribution M is assumed as Gaussian. Its mean μ_M and variance σ_M^2 can then be computed as in [22]:

$$\mu_M = \mu_h \quad \sigma_M^2 = \frac{\sigma_h^2 + \sigma_f^2}{2} = \frac{2\sigma_h^2 + c_1 a^2 + 2c_2 a}{2} \tag{27}$$

C. DETECTION AND ESTIMATION THEORETICAL MODEL

Based on the definition of the Jensen-Shannon divergence given in (15), in Gaussian conditions, the specific expression of the JSD can be written as (28), shown at the bottom of the page. Then, combining this latter with (27) and (19), we obtain in (29), shown at the bottom of the page, the theoretical derivation of the JSD value that can be considered in our detection process. It can be noticed that this equation

$$D_{JS}(p,q) = \frac{1}{4}\left[log\frac{\sigma_M^4}{\sigma_h^2\sigma_f^2} + \frac{\sigma_h^2 + \sigma_f^2 + \frac{1}{2}(\mu_h - \mu_f)^2}{\sigma_M^2} - 2\right] \tag{28}$$

$$D_{JS}(p,q) = \frac{1}{4}\left[log\frac{4\sigma_h^4 + 8\sigma_h^2 c_2 a + (4\sigma_h^2 c_1 + 4c_2^2)a^2 + 4c_1 c_2 a^3 + c_1^2 a^4}{4\sigma_h^4 + 4\sigma_h^2 c_1 a^2 + 8\sigma_h^2 c_2 a}\right] \tag{29}$$

$$D_{JS}(a) = D_{JS}(a=0) + a \times \frac{\partial D_{JS}(a=0)}{\partial a} + \frac{a^2}{2} \times \frac{\partial^2 D_{JS}(a=0)}{\partial a^2} + \cdots \tag{30}$$

$$b_1(a) = \frac{\partial D_{JS}(a)}{\partial a} = \frac{1}{4}\left[\frac{8\sigma_h^2 c_2 + (8\sigma_h^2 c_1 + 8c_2^2)a + 12c_1 c_2 a^2 + 4c_1^2 a^3}{4\sigma_h^4 + 8\sigma_h^2 c_2 a + (4\sigma_h^2 c_1 + 4c_2^2)a^2 + 4c_1 c_2 a^3 + c_1^2 a^4} - \frac{2c_1 a + 2c_2}{\sigma_h^2 + c_1 a^2 + 2c_2 a}\right] \tag{31}$$

$$\begin{aligned}b_2(a) = \frac{\partial^2 D_{JS}(a)}{\partial a^2} = \frac{1}{4}[&-\frac{\left(8\sigma_h^2 c_2 + (8\sigma_h^2 c_1 + 8c_2^2)a + 12c_1 c_2 a^2 + 4c_1^2 a^3\right)^2}{\left(4\sigma_h^4 + 8\sigma_h^2 c_2 a + (4\sigma_h^2 c_1 + 4c_2^2)a^2 + 4c_1 c_2 a^3 + c_1^2 a^4\right)^2} + \frac{2c_1^2 a^2 + 4c_1 c_2 a + 4c_2^2 - 2c_1\sigma_h^2}{(\sigma_h^2 + c_1 a^2 + 2c_2 a)^2} \\ &+ \frac{(8\sigma_h^2 c_1 + 8c_2^2) + 24c_1 c_2 a + 12c_1^2 a^2}{4\sigma_h^4 + 8\sigma_h^2 c_2 a + (4\sigma_h^2 c_1 + 4c_2^2)a^2 + 4c_1 c_2 a^3 + c_1^2 a^4}]\end{aligned} \tag{32}$$

is directly linked with the fault severity value in the filtered feature space. In the following this equation can be used for estimating the fault severity in this feature space.

For our incipient fault detection case study, JSD evaluate very slight modifications in the probability distributions, then (29) can be seen as a function of a in the neighbourhood of zero. It is then infinitely derivable in the neighborhood of zero. So, the Taylor development of D_{JS} can be given following (30), shown at the bottom of the previous page, where its first and second order derivative are shown in (31) and (32), at the bottom of the previous page, respectively.

By limiting the Taylor development to its first two order derivatives, we obtain the quadratic equation (33) as:

$$D_{JS}(a) = D_{JS}(0) + a \times b_1(0) + \frac{a^2}{2} \times b_2(0) \tag{33}$$

This equation resolution allows to derive an estimated value of a denoted $\hat{a}$. In healthy conditions, the D_{JS} value is seen as the evaluation for fault severity $a = 0$. In that case the JSD value is assumed to be null ($D_{JS}(0) = 0$), then we obtain for (33) the solution given in (34).

$$\hat{a} = \frac{-b_1(0) + \sqrt{b_1^2(0) + 2 \times b_2(0) \times D_{JS}(a)}}{b_2(0)} \tag{34}$$

Assuming that in healthy conditions, the change corresponding to the fault occurrence a is seen as 0, the initialisation constants $b_1(0)$ and $b_2(0)$ can be simplified as in (35).

$$b_1(0) = 0 \quad b_2(0) = \frac{c_2^2}{2\sigma_h^4} \tag{35}$$

The estimation value of the fault severity in the filtered feature space can then be written as:

$$\hat{a} = \frac{2\sigma_h^2}{c_2}\sqrt{D_{JS}(a)} \tag{36}$$

Based on this latter equation (36), the fault severity can be estimated in the filtered feature space. We will assume that the fault evolution in the original feature space will follow the same behavior with a lowed weighted value considering the different linear operations applied in the process.

VI. CRACKS DETECTION AND ESTIMATION RESULTS

To validate this study, we consider experimental ECT data. For this data, complex impedance values obtained in the healthy and the faulty conditions are provided as impedance maps with size 40×32. We enlarge first each ECT map to a 420×100 one by increasing the healthy data size. With this, we can have sufficient data and are able to reduce the false alarm probability caused by the internal and external perturbations in healthy conditions. For this work we focus on incipient cracks. Then, our study is mainly validated for four selected minor cracks C_1 ($\mathsf{l} = 100\mu m$, $\mathsf{d} = 100\mu m$), C_2 ($\mathsf{l} = 100\mu m$, $\mathsf{d} = 200\mu m$), C_3 ($\mathsf{l} = 200\mu m$, $\mathsf{d} = 100\mu m$), (C_4 with size $\mathsf{l} = 200\mu m$, $\mathsf{d} = 200\mu m$). The environmental disturbance will be created with SNR varying from 0 to 20dB for each crack size.

For this work, we first evaluate the detection capabilities of the two selected methodologies from the literature within the references [7], [44]. For these detection techniques based on the Kullback-Leibler divergence or the CUSUM, we highlight their strength and limitations for the detection of such cracks in noisy environment. Then, in second, we evaluate the performances of our proposed methodology and highlight its benefits. For this latter, we apply the four step process of our proposal for the crack detection and then its size estimation (Fig. 3). In the first step, the complex ECT impedance data are considered. The FASTICA algorithm is then considered for the preprocessing step. It is applied on the two parts of the complex impedance signal for the blind source separation. Two separated noisy source signals are obtained. For the third step, the wavelet transform is then applied to these noisy signals. A Daubechies mother wavelet is considered with a decomposition level equal to five. The approximation based signals are then used as features for the fault detection. The last step is the feature analysis for the fault detection and estimation. The benefit of the Jensen-Shannon divergence is then highlighted and studied in several noise conditions.

A. DETECTION PERFORMANCE ANALYSIS

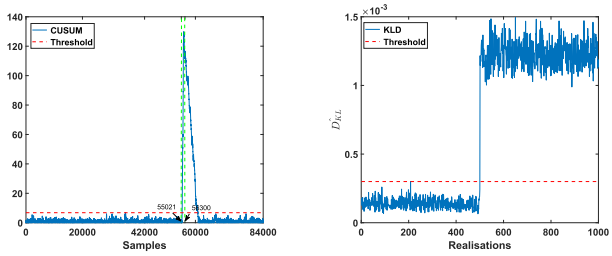
1) USING KLD AND CUSUM PROCESS

In [7], the detection scheme proposed by Jinane Harmouche *et al.* is based on the imaginary impedance values which are known to be most sensitive impedance part to the presence of the crack. Its detection advantage for the minor cracks within a nuisance level equivalent to a 20dB has been well shown compared to the four statistical moments with order 1 to 4. We first compare this approach with the Cumulative Sum (CUSUM) which is well known as an efficient technique for detecting signal change point. The fault detection using KLD is repeated many times. One detection is denoted as one realisation. For this comparison, we proceed to the evaluation of the KLD considering 500 realisations in healthy cases and idem for the faulty ones.

Note that the input data of CUSUM is a vector composed with 84000 samples: the first 42000 ones are obtained considering a healthy plate and the last 42000 samples were evaluated for a plate with the surface crack (faulty plate). For this work, that the faulty samples corresponding to the crack occurrence is located among the samples numbered from 55021 to 56300. The green vertical dash lines highlights these locations in the CUSUM results. The results studied in the following concerns the four previously mentioned cracks, but a highlight will be mainly focused on C_1 and C_4 to settle the comparison of the performances.

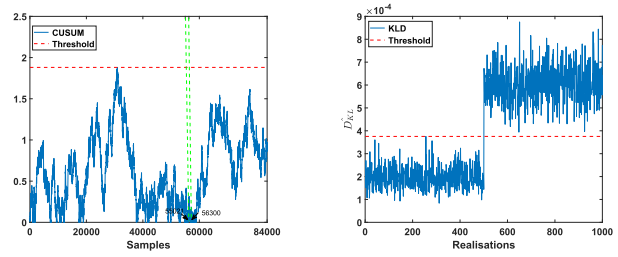
The detection results for 20dB are respectively displayed in Fig.4 and Fig.5 for the considered cracks. Note that the red dash line on the figures marks the detection threshold. For this first part of the study, it has been settled to the maximum of the healthy value (leading to $P_{FA} = 0$).

The detection capabilities of CUSUM and KLD for crack size C_4 in 20dB respectively given in Fig.4-a and Fig.4-b



(a) Detection using CUSUM (b) Detection using KLD

FIGURE 4. Fault detection using CUSUM and KLD for C_4 in 20dB.



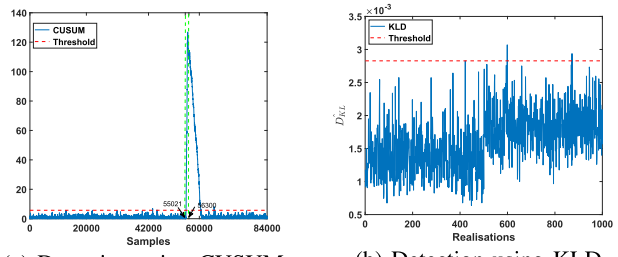
(a) Detection using CUSUM (b) Detection using KLD

FIGURE 5. Fault detection using CUSUM and KLD for C_1 in 20dB.

highlights that both methods can easily allow to detect such fault in this environment. One major benefit of the CUSUM here is its capability for obtaining the fault change point location. This can be useful for obtaining the location of the crack on the plate. Nevertheless, one can notice that the end location of the faulty data is not clear for this conditions.

Otherwise, for the lowest crack size C_1 in the same environment (20dB), CUSUM fails for the detection of the crack (Fig. 5-a) while KLD allows a successful detection: regarding the given threshold, the detection using KLD leads to a probability of detection $P_D = 1$, but for CUSUM $P_D = 0$. This proves the benefit for using KLD in minor fault detection compared to CUSUM in this environment: the changes occurred by the fault are too small to be detected using CUSUM for that conditions.

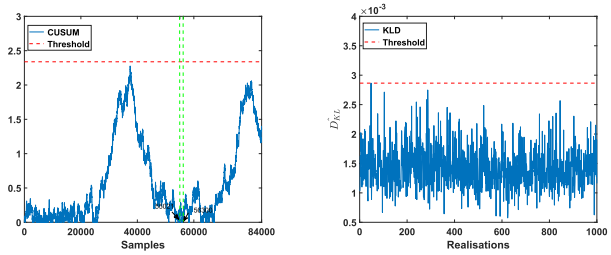
Considering now the changes in the environment where $SNR = 0$dB, we can show the detection ability of the two methods considering C_1 and C_4. The results are displayed on Figure 6 and 7.



(a) Detection using CUSUM (b) Detection using KLD

FIGURE 6. Fault detection using CUSUM and KLD for C_4 in 0 dB.

For the crack size C_4 (Fig. 6), it can be noticed the reduction of the efficiency of KLD for detecting such fault widely influenced by the noise. Lots of missed detection are observed ($P_D = 0.004$) with the settled threshold to reach $P_{FA} = 0$. In order to lower the missed detection, P_{FA} can be slightly increased. Indeed, an optimization between this two latter must be done to consider the best trade-off performances. Even with this optimization the performances are not good enough ($P_{FA} = 0.18$ and $P_D = 0.84$). For the CUSUM, the detection is still accurate, nevertheless the change point is not accurate for the end of the faulty data location. Compared to Fig.4-a, the noise does not influence a lot this crack size detection with the CUSUM.



(a) Detection using CUSUM (b) Detection using KLD

FIGURE 7. Fault detection using CUSUM and KLD for C_1 in 0 dB.

Their detection performances for the crack size C_1 are shown in Fig. 7, we can see that both of the two techniques fails in the detection of the crack. We summarize in the Table 3 the obtained optimal probabilistic detection performances for the crack size C_1 and C_4 in the noise conditions for $SNR = 0dB$ and $SNR = 20dB$.

TABLE 3. Cracks detection using CUSUM and KLD.

Crack Size	SNR=20 dB				SNR=0 dB			
	CUSUM		KLD		CUSUM		KLD	
	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}	P_D
C_1	0	0	0	1	0	0	0	0
C_4	0.11	0.78	0	1	0.11	0.77	0	0.004

It can be noticed that for both of the techniques, the detection capabilities are $P_D = 0$ for C_1 in 0dB. In this case, the crack severity is too low and the noise in the environment is too high for these detection techniques. For this crack size if the $SNR = 20dB$ the detection is possible using KLD. In the case of C_4, the detection performance using CUSUM is correct for both SNR values $P_D \backsim 0.77$ but not perfect and can be improved. Using KLD the detection completely fail ($P_D \backsim 0$) for $SNR = 0dB$ but is accurate enough ($P_D = 1$) for $SNR = 20dB$. Indeed, one of the challenges of our new proposal will be to show how it is possible allow the crack detection with better performances even for these severe noise environment conditions.

2) USING OUR PROPOSED ICA BASED JSD PROCESS

First, the detection performances of CUSUM in the framework of ICA-WAVELET for the incipient cracks are studied, and the results for incipient crack C_1 in $\{0, 5, 20\}$dB are given in Fig.8, Fig.9 and Fig.10. It presents that the detection is possible using ICA-WAVELET-CUSUM when the SNR is

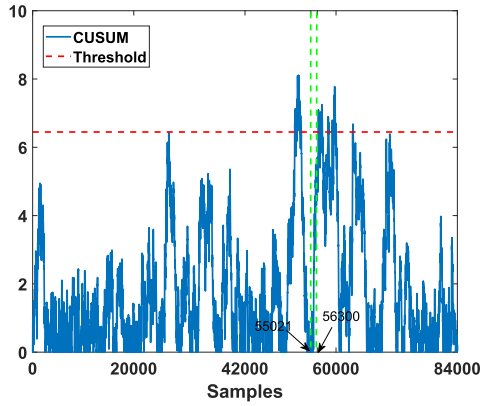


FIGURE 8. Detection performances of ICA-wavelet-CUSUM for C_1 in 0dB.

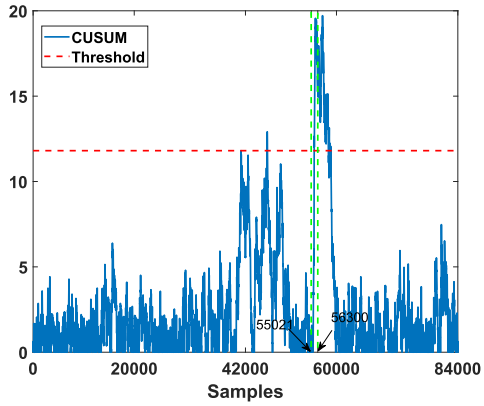


FIGURE 9. Detection performances of ICA-wavelet-CUSUM for C_1 in 5dB.

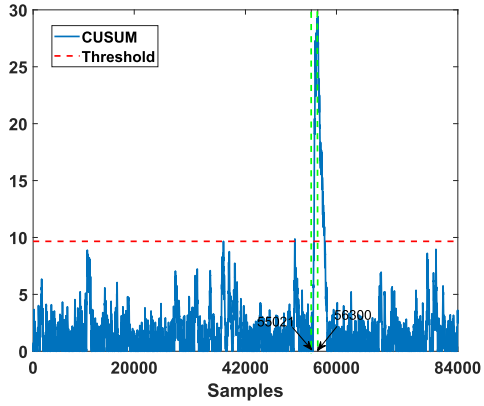


FIGURE 10. Detection performances of ICA-wavelet-CUSUM for C_1 in 20dB.

larger than 5dB, however it will cause a large probability of false alarm even when the SNR is 20dB.

The detection and false alarm probabilities in the case of ICA-WAVELET-CUSUM for incipient crack size C_1 and C_4 in several noise level are listed in Table 4. In that case the detection threshold is optimally obtained based on the healthy plate ECT evaluation. Compared to the performances obtained with the CUSUM technique without the ICA and wavelet feature space, some slight improvement are obtained in the detection performances for this two crack sizes but these are not sufficient. The results demonstrate that for crack

TABLE 4. Cracks detection using ICA-WAVELET-CUSUM.

	SNR=0 dB		SNR=5 dB		SNR=20 dB	
Crack Size	P_D	P_{FA}	P_D	P_{FA}	P_D	P_{FA}
C_1	0	0.05	0.55	0.05	0.58	0.03
C_4	0.76	0.1	0.785	0.093	0.8	0.09

C_4, the detection performance obtained even in 0dB are still good ($P_D \backsim 0.78$) but not perfect even with a correct false alarm probability ($P_{FA} \leq 0.1$). For the crack size C_1, the detection performances are bad for all the SNR conditions ($P_D \leq 0.5$).

For this other step of the study, in order to evaluate the performance of our proposal based on ICA-wavelet-JSD, we have computed the Receiver Operating Characteristic (ROC) curve [53], [54]. Then, we give the evolution of the Probability of detection (P_D) along with the probability of false alarm (P_{FA}) for all the considered SNR values and crack sizes. For these ROC curves, as it is worthless for engineers when $P_{FA} > 0.5$, for this study we will limit the representation of the ROC curves to $P_{FA} < 0.5$. In the following, the obtained performances of our proposal denoted ICA-wavelet-JSD are compared with the previously used KLD ones described in section IV-A but also with the ICA-wavelet-KLD. Figures 11 and 12 respectively display the results obtained for $C1$ and C_4.

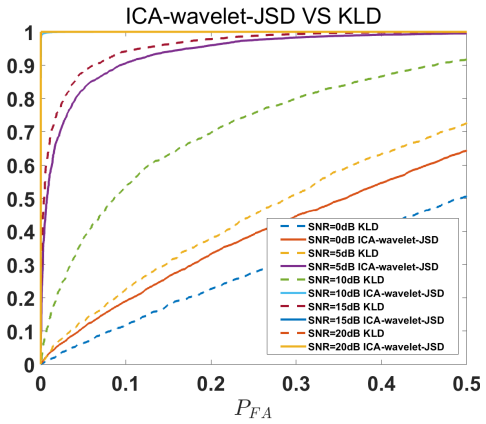


FIGURE 11. Detection performances of ICA-wavelet-JSD and KLD for C_1.

We can notice that the performances are improved along with the SNR (the noise level decreases) and the increase of the crack size. When the noise level is quite low ($SNR = 20dB$) the detection performances are perfect ($P_D = 1$ and $P_{FA} = 0$) for all the crack sizes.

When the noise level increases (SNR decreases) the performances are affected. Nevertheless, our proposal shows better performances than KLD. As an example considering a settled P_{FA} value as $P_{FA} = 0.05$, the P_D, in the same noise level and crack conditions, is higher for the ICA-Wavelet-JSD method than that for the KLD one.

Indeed, when the $SNR \geq 10dB$, the detection capability of ICA-wavelet-JSD is 100% (*i.e.* $P_D = 1$) from an extremely low P_{FA} value whatever the crack size.

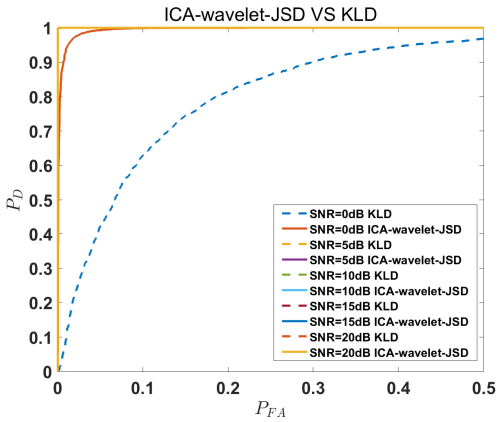


FIGURE 12. Detection performances of ICA-wavelet-JSD and KLD for C_4.

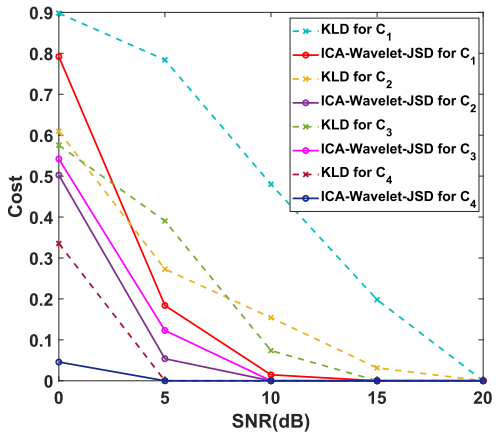


FIGURE 13. Detection performances of ICA-wavelet-JSD and KLD versus SNR.

The proposed methodology based on ICA-wavelet-JSD is then the most efficient one to properly detect the minor cracks C_1, C_2, C_3 and C_4 even in high environmental noise conditions. The results displayed in Fig.13, confirms the benefit of our proposal. In this figure we plot the detection performances obtained for the four sizes cracks in different perturbation level while minimizing the Bayes risk denoted as the cost function: the detection threshold is then optimal. The cost is here calculated as the sum of the optimal P_{FA} and P_{MD} values. The lower cost tends to the better detection performance. The results for each crack size demonstrate that the optimal detection performances of the two methods increases with the SNR increases. Moreover, for the same SNR values, ICA-wavelet-JSD always supports better detection performances than KLD for each crack size.

For further showing the detection advantage caused by Jensen-Shannon divergence, and really appreciate the benefit of the Blind source decomposition coupled with wavelet transform for the evaluation feature space, we compared the detection performances of ICA-wavelet-JSD and ICA-wavelet-KLD. Fig.14 shows the evolution of the Probability of detection along with the probability of false alarm for crack size C_1 and Fig.15 shows the detection performances for the crack size C_4. For these two figures, the SNR values inducing

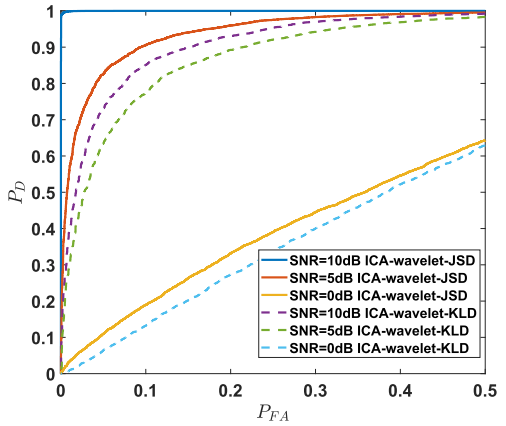


FIGURE 14. Detection performances of ICA-wavelet-JSD and ICA-wavelet-KLD for C_1.

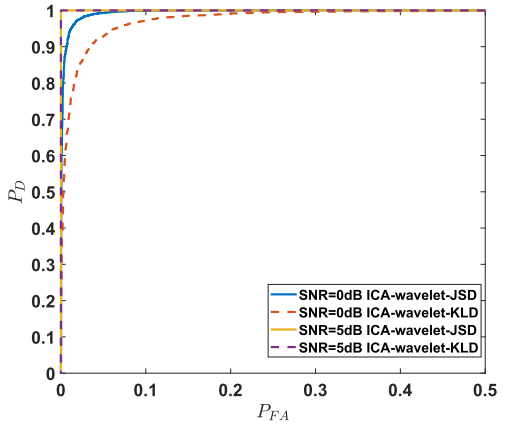


FIGURE 15. Detection performances of ICA-wavelet-JSD and ICA-wavelet-KLD for C_4.

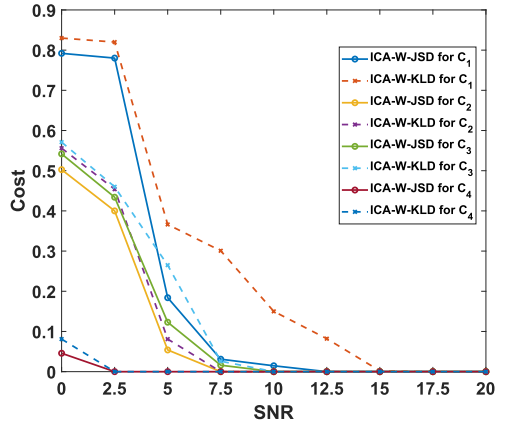


FIGURE 16. Detection performances of ICA-wavelet-JSD and ICA-wavelet-KLD versus *SNR*.

the highest noise severity (SNR = 0dB, 5dB and 10dB) and impacting the detection probability are remained selected for the figures representations. In figure Fig.16, the evolution of the cost is summarised for the different noise levels in the paper considered range ([0dB; 20dB]).

These latter results demonstrate that for a given probability of false alarm and in the same noise level, the detection performance of ICA-wavelet-JSD is better than

ICA-wavelet-KLD. More the crack size is severe and higher is the benefit of the JSD in the ICA-wavelet feature space compared to KLD. This proves then the detection advantage using Jensen-Shannon divergence compared to Kullback-Leibler one and highlight the importance of the blind source decomposition and the wavelet transform for the working feature space.

B. ESTIMATION PERFORMANCE ANALYSIS

Based on the good detection ability of our process, we propose in this section to give an estimation of the crack severity. This estimation is based on the proposed model given in (36). The estimated value $\hat{a}$ in the wavelet filtered signal is directly linked to the original impedance data information, the ICA mixing matrix elements, and the wavelet filtering strength. For our work, this model is given considering several assumptions:

- *The fault amplitude is constant all over the faulty samples.* Considering that we evaluate incipient cracks, very low variations due to the crack will be noticed in the complex impedance signal. This very low information can be considered quite constant in the studied signals.
- *The fault is independent to the noise.* The noise and the fault severity will have independent variations. The noise has then the same influence on the healthy signal than on the faulty one.
- *The healthy and the faulty signal has the same mean value.* The presence of the incipient fault do not induce any changes in the mean considered signal compared to the healthy one.
- *The considered filtered signal is gaussian distributed.* If we consider that with the ICA one of the source in the mixture can have a normal distribution, the corresponding filtered signal with the wavelet transform can also be considered as gaussian distributed. The model can then be used to obtain a realistic fault amplitude estimation. In the case of this gaussian assumption is not validated, the estimation accuracy will be reduced.

For validating our proposed estimation model, the considered crack sizes are listed in Tab. 5. The corresponding Fault to Noise Ratio (FNR) are calculated to confirm the incipient condition [45] defined as settled with negative or null FNR. The FNR values are obtained considering the variance of the impedance signal in healthy conditions according to the proposal in [7]. For the considered cracks, more the FNR is negative more incipient the fault is. Based on this evaluation, all the considered crack sizes are considered as incipient in a given noise environment.

The calculated healthy and faulty pdf are compared for all the crack size. For the largest crack C_7 the results are displayed in Fig. 17.

We can denote no obvious changes in the mean of two pdfs. Moreover the signal distribution is close to a gaussian one. Thus the estimation model can be used to evaluate the fault severity for cracks smaller or equal to C_7.

TABLE 5. Considered crack surface and experimental FNR.

Notation	Area (mm^2)	l (mm)	d (mm)	FNR (dB)
C_1	0.01	0.1	0.1	-29.3
C_2	0.02	0.1	0.2	-20.1
C_3	0.04	0.2	0.2	-16.8
C_4	0.08	0.4	0.2	-9.63
C_5	0.12	0.6	0.2	-5.85
C_6	0.16	0.4	0.4	-4.5
C_7	0.24	0.6	0.4	0.21

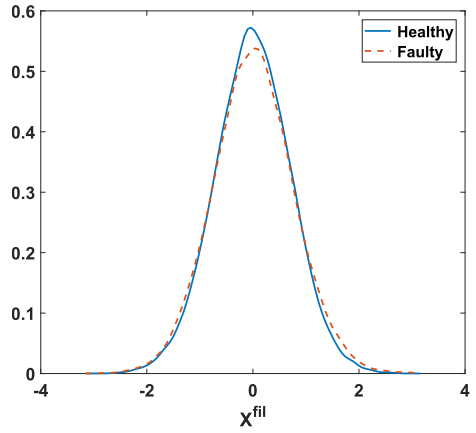


FIGURE 17. Pdfs for crack size C_7.

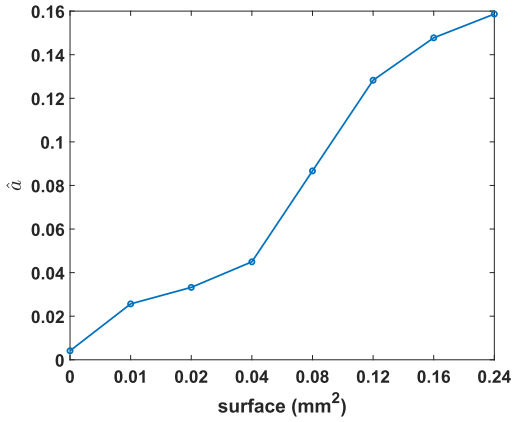


FIGURE 18. Estimated fault severity versus crack surfaces.

In Fig.18 we present the calculated fault severities obtained using our proposed estimation model. It shows that the estimated fault severities increase along with the surface area. This results proves that the derived estimation model can allow to evaluate the incipient fault level successfully.

VII. CONCLUSION

In this paper, we proposed a new fault diagnosis method based on the ICA-wavelet-JSD combination for the minor cracks detection and estimation in high perturbation levels. The benefit of using a blind source separation function coupled with wavelet transform and Jensen-Shannon Divergence is discussed and validated. The limitations of the literature are first highlighted for methodologies based on Kullback-Leibler divergence and Cumulative Sum. The benefit of our proposal is discussed and its detection performances is proved as the most efficient one in high noise levels. For estimation

purpose, we derive a theoretical model for our process and we show its efficiency for several incipient crack severities. Our proposed diagnosis process is then well validated for experimental crack detection.

ACKNOWLEDGMENT

The authors would like to thank China Scholarship Council for funding and also Prof. Y. Le Bihan for providing the experimental ECT data.

REFERENCES

[1] E. L. Chaing, L. H. Russel, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. London, U.K.: Springer, 2001.

[2] R. Hamia, C. Cordier, and C. Dolabdjian, "Eddy-current non-destructive testing system for the determination of crack orientation," *NDT & E Int.*, vol. 61, pp. 24–28, Jan. 2014.

[3] A. Rosell and G. Persson, "Finite element modelling of closed cracks in eddy current testing," *Int. J. Fatigue*, vol. 41, pp. 30–38, Aug. 2012.

[4] *Standard Terminology for Nondestructive Examinations*, Standard ASTM E1316-19b, 2019. [Online]. Available: http://www.astm.org

[5] P. Zhu, Y. Cheng, P. Banerjee, A. Tamburrino, and Y. Deng, "A novel machine learning model for eddy current testing with uncertainty," *NDT & E Int.*, vol. 101, pp. 104–112, Jan. 2019.

[6] M. R. Bato, A. Hor, A. Rautureau, and C. Bes, "Impact of human and environmental factors on the probability of detection during NDT control by eddy currents," *Measurement*, vol. 133, pp. 222–232, Feb. 2019.

[7] J. Harmouche, C. Delpha, D. Diallo, and Y. Le Bihan, "Statistical approach for nondestructive incipient crack detection and characterization using Kullback–Leibler divergence," *IEEE Trans. Rel.*, vol. 65, no. 3, pp. 1360–1368, Sep. 2016.

[8] H. Chen and B. Jiang, "A review of fault detection and diagnosis for the traction system in high-speed trains," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 2, pp. 450–465, Feb. 2020.

[9] H. Chen, B. Jiang, S. X. Ding, N. Lu, and W. Chen, "Probability-relevant incipient fault detection and diagnosis methodology with applications to electric drive systems," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 6, pp. 2766–2773, Nov. 2019.

[10] X. Chen, J. Wang, and J. Zhou, "Probability density estimation and Bayesian causal analysis based fault detection and root identification," *Ind. Eng. Chem. Res.*, vol. 57, no. 43, pp. 14656–14664, Oct. 2018.

[11] J. Zhou, S. Zhang, and J. Wang, "A dual robustness projection to latent structure method and its application," *IEEE Trans. Ind. Electron.*, early access, Feb. 5, 2020, doi: 10.1109/TIE.2020.2970664.

[12] J. Wang, W. Zhang, and J. Zhou, "Fault detection with data imbalance conditions based on the improved bilayer convolutional neural network," *Ind. Eng. Chem. Res.*, vol. 59, no. 13, pp. 5891–5904, Apr. 2020.

[13] W. Y. Fan, G. P. Wang, H. D. Liu, Z. J. Xie, Y. P. Chen, Z. F. Yang, and H. C. Wang, "Fault diagnosis of bearing based on KPCA and KNN method," *Adv. Mater. Res.*, vol. 986, pp. 1491–1496, Jul. 2014.

[14] T. Liu, W. Wang, W. Qiang, and G. Shu, "Mechanical properties and eddy current testing of thermally aged Z3CN20.09M cast duplex stainless steel," *J. Nucl. Mater.*, vol. 501, pp. 1–7, Apr. 2018.

[15] J. Harmouche, C. Delpha, and D. Diallo, "Incipient fault detection and diagnosis based on Kullback–Leibler divergence using principal component analysis: Part I," *Signal Process.*, vol. 94, pp. 278–287, Jan. 2014.

[16] A. Youssef, C. Delpha, and D. Diallo, "Performances theoretical model-based optimization for incipient fault detection with KL divergence," in *Proc. 22nd Eur. Signal Process. Conf. (EUSIPCO)*, 2014, pp. 466–470.

[17] H. Chen, B. Jiang, and N. Lu, "An improved incipient fault detection method based on Kullback–Leibler divergence," *ISA Trans.*, vol. 79, pp. 127–136, Aug. 2018.

[18] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[19] Y. Du and D. Du, "Fault detection and diagnosis using empirical mode decomposition based principal component analysis," *Comput. Chem. Eng.*, vol. 115, pp. 1–21, Jul. 2018.

[20] M. A. B. Shams, H. M. Budman, and T. A. Duever, "Fault detection, identification and diagnosis using CUSUM based PCA," *Chem. Eng. Sci.*, vol. 66, no. 20, pp. 4488–4498, Oct. 2011.

[21] C. Delpha, D. Diallo, H. Al Samrout, and N. Moubayed, "Multiple incipient fault diagnosis in three-phase electrical systems using multivariate statistical signal processing," *Eng. Appl. Artif. Intell.*, vol. 73, pp. 68–79, Aug. 2018.

[22] X. Zhang, C. Delpha, and D. Diallo, "Performance of Jensen Shannon divergence in incipient fault detection and estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 2742–2746.

[23] X. Zhang, C. Delpha, and D. Diallo, "Incipient fault detection and estimation based on Jensen–Shannon divergence in a data-driven approach," *Signal Process.*, vol. 169, Apr. 2020, Art. no. 107410.

[24] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 145–151, Jan. 1991.

[25] W. Yang, H. Song, X. Huang, X. Xu, and M. Liao, "Change detection in high-resolution SAR images based on Jensen–Shannon divergence and hierarchical Markov model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 8, pp. 3318–3327, Aug. 2014.

[26] S. Kolouri, S. R. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, "Optimal mass transport: Signal processing and machine-learning applications," *IEEE Signal Process. Mag.*, vol. 34, no. 4, pp. 43–59, Jul. 2017.

[27] C. Vilani, *Topics in Optimal Transportation* (Graduate Studies in Mathematics), vol. 58. Providence, RI, USA: American Mathematical Society, 2003.

[28] A. Hyvärinen, "Independent component analysis in the presence of Gaussian noise by maximizing joint likelihood," *Neurocomputing*, vol. 22, nos. 1–3, pp. 49–67, Nov. 1998.

[29] A. Hyvarinen, "Gaussian moments for noisy independent component analysis," *IEEE Signal Process. Lett.*, vol. 6, no. 6, pp. 145–147, Jun. 1999.

[30] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis, Adaptive and Learning Systems for Signal Processing, Communications, and Control*, vol. 1. Hoboken, NJ, USA: Wiley, 2001, pp. 11–14.

[31] S. Ikeda and K. Toyama, "Independent component analysis for noisy data—MEG data analysis," *Neural Netw.*, vol. 13, no. 10, pp. 1063–1074, Dec. 2000.

[32] V. Nassiri, M. Aminghafari, and A. Mohammad-Djafari, "Solving noisy ICA using multivariate wavelet denoising with an application to noisy latent variables regression," *Commun. Statist.-Theory Methods*, vol. 43, nos. 10–12, pp. 2297–2310, May 2014.

[33] P. E. Tikkanen, "Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal," *Biol. Cybern.*, vol. 80, no. 4, pp. 259–267, Apr. 1999.

[34] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.

[35] L. Satish and B. Nazneen, "Wavelet-based denoising of partial discharge signals buried in excessive noise and interference," *IEEE Trans. Dielectrics Electr. Insul.*, vol. 10, no. 2, pp. 354–367, Apr. 2003.

[36] U. Kruger and L. Xie, *Advances in Statistical Monitoring of Complex Multivariate Processes*. New York, NY, USA: Wiley, 2012.

[37] J. Harmouche, C. Delpha, and D. Diallo, "A theoretical approach for incipient fault severity assessment using the Kullback–Leibler divergence," in *Proc. 21st Eur. Signal Process. Conf. (EUSIPCO)*, 2013, pp. 1–5.

[38] Y. Le Bihan, J. Pávó, and C. Marchand, "Characterization of small cracks in eddy current testing," *Eur. Phys. J. Appl. Phys.*, vol. 43, no. 2, pp. 231–237, Aug. 2008.

[39] A. Sophian, G. Y. Tian, D. Taylor, and J. Rudlin, "A feature extraction technique based on principal component analysis for pulsed eddy current NDT," *NDT & E Int.*, vol. 36, no. 1, pp. 37–41, Jan. 2003.

[40] Y. He, M. Pan, F. Luo, D. Chen, and X. Hu, "Support vector machine and optimised feature extraction in integrated eddy current instrument," *Measurement*, vol. 46, no. 1, pp. 764–774, Jan. 2013.

[41] X. Chen, D. Hou, L. Zhao, P. Huang, and G. Zhang, "Study on defect classification in multi-layer structures based on Fisher linear discriminate analysis by using pulsed eddy current technique," *NDT & E Int.*, vol. 67, pp. 46–54, Oct. 2014.

[42] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Evanston, IL, USA: Routledge, 2018.

[43] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.

[44] D. M. Hawkins and D. H. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*. New York, NY, USA: Springer, 2012.

[45] C. Delpha and D. Diallo, "Incipient fault detection and diagnosis: A hidden information detection problem," in *Proc. IEEE 24th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2015, pp. 837–842.

[46] A. Youssef, J. Harmouche, C. Delpha, and D. Diallo, "Capability evaluation of incipient fault detection in noisy environment: A theoretical Kullback–Leibler divergence-based approach for diagnosis," in *Proc. 39th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Nov. 2013, pp. 7364–7369.
[47] M. Kawanabe and N. Murata, "Independent component analysis in the presence of Gaussian noise," Univ. Tokyo, New Delhi, India, Tech. Rep. METR 2000-W, 2000.
[48] A. Hyvarinen, "Fast ICA for noisy data using Gaussian moments," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 5, May/Jun. 1999, pp. 57–61.
[49] P. M. Crowley, "An intuitive guide to wavelets for economists," Bank Finland Res. Discuss. Papers, Bank Finland, Helsinki, Finland, Tech. Rep. 1/2005, 2005.
[50] X. Dai, L. Z. Cheng, J.-C. Mareschal, D. Lemire, and C. Liu, "New method for denoising borehole transient electromagnetic data with discrete wavelet transform," *J. Appl. Geophys.*, vol. 168, pp. 41–48, Sep. 2019.
[51] R. Sharbati, F. Khoshnoudian, H. R. Ramazi, and H. R. Amindavar, "Stochastic modeling and simulation of ground motions using complex discrete wavelet transform and Gaussian mixture model," *Soil Dyn. Earthq. Eng.*, vol. 114, pp. 267–280, Nov. 2018.
[52] I. Eisenberger, "Genesis of bimodal distributions," *Technometrics*, vol. 6, no. 4, pp. 357–363, Nov. 1964.
[53] C. D. Brown and H. T. Davis, "Receiver operating characteristics curves and related decision measures: A tutorial," *Chemometric Intell. Lab. Syst.*, vol. 80, no. 1, pp. 24–38, Jan. 2006.
[54] D. L. Streiner and J. Cairney, "What's under the ROC? An introduction to receiver operating characteristics curves," *Can. J. Psychiatry*, vol. 52, no. 2, pp. 121–128, Feb. 2007.

XIAOXIA ZHANG received the bachelor's and master's degrees in mechanical engineering from the Wuhan University of Technology, Wuhan, China, in 2014 and 2017, respectively. She is currently pursuing the Ph.D. degree with the Doctoral School Sciences et Technologies de l'Information et de la Communication, Laboratoire des Signaux et Systèmes, CentraleSupélec, Université Paris Saclay, France, under the supervision of Prof. C. Delpha. Her current research interests include statistical signal processing, data-based incipient fault detection and diagnosis for complex systems, and their applications to mechanical and electrical processes.

CLAUDE DELPHA (Senior Member, IEEE) graduated in electrical and signal processing engineering. He received the Ph.D. degree in instrumentation and measurements and signal processing from the Université de Metz, with an application on smart sensors-based systems. Since 2001, he has been with the Laboratoire des Signaux et Systèmes (UMR CNRS 8506, CentraleSupelec, France). He is currently an Associate Professor with the Université Paris Saclay, France. He also works in the field of signal processing for complex systems security and process monitoring. His main areas of interests are multidimensional and statistical signal processing, fault diagnosis and prognosis (modeling, detection, estimation), PHM, data hiding (watermarking, steganography), pattern recognition (machine learning and deep learning).

DEMBA DIALLO (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical and computer engineering from the National Polytechnic Institute of Grenoble, France, in 1990 and 1993, respectively. He is currently a Full Professor with Université Paris Saclay and the Director of the French National Research Network on Electrical Engineering. He also develops his research activities within the Group of Electrical Engineering Paris. His current area of research includes fault diagnosis, fault tolerant control, and energy management. The applications are related to more electrified transportation systems (EV and HEV) and microgrids with renewable energies. He has published more than 150 journal and international conference papers and three book chapters.