Development of a robust inversion methodology in nondestructive eddy current testing

Shamim Ahmed

université PARIS-SACLAY

NNT : 2018SACLS043

Thèse de doctorat

UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir®

Développement d'une méthodologie robuste d'inversion dédiée au CND par courants de Foucault

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°580 Sciences et Technologies de l'Information et de la Communication (STIC)
Spécialité de doctorat: Traitement du Signal et des Images

Thèse présentée et soutenue à CEA, Saclay, Gif-sur-Yvette, le 5 mars 2018, par

Shamim Ahmed

Composition du Jury:

Dominique Lesselier Directeur de Recherche, CNRS, L2S UMR 8506	Président
Andrea Massa Professeur, University of Trento (DIGITEO Chair)	Directeur de thèse
Bruce Drinkwater Professeur, University of Bristol	Rapporteur
Yiming Deng Professeur Associé, Michigan State University	Rapporteur
Pierre-Emile Lhuillier Ingénieur Chercheur, Electricité de France	Examinateur
Roberto Miorelli Ingénieur Chercheur, CEA-LIST	Encadrant
Pierre Calmon Directeur de Recherche, CEA-LIST	Invité
Marco Salucci Chercheur, ELEDIA, University of Trento	Invité

This work is dedicated to my mother Hosneara Rashid and my father Md Abdur Rashid.

Acknowledgement

This work has been partially supported by the SIRENA project (2014-2017) funded by DIGITEO (France) under the Call for Chairs 2014, the French project ANR-ByPASS, and by the Italian Ministry of Foreign Affairs and International Cooperation, Directorate General for Cultural and Economic Promotion and Innovation within the SNATCH Project (2017-2019).

I would like to thank all of my colleagues from CEA (Saclay, France) and ELEDIA Research Center (Trento, Italy and L2S, France) for their continuous support and collaboration.

Conferences and publications

Peer reviewed proceedings

[1] S. Ahmed, M Salucci, R. Miorelli, N. Anselmi, G. Oliveri, P. Calmon, C. Reboud and A. Massa, "Real time Groove Characterization Combining Partial Least Squares and SVR Strategies: Application to Eddy Current Testing," *Journal of Physics: Conference Series*, vol. 904, Oct. 2017.

Peer reviewed paper in edited book

[2] S. Ahmed, R. Miorelli, C. Reboud, P. Calmon, N. Anselmi and M. Salucci, "Fast Characterization of Multiple Cracks in Conductive Media based on Adaptive Feature Extraction and SVR," *Studies in Applied Electromagnetics and Mechanics: Electromagnetic Nondestructive Evaluation (XXI)* [Accepted].

[3] S. Ahmed, R. Miorelli, M. Salucci, and A. Massa, "Real-time Flaw Characterization Through Learning-by-Examples Techniques: A Comparative Study Applied to ECT," *Studies in Applied Electromagnetics and Mechanics: Electromagnetic Nondestructive Evaluation (XX),* vol. 42, 2017, pp. 228-235, Jun. 2017.

Other conferences and proceedings

[4] S. Ahmed, R. Miorelli, P. Calmon, N. Anselmi and Marco Salucci, "Real Time Flaw Detection and Characterization in Tube through Partial Least Squares and SVR: Application to Eddy Current Testing," The 44th International Annual Review of Progress in QNDE, USA, Jul. 2017. [Under review process].

[5] Salucci, S. Ahmed, N. Anselmi, G. Oliveri, P. Calmon, R. Miorelli, C. Reboud, A. Massa, "Real-Time Crack Characterization in Conductive Tubes Through an Adaptive Partial Least Squares Approach," *IEEE Interna-*

tional Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, pp. 21-22, Jul. 2017

[6] M. Salucci, S. Ahmed, A. Massa, "An adaptive Learning-by-Examples strategy for efficient Eddy Current Testing of Conductive Structures," *The 10th European Conference on Antennas and Propagation* (*EuCAP*), pp. 1-4, Apr. 2016.

[7] S. Ahmed, R. Miorelli, P. Calmon, M. Salucci, and A. Massa, "Learning by Example Techniques for Fast Flaw Characterization During Eddy Current Testing Inspection," *The 50th French Confederation for Non Destructive Testing*, Strasbourg, May. 2017.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The thesis focuses on the development of robust and innovative strategies for the solution of inverse problems arising in the field of Non-Destructive Testing and Evaluation (NDT&NDE). In particular, inverse problems dealing with Eddy current testing (ECT) NDT has been emphasized. In this chapter, an overview on the different NDT methods that are used in industries for inspecting defect(s)/flaw(s) within the inspected medium are discussed. A review on the state of art of the applications of ECT has been discussed with the definition of forward and inverse problems. Among different inverse problem solutions, non-iterative Learning-By-Examples (LBE) strategy has been briefly introduced which is followed by defining the goal of this thesis.

1.1 Nondestructive testing and evaluation

Nondestructive Testing (NDT) deals with the methods which are used to test material characteristics such as defects (e.g., cracks, voids, corrosion etc.) electromagnetic and/or mechanical properties of a material or system without causing any impairment for its future uses. Nondestructive Evaluation (NDE), on the other sense, is used to envision inside the material to locate and characterize of any possible defects. Usually NDE procedures must have an illumination source such as visible light, ultrasonic waves, electromagnetic waves, etc. These methods are used to guarantee quality control standards and to perform integrity test during the whole manufacture lifetime. NDE can be divided in two broad aspects (qualitative and quantitative). Qualitative NDE is applied to investigate a defect inside the material. Quantitative NDE is concerned for locating and sizing the defects or by providing the numerical measures of the defect or material properties.

Several inspection techniques have been developed for performing NDT/NDE analyses, such as [1]:

- Visual Testing (VT): this testing is done either through direct observation or using optical instrument by a principle inspector. This procedure is widely used

for detecting surface anomalies associated with various structural failures. The application of this procedure can be used for detecting surface discontinuities, cracks, corrosions, misalignments, physical damages etc.

- Liquid Penetrate Testing (LPT): this method involves penetrating liquid for testing discontinuities in solid and nonporous materials. Principal use cases of this method includes post-fabrication testing, in-process and quality control. Usually surface discontinuities, seams, cracks, laps, porosity leaks paths are mostly investigated by LPT.

- Ultrasonic Testing (UT): beams of sound waves having higher frequency span from 100 kHz to 50 MHz are injected into the inspected materials for testing discontinuities in the materials and assess material characteristics. These waves travel through the materials, are being attenuated partially and reflected at the interfaces. These echos are analyzed to define the presence and location of the discontinuities. UT is widely used for identifying cracks, holes, laminations, slags, cavities, porosities, flakes, lack of penetration and other discontinuities that produce sharp boundaries.

- Acoustic Emission Testing (AET): mechanical waves produced by sudden movement in stressed materials are known as acoustic emission. The main sources of acoustic emission are discontinuity related deformation process such as crack growth and plastic deformation. Sudden movement at the source produces a stress wave that radiates out into the structures and excites a sensitives piezoelectric sensor. Emissions are generated by the rising stress in the material. Through this process, the highly sensitive equipment can not only detect the crack growth but also processes such as solidification, friction, impacts and phase transformation. Hence acoustic emission testing is also used for in-process weld monitoring, detecting loose parts and loose particles, cavitation, weld monitoring and leak testing.

- Infrared and Thermal Testing (ITT): thermal testing involves the measurement or mapping of surface temperature while heat is passing through the test objects. Surface temperature varies according to the heat transferred characteristics of an object. On basis of the heat pattern, discontinuities of the object can be identified. This testing is used in transmission and distribution lines, transformers, cable tray, friction in rotating equipment, insulation in boilers, furnaces kilns, testing delamination and voids in composite material.

- X-Ray (XR): the concept of differential adsorption of penetrating radiation to inspect material and components is widely used in the domain X-ray NDT. With the material discontinuity (e.g., void or change in configuration), the degree of radiation absorption is changed. Therefore, due to the differences in density, thickness, shapes, sizes, or absorption characteristics of the structure under testing, the amount of absorbing radiation is different. The unabsorbed radiation

passed through the irradiated specimen is recorded on different type of detectors such as film, fluorescent screens, or other radiation monitors. Finally, the condition of the tested medium appears as the variants of black/white/gray contrasts on exposed film, or variants of color on fluorescent screens. X-ray inspection can reveal hidden voids and defects in a variety of solid materials, including metal, plastic, composites, silicone and rubber etc.

- Electromagnetic Testing (ET): electromagnetic signals are used for inspecting the material condition of the inspected object. It is commonly used for inspecting alloy content, cavity, local strain, composition, contamination, cracks, heat treatment, assessment of mechanical and physical properties etc.

The main focus of this thesis is to deal with eddy current testing inspection which is widely applied within the ET domain. The following sections briefly describe some of the ET methods.

1.2 Electromagnetic testing (ET)

In case of electromagnetic testing, electromagnetic signals are induced into the inspected object. The induced energy interacts with the material of the object and the interaction is analyzed to know the condition of the material. On basis of the different excitation frequencies, different nature of the transducers are used. Thus the set of electromagnetic testing methods that can be divided by considering the bands of the frequencies employed for inspecting the Structure Under Test (SUT) [1, 2, 3]:

- Magnetic Flux Leakage Testing (MFLT) [1Hz-10Hz],
- Eddy Current Testing (ECT) [10Hz-10MHz],
- Microwave and Terahertz Testing (MTT) [0.3GHz-10THz].

1.2.1 Magnetic flux leakage testing

This testing is frequently used in industry for testing ferromagnetic parts and components. The test object is magnetized by a permanent magnet or by passing an excitation current directly through an electromagnet. The magnetic field lines are disturbed in case of discontinuity on or near the surface of the test object. As a result local leakage field is introduced surrounding the discontinuities. This leakage field is detected and imaged by various magnetic flux imaging techniques for illustrating the discontinuities within the test object (Figure 1.1).

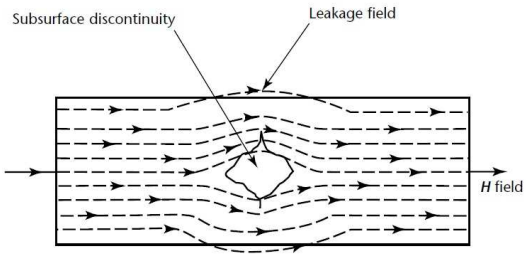


Figure 1.1: Magnetic flux leakage testing phenomenon [1].

1.2.2 Eddy current testing

Eddy current testing (ECT) is the widely used electromagnetic testing in NDT. ECT relies on the principles of magnetic induction to interrogate the materials under test. Eddy current testing is based on the fact that, when a coil excited by an alternating current is brought close to a material, the terminal impedance of the coil changes. The primary field set up by the eddy current coil induces eddy currents within the electrically conducting specimen that brings makes the associated changes [1]. Due to the possibility to design very compact and contact-less probes, which consist of particular coil arrangement and/or magnetic sensors (e.g., Giant magnetic resistance GMR, Hall effect sensor etc.), ECT is well suited for detecting small cracks within the test object. Specially in nuclear power plants for inspecting steam generator tube, in oil and gas applications for testing corrosions under coating layers in pipes etc. In aeronautics, ECT probes are used to inspect micro-cracks appearing close to the fastener sites by inspecting area surrounding the rivet.

1.2.3 Microwave and terahertz testing

Electromagnetic waves having frequencies between 0.3 GHz and 300 GHz are used for this kind of testing in NDT for inspecting weakly conductive or dielectric media. Microwaves can propagate through nonmetallic materials while they are almost completely reflected by metallic materials. The basic theory is to illuminate test set object with microwaves through a suitable antenna. The signal travelling through the test object is received by a receiving antenna. The acquired signal is analyzed and compared with the amplitude and phase of the incident signal. By comparing the signals, the discontinuities of the test object are examined.

Terahertz waves testing nondestructive evaluation is an inspection method that uses the radiation between the frequency range within 0.1 THz and 10 THz. It offers a non-contact and high-resolution means of inspection for corrosion effects that are hidden under a dielectric layer such as non-conductive paint [4]. Like microwaves, this radiation is able to penetrate and inspect non-conductive material. The radiation also reflects from interfaces that exhibit material discontinuities such as inclusions and

voids. Due to the use of shorter wavelength, higher spatial and temporal resolution can be obtained by using terahertz testing compared to conventional microwave testing.

1.3 An overview of eddy current principle

Among all the electromagnetic NDT (E-NDT) methods, ECT is the most suitable for detecting surface breaking cracks. Nevertheless, ECT is also employed for detecting subsurface cracks and the cracks affecting in-homogeneous medium (e.g., welds planner and cylindrical medium) which are common in nuclear, aerospace and aeronautical domains. Eddy currents are induced in a test-piece by a probe coil. When the coil is surrounded by air, it is characterized by the impedance parameter Z_0, which is a complex number defined as in Eq. (1.1). This represents the voltage-current ratio (V_0/I_0) at the time harmonic excitation frequency f [2].

$$Z_0 = \frac{V_0}{I_0} = R_0 + jX_0 = R_0 + j2\pi f L_0 \tag{1.1}$$

Eddy current inspection is based on Faraday's electromagnetic induction law. Faraday discovered that a time-varying magnetic flux density induces currents in an electrical conductor. When an alternating energized coil of impedance Z_0 approaches an electrically conductive material, the primary alternating magnetic field penetrates the material and generates continuous and circular eddy current. The induced current flowing within the test piece generates a secondary magnetic field that tends to oppose the primary magnetic field, as shown in Fig. 1.2. In effect, the new imaginary part of the coil impedance decreases proportionally when the eddy current intensity in the test piece increases. Eddy currents also contribute to the increase of power dissipation that affects the real part of coil impedance.

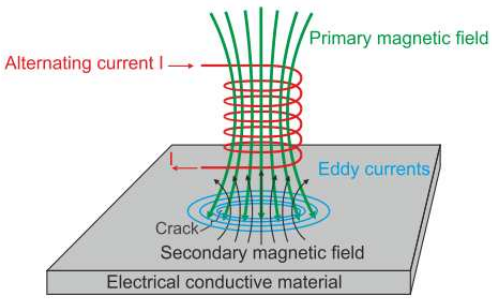


Figure 1.2: Eddy current testing setup [2].

When an energized coil probe approaches the conductive test material, eddy current appears in the test piece. It creates a secondary field that interacts with the primary field. As a result, the new impedance of the coil, Z_{object} can be expressed by Eq. (1.2)

$$Z_{object} = R_{object} + jX_{object} = R_{object} + j2\pi f L_{object} \tag{1.2}$$

R_{object} and jX_{object} represent the real and the imaginary parts of Z_{object}, which is proportional to frequency f and the induction coefficient L_{object} when the coil approaches the test piece. If the test piece is affected by any flaw/defect or material discontinuity, the impedance of the coil changes while inspecting the flawed region and that can be denoted by $Z_{object+defect}$. Therefore, by measuring the coil impedance variation denoted by $\triangle Z$ that is defined in Eq. (1.3) through monitoring either the voltage or the current signals, one can infer the changes in medium thickness, conductivity, permeability, material discontinuities, etc. That is, defect(s) or geometric variation(s) in the test piece cause the variation of the electromagnetic field induced within the medium (i.e., how eddy currents flow). This causes the impedance variation of the probe while it is scanning the work-piece. This variation is the quantity of interest in eddy current testing. It can be described as

$$\triangle Z = Z_{object} - Z_{object+defect} \tag{1.3}$$

where

- Z_{object} is the probe impedance due to test object,
- $Z_{object+defect}$ is the probe impedance due to the presence of the defect in the test object.

Similar to $\triangle Z$, ECT signal can also be represented in terms of voltage variation $\triangle V$ $[V]$ and magnetic flux density variation $\triangle \mathrm{B}$ $[T]$ (when the probe(s) are sensitive to the magnetic flux density variation).

The main advantages of eddy current inspection include:

- relatively fast scanning speed (compare to X-Ray, Ultrasound testing etc.)
- it can be applied for inspections in severe operational conditions (very high temperatures)
- minimal preparation is required for the surfaces (rough surface)
- it can be used through several millimeters of coating
- small cracks can be detected if they lie few skin depth below the surface of the inspected specimen
- complex geometries (physically) can be investigated
- testing devices are portable and immediate feedback can be provided

1.4 Applications of eddy current testing

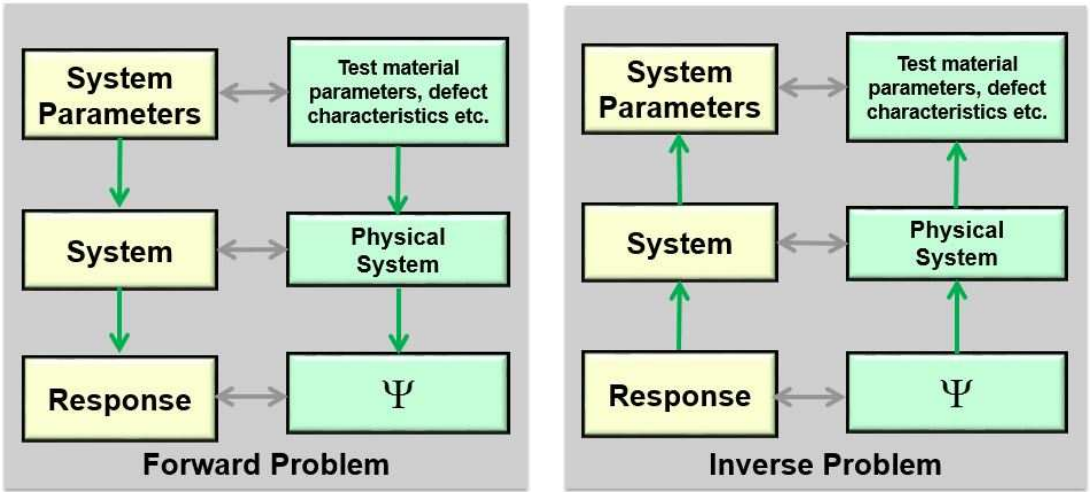
ECT inspection is frequently applied in the aeronautical, nuclear, petrochemical and metallurgic industries for crack detection, localization and characterization. Moreover, ECT is also applied for material characterization and ECT data are also useful for classification tasks. Life span of steam generator tubes (in nuclear power plant) is subjected to various degradation processes that lead to circumferential cracks, notches, pitting, wall thinning, leakage or rupture, stress corrosion etc. [5]. ECT probes are pushed and pulled through the tubes to be inspected at higher speeds than with other NDE methods. Eddy current inspection is carried out regularly on aircraft under carriage wheel, in different tubes, tublar components of aircraft and engine for detecting corrosions, surface and subsurface defects in fastener site. Surface defects and conductivity testing are done by applying high frequency ECT signals whereas subsurface defects are detected through low frequency ECT signals. Having high level of sensitivity, ECT is utilized for material identification and for the characterization of the micro-structure state [2]. The physical parameters of the test pieces are related to the electrical conductivity and magnetic permeability which can be measured by impedance variation of coil. The relation between hardness and these parameters permits eddy current testing to be applied on heat damage detection and heat treatment control [6]. In [7], the advantage of lift-off variation allows the ECT to measure the coating thickness of non-conductive materials or the oxide thickness of conductive materials that was explored. Authors in [8] have shown railroad track surface testing at train speeds of 70 km/h by applying ECT in the field of transportation. In metallurgic domain, defects can appear either randomly or periodically during material production lines. A poor overall quality of the material can be identified by detecting the presence of random defect(s) in the raw material or flaw(s) in the general production process. Periodic defects can be generated at regular intervals by the damaged rollers or guide rollers in the production line. ECT can be used for inspecting residual stresses in engineering structures that can provide early indications of stress status and eventual failure is a rapidly growing area in non-destructive testing [9]. Eddy current coil probes are used in [10] to detect very small stress variations in ferromagnetic steels due to the magneto-elastic effect based on the measurement of the changes in impedance.

1.5 Forward and inverse problems

In NDE, simulation tools are widely applied in order to assess the performance of the inspection procedure for designing inspecting probes and to help experts in interpreting measured signals. We can distinguish two main problems that can be addressed through analytical and/or numerical simulation tools. The first one is known as forward problem and the latter one is known as inverse problem. In the following sections we provide an overview of both of them in the context of ECT simulation.

1.5.1 Forward problem

The development of the physical models by having system parameters to accurately predict the behavior of the system in terms of the NDE data is known as forward problem. For example, provided a set of parameters describing a given scenario (e.g., definition of test material, defect characteristic, coil etc.) dealing with the calculation of the quantity of interest (e.g., ECT signals/data in terms of impedance variation ($\triangle Z$) of the probe) is known as a froward problem. For the sake of simplicity, we refer Ψ to represent the impedance variation of the ECT probe as ECT signal. Ψ can also be represented for other NDT signals/data (e.g., magnetic flux density, ultrasound field pressure etc.).



(a) (b)

Figure 1.3: Block diagram of (a) forward problem and (b) inverse problem.

If the response of a system always gives the same output for a same set of input parameters, then it is called deterministic forward problem and it can be defined that system response is the function of the set of given inputs. Fig. 1.3(a) shows the definition of forward problem. In ECT, the system accounts for all interactions that happened between parameters (i.e., the probe, specimen, presence of the flaws etc.). The forward problem can be solved in two different ways:

1) By performing the measurement on the configuration of interest.

2) By using an appropriate mathematical/physical model and carrying out numerical simulations.

Various categories of numerical forward solver such as Finite Element Methods (FEM), Method of Moments (MoM) scheme and discretization techniques for integral equations, Finite Difference (FD), Finite Difference Time Domain (FDTD) methods, Finite Integration Techniques (FIT) etc. [11] are used for numerical simulation. Nowadays a large set of commercial softwares are used as physical simulators for

providing numerical results. Among the most used, one can mention COMSOL Multiphysics [12] and FLUX 3D (FEM softwares) [13], CST (FIT software)[14], CIVA (integral equation based solution) [15].

1.5.2 Inverse problem

The task of mapping the NDE data (e.g., ECT signals/data) to determine the system parameters which generate the signals is known as inversion. More specifically, knowing the impedance variation $\triangle Z$ of probe (i.e., Ψ), the retrieval of the state of the material in terms of degradation such as defects/cracks and or material properties is an inverse problem. That means, inverse problem predicts the system parameters from the known system response (Figure 1.3 (b)). For a particular response if the predicted system parameter is always giving same values (unique) then this problem is called well-posed. Unfortunately most of inverse problems deal with practical applications that are not well-posed. In the next section we describe more in details the relationship between forward and inverse mapping.

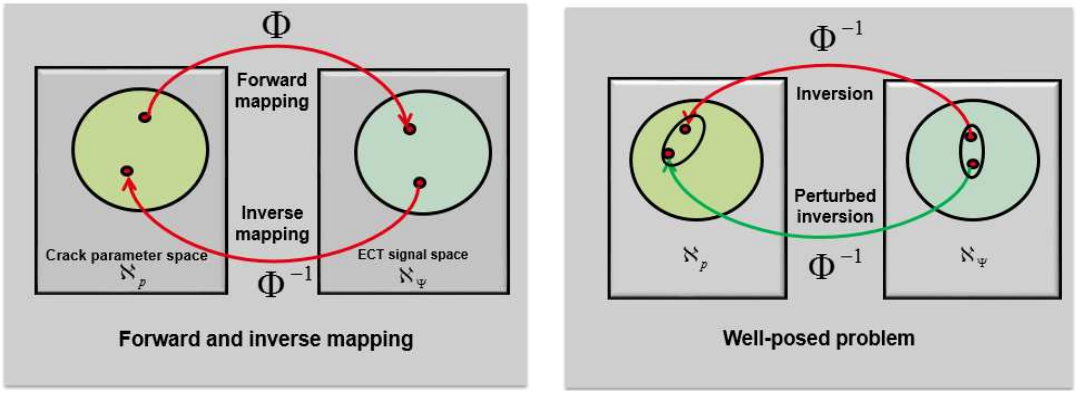
1.5.3 Forward and inverse mapping

Forward and inverse problems can be represented in terms of mapping between certain sets of functions. For simplicity, lets the forward problem consists of finding an appropriate function Φ, for mapping an input parameter p (e.g., crack/defect parameter) forming a system parameter set $\aleph_p$ to the output response Ψ (e.g., in ECT signal, $\triangle Z$) forming a response set $\aleph_\Psi$. $\aleph_p$ and $\aleph_\Psi$ describe the defect/crack parameter space and ECT signal space, respectively. Φ being a mapping operator (known also forward operator), acts on the crack parameter $p \in \aleph_p$ to produce ECT signal $\Psi \in \aleph_\Psi$, which can be defined as $\Phi : \aleph_p \rightarrow \aleph_\Psi$. Given a parameter $p \in \aleph_p$, the ECT signals associated to $\Psi \in \aleph_{\mathbf{\Psi}}$ can be formally described as follow.

$$\aleph_\Psi = \{\Psi \ : \ \Phi(p) \rightarrow \Psi\}. \tag{1.4}$$

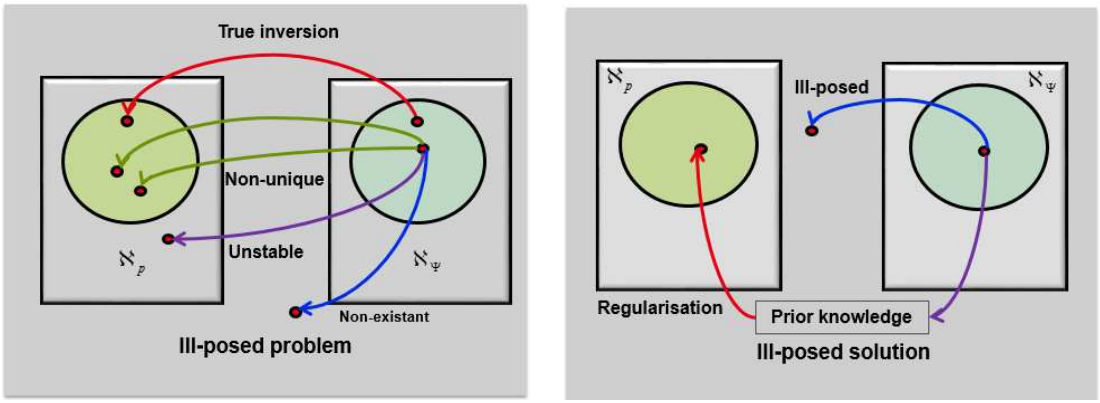
Equation 1.4 states that function operator Φ maps p to Ψ. Hence, the solution of the inverse problem can be defined as finding inverse mapping or inverse operator $\Phi^{-1} : \aleph_\Psi \rightarrow \aleph_p$ for constructing system parameter set $\aleph_p$ from response set $\aleph_\Psi$. Here Φ^{-1} simply reconstructs the system parameter set from known response set which can be represented as

$$\aleph_p = \{p \ : \ \Phi^{-1}(\Psi) \rightarrow p\}.$$



(a) (b)

Figure 1.4: In (a) forward and inverse mapping and in (b) well-posed (stable) inverse problem



(a) (b)

Figure 1.5: (a) Ill-posed problem and (b) ill-posed solution.

Figure 1.4 (a) shows both the forward and inverse mapping operations in pictorial view. Sometimes the available information (responses) is not enough to characterize the system parameters. Inverse mapping should map elements from response set to those elements which are proper members of the system parameter set. Inversion should also be unique, in the sense that having one response should correspond to one element in the system parameter set. Inverse problem which has unique solutions and stability is known as well-posed problem [16] (Fig. 1.4 (b)). Here stability of the mapping defines that a small perturbation of the response grants the inversion a small perturbation on reconstructed parameter.

Sometimes the inversion of the response cannot give unique point in the system parameter set $\aleph_p$. Such a phenomenon may be observed when there is no solution,

or the solution is not unique or it may be unstable. This phenomenon is called ill-posedness problem of the inverse problem. The inversion maps response to the points which lie outside the system parameter set represents a non-existing solution. An ill-posed problem can be present in spite of having an exact knowledge of response, for instance, when noise corrupts the values in the response set. However, the lack of information which is one of the main causes of ill-posedness, can be reduced by providing a priory knowledge (regularization in mathematical terms) by inferring priory information from the system parameters. Figure 1.5 (a) illustrates the graphical representation of the ill-posed problem and a probable solution by providing a-priori information is shown in Fig. 1.5 (b).

1.5.4 State of art of inverse problem solution

In particular, inverse problem can be handled in iterative and non-iterative procedures.

1.5.4.1 Iterative inverse problem solutions

Iterative inversion algorithms, consisting in minimizing a cost function related to the discrepancy between simulation and observation which can be employed for defect characterization and localization. Iterative based inverse solutions can be categorized by the most widely employed ones:

- Deterministic based inversion: gradient-based and gradient-free schemes [17].
- Stochastic based inversion: particle swarm optimization (PSO), evolutionary algorithms (EA)[18, 19].
- Signal processing based inversion: wavelet-transform approach [20].
- Statistical based inversion: Bayesian inference approach [21].

Iterative inverse solutions are useful for any kinds of crack shape reconstruction in case of well posed problem. Some stochastic algorithms such as PSO are robust to deal with noise. Moreover, they can also be applied for complex problems. Very often, iterative solutions suffer from ill-posedness and non-linear problems during the minimization of the cost function. In addition, the iterative minimization of the cost function make these algorithms computationally expensive and some of them can be trapped in area realizing local minima of the cost function under minimization [17, 18, 19, 22] for the evaluation of several forward problems.

1.5.4.2 Non-iterative inverse problem solutions

In general, nowadays one of the biggest challenges in NDT consist in obtaining faster and reliable inverse solutions, ideally during the inspection phase. These are also considered mandatory steps in view of automatic inspection procedures. Non-iterative

inversion strategies such as Multiple Signal Classification (MUSIC) [23], monotonicity [24] and Learning-by-Examples (LBE) [25, 26, 27, 28, 29, 30, 31] have shown to be computationally effective.

Though MUSIC can be applied for real time inverse solution, it is suitable for small cracks characterization. Therefore, it has limitations when dealing with larger cracks and multi-cracks problems. Conversely, monotonicity is suitable for the shape reconstruction of any kinds of cracks under certain hypothesis. Due to the limitations on certain probe assessment, monotonicity cannot be applied directly to any kind of NDE problems.

Heading towards a real-time diagnosis solution in terms of inversion, LBE approach can be a suitable choice to be applied when a parameterized description of the complex problems (i.e. set of input output pairs) is possible. Special probe assessment is not required, while relatively much faster prediction can be obtained. Larger and multi-crack characterization problems can be handled by LBE approach. Moreover, LBE can also be applied on different NDE methods such as ECT, UT etc. by suitably treating signals/data. In this way, LBE can be a suitable parametric inversion strategy for the NDT test cases that can be described through a discrete set of parameters. According to our experiences we could find this as the only limitation of this approach.

1.6 Thesis goal and outlines

The aim of this thesis is to develop a robust and innovative inversion strategy applied to eddy current testing signals based on learning by examples (LBE) paradigm. LBE has been employed to solve the inverse problem for estimating the cracks/flaws, probe(s) etc. parameters in complex structure in almost real-time.

Though the primary focus of thesis is to deal with nondestructive ECT testing, the methodologies developed within the scope of LBE has preliminarily applied to other NDT methods (e.g., Ultrasound Testing (UT)). Moreover, the robustness of the LBE strategy has also been treated on multi-physics applications combining ECT and UT signals/data, which enhances the defect detection and localization capabilities by taking the advantages from both of the NDT techniques. The problems addressed in this thesis are close to practical application cases faced by industries. The proposed LBE approach has been tested against real experiments with promising results. Moreover, the quasi real time prediction capability enabled by LBE approach opens a way to possible application in view of online defect localization/detection and characterization.

The thesis is organized as follows. Firstly, LBE strategy is introduced in Chapter 2 for solving inverse problem. In Chapter 3, different sampling strategies within the framework of LBE are described for training set generation. Chapter 4 shows inverse model generation by using learning algorithms. In Chapter 5 and Chapter 6, applications of LBE on different test cases using ECT signals are presented. Chapter 7 is dedicated for the application of LBE in the context multi-physics approach. Finally, Chapter 8 shows some concluding remarks with the perspectives of the LBE strategy in NDT.

Chapter 2

Inverse problem solution through learning by examples

The main idea behind Learning by examples (LBE) paradigm represents a methodology development in the framework of the statistical learning theory. LBE approach consists in solving the problem of learning by building a fast and accurate model through input-output mappings from empirical data (the training data sets). This type of learning is also called supervised learning. Such a model is then used to predict the output associated to a set of unknown test samples. Within the framework of LBE, different sampling and features extraction strategies can be applied for generating training and testing sets by extracting suitable features from the actual data set. Different (machine) learning algorithms can be fitted on the generated training set for developing a suitable prediction model. Depending on the characteristics of the output, the problem is defined as either regression or classification [32, 33, 34]. If the outputs are continuous with respect to the input data, it is called regression. Whereas, if the output represents discrete values (e.g., a set of categories) with respect to input data, it is called classification. Classification deals with the discrete output (e.g., presence of defect inside the structure under test (SUT), or cracks type etc.) associated to the input data (e.g., NDT signals/data). The fitted model is used to classify a set of unseen test samples into a finite set of categories. Where regression predicts the output associated to the unseen test sample based on the learned model by mapping the input to the corresponding continuous output (e.g., defect characteristics, position/location etc.). Next sections describe the brief overview of machine learning applications on NDT community.

2.1 Machine learning in E-NDT community

Machine learning algorithms started to appear in E-NDT almost three decades ago. Authors in [35] showed eddy current signals interpretation dealing with classification problem. Longitudinal and transverse surface-breaking notches detection and characterization are shown in [36] by using time-domain parameters as the functions of

digitized in-phase and quadrature components of probe impedance. Reconstruction of stress map in the strained metallic plates by using electromagnetic measures have been exploited in [37] through support vector regression (SVR). An eddy current imager signal processing method is presented in [38] based on a principal component analysis (PCA) followed by a maximum likelihood (ML) approach for characterizing the defects in aeronautical riveted lap-joints. Combination of PCA feature extraction and k-mean algorithms is used [39] for cracks detection and classification by applying pulse eddy current testing. Defect localization, and classification tasks are explored in [40] by applying Fisher linear discriminant analysis (FLDA) feature extraction with support vector machine (SVM) by applying pulse eddy current testing. In [41], Tikhonov regularization has been applied as an inversion algorithm for visualization of the geometric profile and size of the cracks through the determination of the 2D image of the eddy current density.

In this thesis, innovative strategies have been addressed within the framework of LBE. The systematic study of different strategies and solutions are adopted within LBE in order to achieve simultaneously good accuracy and computational time efficiency for estimation of defect parameters.

2.2 A gentle introduction to Learning by examples

In general, learning by examples is a two phases approach. During the preliminary phase (so called offline phase), a fast and accurate model is built based on a training set made of input-output (I/O) pairs by regression/prediction techniques. The model is also known as inverse/trained model. The developed (trained) model from offline phase is then used to predict the output associated to an unknown test sample during the second phase (online phase). Both of these phases fulfill the following three steps.

- Feature selection and feature extraction,
- Exhaustive representation of extracted feature space,
- Choice of prediction technique.

For the sake of simplicity and for introducing the mathematical notation needed in this thesis, we consider hereafter an easy ECT problem (toy problem) aiming to better explain the LBE approach. From Chapter 5 to Chapter 7 more realistic and thus complicated problems will be addressed.

2.2.1 Problem definition

Let us consider a $2D$ axisymmetric configuration made by conductive tube having conductivity σ, relative permeability $\mu_r = 1$ (Fig. 2.1). The tube is inspected by two coils excited by a time-harmonic current. The coils centred at $x = 0$ and $z = 0$ are working in differential mode and moving along the y axis of the tube. The tube

is affected by a single axisymmetric groove (i.e., defect) that occupies a volumetric region Ω (Fig. 2.1) within the SUT (i.e., $\sigma(\mathbf{r}) \neq \sigma \; \forall \mathbf{r} \in \Omega$).

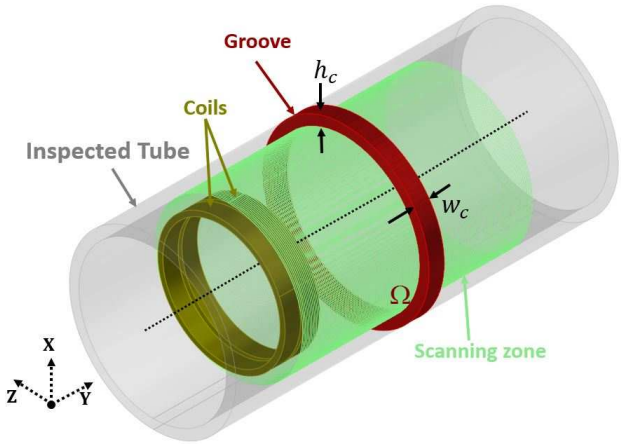


Figure 2.1: Example of studied tube geometry.

2.2.2 Feature selection and feature extraction

2.2.2.1 Feature selection

Features selection is a process by which the ECT signals are collected in terms of impedance variation of the inspecting coils from the SUT. These features represent the set of meaningful measurements in the ECT signal space. Since all along this thesis, we have employed CIVA simulator as a forward operator $\Phi\{.\}$ (see Sec 1.5.3) to generate synthetic ECT signals, hereafter we briefly introduce the Volume Integral Method (VIM) developed in CIVA. The ECT signal in terms of impedance variation at the k-th ($k = 1, ..., K$) scanning position is given by

$$\Psi_k = \frac{1}{I^2} \int_{\Omega} \mathbf{E}^{inc}(\mathbf{r}|\mathbf{r}_k).\rho(\mathbf{r}|\mathbf{r}_k)d\mathbf{r}. \tag{2.1}$$

I is the current flowing inside the coils while $\mathbf{E}^{inc}(\mathbf{r}|\mathbf{r}_k)$ is the incident field generated at position $\mathbf{r}$ in the unflawed tube ($\mathbf{r}_k = y_k$ represents the k-th coil position along the tube). $\rho(\mathbf{r}|\mathbf{r}_k)$ represents the unknown induced current dipole density, which represents the presence of the groove. It is related to the total field, $\mathbf{E}^{tot}(\mathbf{r}|\mathbf{r}_k)$ that can be expresses by

$$\rho(\mathbf{r}|\mathbf{r}_k) = [\sigma(\mathbf{r}) - \sigma]\mathbf{E}^{tot}(\mathbf{r}|\mathbf{r}_k). \tag{2.2}$$

More details of the VIM approach can be found in [42, 43]. We define $\underline{\Psi}$ as the ECT signal vector associated to the selected features as $\underline{\Psi} = (\Psi_k;\, k = 1, ..., K)$ for the set of defect descriptors expressed via the vector $\underline{p} = (p_q; q = 1, ..., Q)$. More concisely, $\Psi_k = \Phi\{\mathbf{r}_k, \underline{p}\}$ satisfies the relationship of Eq. (2.1). The groove considered in the presented problem has a region Ω that can be defined by a finite set of

$Q = 2$ parameters $\underline{p} = (h_c,\, w_c)$. h_c and w_c are the height and width of the groove, respectively. The groove is placed along the tube axis with an angular extension of $360°$(Fig. 2.1). Due to the complex nature of the ECT signals, the associated signal vector can be written as $\underline{\Psi} = \{(\Re\{\Psi_k\}; \Im\{\Psi_k\}); k = 1, ..., K\}$ with $\underline{\Psi} \in \mathbb{R}^{1\times F}$. Therefore, $F = 2K$ features represent the ECT signal space $\aleph_\Psi$. Hence, the ECT signal space has a dimension $dim(\aleph_\Psi) = 2K$. Figure 2.2 shows the absolute value map of ECT signals that are collected along the tube axis, for a single groove configuration according to the test case shown in Fig. 2.1.

Dealing with larger set of ECT signals (i.e., for instance, the impedance variation collected from higher number of measurement points K, increases the number of treated ECT features F). Higher the number of features more complicated the training phase is, thus higher number of training samples are needed for optimal model development during training phase. This problem is also known as curse of dimensionality for learning algorithms. On the other hand, higher number of training samples increases the training time as well as the complexity of the learning task in order to built optimal training model. To facilitate this problem, feature extraction by means of dimensionality reduction techniques are widely used.

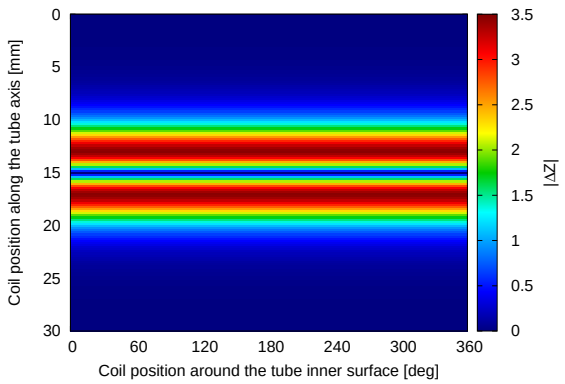


Figure 2.2: The ECT signal map representation in terms of absolute value of coil impedance variation.

2.2.2.2 Feature extraction

The aim of the application of the feature extraction technique is to reduce the $F = 2K$ dimensional ECT feature space to (a much lower dimension) extracted feature space $\aleph_T$. That is, $\aleph_T$ represents the reduced feature space that contains extracted features $\underline{T}$ of dimension $dim(\aleph_T) = J$ where $J \ll F$. Figure 2.3 shows the pictorial view of the feature extraction process. Whereas, the ECT signals are generated based on the parameters space and high dimensional ECT signals are projected to a lower dimensional space $\aleph_T$ by a feature extraction model $\xi\{.\}$. $\xi\{.\}$ defines the type model such as linear, nonlinear, kernel based that can be applied for dimensionality reduction. For example, in Fig. 2.4, we have shown the representation of a learning task when

Partial Least Square (PLS) feature extraction technique is applied for reducing the dimension of ECT signals. More details of PLS feature extraction are available in Sec. 3.3.2.

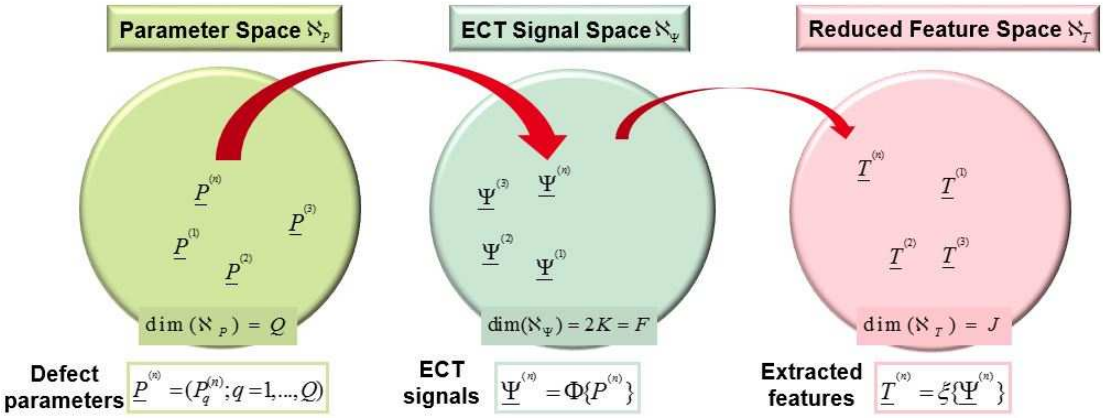
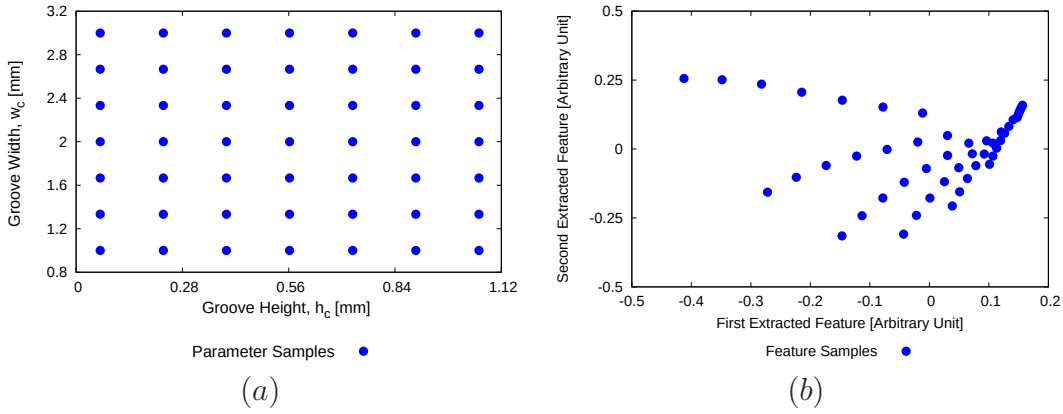


Figure 2.3: Mapping between parameter space to ECT signal space and reduced Feature space.



(a) (b)

Figure 2.4: Samples locations mapped on $a)$ parameter space $\aleph_p$ and $b)$ the resultant extracted feature space $\aleph_T$.

In particular, Fig. 2.4 (a) represents the 2D pictorial view of the parameter space $\aleph_P$ containing groove parameters formed by GRID sampling strategy. The resultant extracted features space $\aleph_T$ with the extracted features $\underline{T}$ from the actual ECT signal space $\aleph_\Psi$ are shown in Fig. 2.4 (b) . For graphical illustration, the extracted feature space is shown for the first two (i.e., $J = 2$) extracted features. From Fig. 2.4 (b), we can observe that the the extracted features in $\aleph_T$, exhibit a pattern with over sampled zones and some under sampled zone. This phenomenon depends on the actual ECT signal space and the distribution of parameter space. The over sampled zone contains redundant information for the learning algorithms which also increases the number of training samples to be treated. Hence, by avoiding the over sampled zone in $\aleph_T$, and uniformly sampling extracted feature space, we can increase the learning ability of

the prediction techniques and obtain better inversion result. In the next section, a brief description of this strategy is illustrated. Other types of feature extraction techniques such as principle component analysis (PCA), kernel PCA, Kernel PLS etc. can also be applied for reducing higher dimensional ECT signal space [33, 34]. However, the obtained pattern of the extracted features shown in Fig. 2.4 (b) can vary depending on the feature extraction method applied for reducing the feature space [44].

2.2.3 Exhaustive representation of extracted feature space

ECT signal varies with the variation of different defect parameters which is not uniform for all parameters. For example, variation of groove height has stronger influence for changing ECT signals compare to variation the width of the groove. Therefore, if a certain budget of simulation is provided, it is essential to select optimal set of training samples of defect (i.e., groove) parameters and the associate ECT signals for building accurate model. By looking at Fig. 2.4, one can notice that an even distribution of samples in the parameter space contributes for the oddly distribution of the extracted feature space on which learning task is performed. To have a suitable and more accurate trained model, uniform distribution of resultant extracted feature space is required in order to represent $\aleph_T$ by the most significant information. Moreover, by uniformly distributing the extracted feature space, redundant and unnecessary samples (i.e., too close features points) for different configuration of defect parameters can be avoided. As a consequence, a more accurate model can be obtained for a given number of training samples. In this thesis, we have applied features extraction technique (e.g. PLS) with modified version of output space filling (OSF) [45] for evenly filling the extracted feature space (Fig. 2.5 (b)) leading to a parameter space which is oddly filled (Fig. 2.5 (a)). The description of PLS-OSF sampling strategy is detailed in Chapter 3.

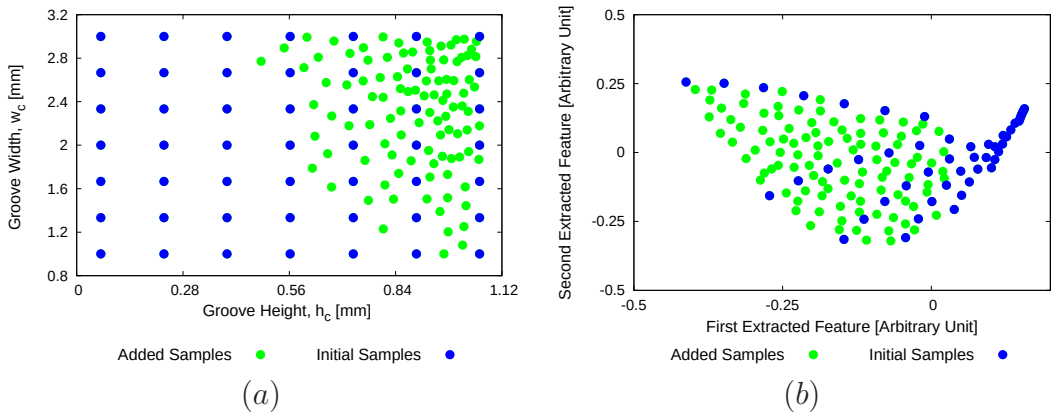


Figure 2.5: Training samples locations mapped on a) parameter and b) the resultant extracted feature space.

2.2.4 Choice of prediction technique

The choice of prediction technique (so called learning algorithm) is one of the crucial steps in LBE. Dealing with a parametric problem, the training model is developed by mapping the NDT signals with the corresponding crack/defect parameters. Within the framework of LBE, machine learning algorithms such as Support Vector Regressor (SVR) [46], Kernel Ridge Regressor (KRR) [47], Relevance Vector Machine (RVM) [48] and Augmented Radial basis function (A-RBF) [49] interpolator have been considered.

In general RBF is an interpolator and A-RBF is obtained by adding a polynomial term to the standard RBF. The polynomial term of A-RBF makes it easier to fit a learning model by using the training data. The more samples are used in training model, the more accurate is the model. However, this leads to over-fitting problem and the learning model becomes more complex in presence of outliers or noise. Therefore, prediction accuracy of A-RBF is higher while treating with noiseless data and conversely decreases while the unknown test set is corrupted.

Having regularization parameters, KRR, RVM and SVR are suitable to deal with noisy data that were addressed in [25] within the framework of LBE. Given a data set $D_{N_q} = \left\{ \left(\underline{\Psi}^{(n)}; p_q^{(n)} \right) ;\, n = 1, ..., N;\, q = 1, ..., Q \right\}$, the aim of the regression task is to find a function $\Theta_q\{.\}$ that estimates the relationship between $\underline{\Psi}^{(n)}$ and $p_q^{(n)}$ such that $\Theta_q(\underline{\Psi}^{(n)})$ provides a predicted/estimated output which is close to $p_q^{(n)}$ for all n. This task is usually done by employing a loss function $\mathcal{L}$ that measures the deviation between the actual $p_q^{(n)}$ and the predicted value $\Theta_q(\underline{\Psi}^{(n)})$ and minimizes the so called Empirical Risk. For example, KRR partially adopts the so called least square error function $\mathcal{L} = \sum_{n=1}^{N} \left(p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)}) \right)^2$. This squared loss function is also known as ℓ_2 loss function, which is sensitive to the presence of outliers or noise. For example, KRR concerns the minimization of this error function $\mathcal{L}$ by solving some of KRR's model parameters. Therefore, any $\underline{\Psi}^{(n)}$ for the corresponding parameter $p_q^{(n)}$ differs a lot from the rest of ECT signals within $\underline{\underline{\Psi}}$ (where $\underline{\underline{\Psi}} = (\underline{\Psi}^{(n)};\, n = 1, ..., N)$) will have disproportionately larger effect on the resulting KRR parameters. Thus, the prediction ability of KRR degrades significantly in presence of noise [34] due to over-fitting.

As it uses a ℓ_1 loss function, this effect is not dominant for applying SVR. Though SVR has a great success in the supervised learning scheme, it has the following disadvantages.

- The number of support vectors grows linearly with the size of the training set. Some form of post processing is also required for reducing computational complexity.
- SVR outputs a point estimation for regression and a hard binary decision for classification. Probabilistic approach is not used.

RVM was introduced by using Bayesian treatment which does not suffer from all of these limitations. It uses a prior over the model weights governed by a set of

hyper-parameters, each one is associated with a particular weight. The most probable values are iteratively estimated from the data. The posterior distributions of many of the weights are peaked around zero which confirm the sparsity. The training vector associated with those remaining non-zero weights are termed as relevance vectors. The important feature of RVM is that it typically utilizes fewer kernel functions.

For evaluating various tuning parameters, RVM is the most time consuming estimator for training a learning model. By properly treating the training set, SVR and RVM both are showing almost similar prediction accuracy. SVR has been successfully applied in Electromagnetic [50, 51] for long time for solving inverse problems and its robustness on noisy data is well known. Moreover, in [25], authors have shown that SVR outperforms than RVM for treating real experimental data in NDE problems. Therefore, we consider SVR as the most suitable prediction technique and we have used SVR through out this research work. Moreover, it is to be noted that the prediction accuracy and robustness of different learning algorithms in presence of noisy data depend on the generation of suitable training set, different defect parameter (i.e., relationship between the ECT signals for particular defect parameter) and the generalization capabilities of the treated algorithm. For higher dimensional problems, it is quite difficult to use other popular machine learning algorithms such Neural Network (NN) as a prediction technique. However, by suitably processing the ECT signals through applying feature extraction and/or applying adaptive sampling, NN can also be used as a prediction technique.

Chapter 3

Different sampling strategies within LBE

This chapter is devoted to show different sampling strategies that can be applied within the scope of learning by examples (LBE) paradigm. Standard sampling approaches such as full factorial GRID, Latin Hyper-cube Sampling (LHS), Output Space Filling (OSF) sampling will be briefly overviewed. In order to build optimal training set, the extension of these sampling approaches with feature extraction method will be explored with pros and cons.

3.1 Problem definition

Let us consider the same problem (toy problem) described in Chapter 2 of a $2D$ axisymmetric configuration made by conductive non-ferromagnetic tube (Fig. 2.1). The coils centred at $x = 0$ and $z = 0$ are working in differential mode and moving along the tube y axis. The tube contains a single axisymmetric groove (i.e., defect placing along the tube axis with an angular extension of $360°$) that occupies a volumetric region Ω (Fig. 2.1) which can be defined by a finite set of $Q = 2$ parameters $\underline{p} = (h_c, w_c)$. h_c and w_c are the height and width of the groove respectively. The ECT signal collected by the coils at the k-th ($k = 1, ..., K$) scanning position is represented by $\underline{\Psi} = \{(\Re\{\Psi_k\}; \Im\{\Psi_k\}); k = 1, ..., K\}$, where $\underline{\Psi} \in \mathbb{R}^{1\times F}$. Hence, the ECT signal space $\aleph_\Psi$ having $dim(\aleph_\Psi) = 2K$, contains $F = 2K$ features. The groove region Ω can be defined by a finite set of $Q = 2$ parameters $\underline{p} = (h_c, w_c)$.

3.2 Training set generation through standard sampling strategies

Prediction ability of the learning algorithms depends on the accurate training (inverse) model generation. The accuracy of the inverse model is dependent on the suitable distribution of defect(s) parameters in the parameter space $\aleph_p$ (i.e., the samples correspondind to the the defect(s) configurations), the associated ECT signals/data on the signal space $\aleph_\Psi$ and extracted features on the feature space $\aleph_T$. Therefore, sampling strategy plays a major role for suitable training set generation and the overall performance of LBE. In this section, different sampling strategies are described in order to generate numerical training and test sets.

3.2.1 Full factorial sampling (GRID)

Full factorial sampling is also known as GRID sampling that contains all possible combinations of a set of factors (e.g., defect parameters). Whereas, each axis in a GRID spans one factor and each point in the GRID represents a unique combination of the factors. In our case, one sample is defined by each combination of the quantization levels of all the defect parameters. The total sample size is the product of the numbers of levels of the all defect parameters. A GRID sampling design of Q parameters, where each parameter is quantized into ℓ levels, creates total $N = \ell^Q$ number of distinct samples. Therefore, N number of samples have been generated by using full-factorial sampling (GRID sampling) strategy within the crack parameter space $\aleph_p$. This is represented by a matrix of defect parameters, $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N)$ having $(N \times Q)$ dimension, where $\underline{p}^{(n)}$ is the n-th row of $\underline{\underline{p}}$. By using $\Phi\{.\}$ generate ECT coil signals and fill the $(N \times F)$ feature matrix $\underline{\underline{\Psi}}$ whose n-th row is represented by $\underline{\Psi}^{(n)}$. Due to the complex nature of the ECT signals, the n-th ECT signal vector is represented by $\underline{\Psi}^{(n)} = \left\{ \left(\Re \left\{ \Psi_k^{(n)} \right\}; \Im \left\{ \Psi_k^{(n)} \right\} \right); k = 1, ..., K \right\}$ of $F = 2K$ ECT features. Finally, the q-th training set $(q = 1, ..., Q)$ is build for each defect parameter $D_{N_q} = \left\{ \left(\underline{\Psi}^{(n)}; p_q^{(n)} \right); n = 1, ..., N \right\}$.

GRID sampling is the most straightforward sampling approach that contains atleast the minimum and maximum values of a parameters. GRID sampling is generally used for training set generation which contains the boundary (e.g., minimum and maximum) values of each defect parameter. Though this sampling approach is easy to implement, it is the most costly sampling strategy where number of total samples increases exponentially with the increase of quantization level of the parameters. In [25], GRID sampling has been shown for training set generation in order to develop training model for crack characterization problem. Due to the high number of samples for increasing number of parameters, GRID sampling has limitation on high dimensional parametric inversion. On the other hand, dealing with higher number of ECT features requires us to deal with high number of training samples for the learning

algorithms to develop accurate training models. This is also known as the curse of dimensionality problem. This also increases the complexity as well as simulation time during training model creation. In Fig. 3.1 (a), a full factorial design of the groove problem having height and width of seven quantization levels (for each parameter) having total $N = 7^2 = 49$ distinct samples are shown in the parameter space. In the literature, other types of deterministic design of experiment sampling strategies are also available. The most well known is the central composite design [52].

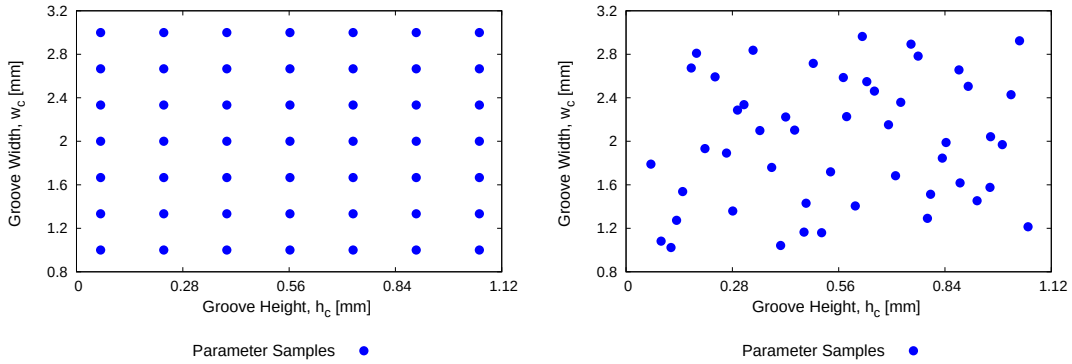


Figure 3.1: Sample locations mapped on the parameter space $\aleph_p$ by (a) GRID sampling and (b) LHS sampling.

3.2.2 Latin hypercube sampling (LHS)

By performing Latin Hypercube Sampling (LHS), the parameter space is randomly sampled in order to fulfill the suitable properties of the sample distribution through generation of a pseudo-random sequence [53]. In this method, the domain of each parameter is randomly decomposed into intervals and same probability is assigned to all the intervals. The number of intervals depends on how many samples would be generated for each parameter. For sampling N different values from each of Q defect parameters, the range of each parameter is divided into N non-overlapping intervals on the basis of equal probability. One value from each interval is selected at random with respect to the probability density in the interval. The N values of each q-th defect parameter is thus obtained through a random manner (equally likely combinations) by pairing with the N values of the other parameter from Q parameters set. It is convenient to think that these samples form an $(N \times Q)$ matrix of input where the $n-$th row contains specific values of each of the Q input parameters. Similar to GRID sampling, N number of samples can be generated by using LHS strategy within the crack parameter space $\aleph_p$. This is represented by a matrix of defect parameters, $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N_0)$ of $(N \times Q)$ dimension, where $\underline{p}^{(n)}$ is the n-th row of $\underline{\underline{p}}$. By using $\Phi\{.\}$, generate ECT coil signals and fill the $(N \times F)$ feature matrix $\underline{\underline{\Psi}}$ whose n-th row is represented by $\underline{\Psi}^{(n)}$. Due to the complex nature of the ECT signals, the $n-th$ ECT

signal vector is represented by $\underline{\Psi}^{(n)} = \left\{ \left(\Re \left\{ \Psi_k^{(n)} \right\}; \Im \left\{ \Psi_k^{(n)} \right\} \right); k = 1, ..., K \right\}$ of $F = 2K$ ECT features.

Similar to GRID sampling, LHS method is easy to apply. As, LHS does not depend on the uniform spacing of each defect parameter, the curse of dimensionality problem of GRID sampling strategy in $\aleph_p$ can be mitigated by LHS. Like GRID sampling, it does not keep into account how ECT signals/data are varied in the parameter space. Therefore, it requires high number of training samples for obtaining accurate training model. Due to the random sampling ability of the parameter space, we employ LHS for unknown test set generation for evaluating the developed training model [25, 26, 31]. Fig. 3.1 (b) shows distribution of groove parameters (i.e., height and width) in the parameter space $\aleph_p$ by applying LHS strategy for $N = 49$ samples. It is worth to mention that other kinds of pseudo-random sequences can be employed instead of LHS with LBE paradigm. That is, Halton and Sobol sequences [54] are widely applied in the standard design of experiment sampling schema.

3.2.3 Adaptive sampling

The distribution of parameters in the parameter space plays important role for suitable training set generation. In [45], authors have described an adaptive sampling strategy through sequential sampling algorithm by which an initial input space (i.e., parameter space) is generated and the corresponding output space (i.e., ECT signal space) is computed. On the parameter space each new sample is added such that the resultant new output sample is placed as far as from the existing output samples in the output space. This adaptive schema is also known as output space filling (OSF).

3.2.4 Limitations of standard sampling strategies

Due to the use of higher number of ECT features, by applying GRID, LHS and OSF sampling, we need higher number of training samples for accurate training model development by the learning algorithm. That means all of these sampling methods suffer for curse of dimensionality problem. This makes it difficult for the learning algorithm to build optimal training model by increasing learning complexity as well as training time (offline phase). This reflects higher prediction error during online phase for an unknown test sample. In section 3.3, we have shown how these sampling techniques can be improved by applying feature extraction techniques.

3.3 Training set generation by applying feature extraction and advanced sampling strategy

In this section, the described standard sampling strategies have been enhanced by applying feature extraction technique.

3.3.1 Importance of feature extraction

Application of feature extraction is very common in the machine learning community. Learning algorithms suffer from dealing with higher dimensional data set in order to develop the learning model. A plethora of feature extraction techniques are available in literature which are useful in different contexts and can be differentiated as supervised and unsupervised feature extraction techniques. Among different feature extraction methods, Principal Component Analysis (PCA) [55], Partial Least Squares (PLS) [56], Linear discriminant analysis (LDA) [57], Canonical Correlation Analysis (CCA) [58] and their nonlinear extension [44] are commonly used in machine learning. Applying PCA for reducing higher dimensional problem considers the exploitation of the input features (i.e., ECT signals) by disregarding the output (e.g., defect parameters). This is also known as an unsupervised dimensionality reduction approach. PCA exploits the correlations between the ECT features and maximize the variance in order to retrieve the orthogonal set of vectors (eigen vectors). Since, we deal with supervised learning algorithm, it is convenient to deal with supervised feature extraction. That means, dealing with parametric problem, we need to consider both the ECT signals and their associated parameters. Therefore, we are interested to have the projections that account for both the variation of ECT signals and the associated defect parameters. PLS extracts the projections that maximize the co-variance between projected input features and output parameters. Therefore, we have chosen PLS as the suitable feature extraction technique in our research work. Other types of supervised feature extraction techniques such as CCA projects input features by maximizing the correlation between the input features and output parameters. The analyzed scheme can be straightforwardly applied to any of the aforementioned methods. However, in the next section, application of PLS feature extraction technique is explored within the framework of LBE.

3.3.2 PLS feature extraction

PLS regression algorithm was first introduced by Wold [56] as an alternative approach to ordinary least squares regression. A dimension reduction technique is processed at the core of PLS regression algorithm. The reduction of the original ECT feature space into a smaller dimension $J < F$ can be formulated by means of a customized version of the PLS algorithm as follows.

Let have a matrix of defect parameters, $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N)$ having $(N \times Q)$ dimension, where $\underline{p}^{(n)}$ is the n-th row of $\underline{\underline{p}}$. By using $\Phi\{.\}$ generate ECT coil signals

and fill the $(N \times F)$ feature matrix $\underline{\underline{\Psi}} = (\underline{\Psi}^{(n)}; n = 1, ..., N)$ whose n-th row is represented by $\underline{\Psi}^{(n)} = \left\{\left(\Re\left\{\Psi_k^{(n)}\right\}; \Im\left\{\Psi_k^{(n)}\right\}\right); k = 1, ..., K\right\}$ of $F = 2K$ ECT features. A $(N \times F)$ matrix $\underline{\underline{\Psi}}'$ is built by subtracting to each f-th column of the original ECT feature matrix $\underline{\underline{\Psi}} = \overline{\Psi}_f\,(f = 1, ..., F)$ its mean value, $\mu(\overline{\Psi}_f)$ through Eq. (3.1). Similarly, a $(N \times Q)$ matrix $\underline{\underline{p}}'$ is built by subtracting to each q-th column of the parameter matrix $\underline{\underline{p}} = \overline{p}_q;\ (q = 1, ..., Q)$ its mean value $\mu(\overline{p}_q)$ and can be defined in Eq. (3.2).

$$\Psi'_{nf} = \Psi_{nf} - \mu(\overline{\Psi}_f);\ f = 1, ..., F;\ n = 1, ..., N. \tag{3.1}$$

$$p'_{nq} = p_{nq} - \mu(\overline{p}_q);\ q = 1, ..., Q;\ n = 1, ..., N. \tag{3.2}$$

where, $\overline{\Psi}_f = \Psi_{nf};\ n = 1, ..., N$ and $\overline{p}_q = p_{nq};\ n = 1, ..., N$ are column vectors with the corresponding average, $\mu(\overline{\Psi}_f) = \frac{1}{N}\Sigma_{n=1}^{N}\Psi_{nf};\ (f = 1, ..., F)$ and $\mu(\overline{p}_q) = \frac{1}{N}\Sigma_{n=1}^{N}p_{nq};\ (q = 1, ..., Q)$. Apply the PLS algorithm to linearly decompose $\underline{\underline{\Psi}}'$ and $\underline{\underline{p}}'$ as follows

$$\underline{\underline{\Psi}}' = \underline{\underline{T}} \times \underline{\underline{S}} + \underline{\underline{Y}};\quad \underline{\underline{P}}' = \underline{\underline{U}} \times \underline{\underline{Z}} + \underline{\underline{G}}. \tag{3.3}$$

$\underline{\underline{T}} = (\underline{T}^{(n)}; n = 1, ..., N)$ in Eq. (3.3) is the $(N \times J)$ matrix of Ψ-scores [$\underline{T}^{(n)} = \left(T_j^{(n)}; j = 1, ..., J\right)$]. It is obtained from $\underline{\underline{\Psi}}'$ through the $(F \times J)$ weight matrix $\underline{\underline{W}}$. J is the number of extracted features ($J < F$). $\underline{\underline{Y}}$ and $\underline{\underline{G}}$ contain the $(N \times F)$ and $(N \times Q)$ residuals of the linear decomposition while $\underline{\underline{S}}$ and $\underline{\underline{Z}}$ are the $(J \times F)$ and $(J \times Q)$ matrices of loadings. The decomposition in Eq. (3.3) is aimed at maximizing the co-variance between the corresponding columns of $\underline{\underline{T}}$ [i.e., $\left(\overline{T}_j; j = 1, ..., J\right)$] and of the $(N \times J)$ matrix of p-scores $\underline{\underline{U}}$ [i.e., $\left(\overline{U}_j; j = 1, ..., J\right)$]. This guarantees all the information about the ECT signal embedded inside $\underline{\underline{\Psi}}'$ (i.e., inside $\underline{\underline{\Psi}}$) is compressed into $\underline{\underline{T}}$. Among different iterative algorithms, we have used SIMPLS algorithm [59] to obtained the weight matrix $\underline{\underline{W}}$. The main objective of this algorithm is to find the $\underline{\underline{W}}$ that linearly project $\underline{\underline{\Psi}}'$ into $\underline{\underline{T}}$ where, $\underline{\underline{T}} = \underline{\underline{\Psi}}' \times \underline{\underline{W}}$. The iterative procedure is briefly described below.

SIMPLS Algorithm:

1. A matrix $\underline{\underline{R}} = (\underline{\underline{\Psi}}')^{\mathrm{T}}\underline{\underline{P}}'$ is computed, $()^{\mathrm{T}}$ stands for transpose operator.

2. Run a for loop over j- For $j = 1, ..., J$

- Apply the Singular Value Decomposition (SVD) to $\underline{\underline{R}}$.
  If $j > 1$, $\underline{\underline{R}} = \underline{\underline{R}} - \left(\underline{\underline{S}}\right)^{\mathrm{T}}\left(\underline{\underline{S}}\left(\underline{\underline{S}}\right)^{\mathrm{T}}\right)^{-1}\underline{\underline{SR}}$;

- The j-th column of $\underline{\underline{W}}$ denoted by $\overline{W}_j = (W_{fj};\ f = 1, ..., F)$, is set equal to the first left singular vector, where $\underline{\underline{W}} = \left(\overline{W}_j;\ j = 1, ..., J\right)$.

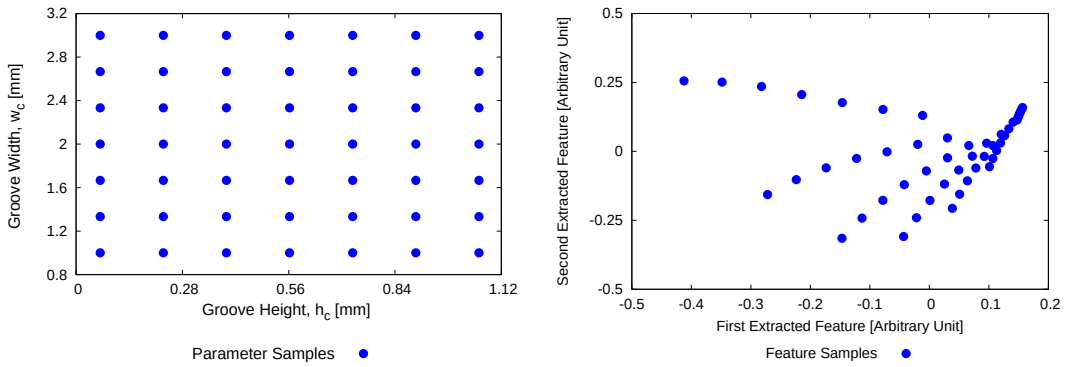
- The j-th Ψ- score is computed by $\overline{T}_j = \underline{\underline{\Psi}}'\overline{W}_j$ while, $\underline{\underline{T}} = \left(\overline{T}_j;\, j = 1, ...,\, J\right)$, $\overline{T}_j = (T_{nj};\, n = 1, ...,\, N)$ and $\overline{W}_j = (W_{fj};\, f = 1, ...,\, N)$.
- Compute the j-th row of $\underline{\underline{S}}$ such that $\underline{S}_j = \left(\left(\left(\underline{\underline{\Psi}}'\right)^{\mathrm{T}}\overline{T}_j\right)^{\mathrm{T}} / \left(\overline{T}_j\right)^{\mathrm{T}}\overline{T}_j\right)$
- The loop continues iteratively until $j = J$.

3.3.3 GRID-PLS sampling

GRID-PLS sampling strategy consists the combination of GRID sampling strategy for sampling the defect parameter space $\aleph_p$ and PLS feature extraction for extracting the suitable features from the ECT signal space $\aleph_\Psi$ through projecting to an extracted feature space $\aleph_T$. At the preliminary stage of GRID-PLS sampling, N samples are generated by using a uniform GRID sampling method within the defect parameter space $\aleph_p$. A matrix of defect parameters $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N)$ having $(N \times Q)$ dimension is formed, where $\underline{p}^{(n)}$ is the n-th row of $\underline{\underline{p}}$. By using $\Phi\{.\}$, ECT coil signals are generated and fill the $(N \times F)$ feature matrix $\underline{\underline{\Psi}}$, whose n-th row is represented by $\underline{\Psi}^{(n)}$. The n-th ECT signal vector is represented by $\underline{\Psi}^{(n)} = \left\{\left(\Re\left\{\Psi_k^{(n)}\right\}; \Im\left\{\Psi_k^{(n)}\right\}\right); k = 1, ...,\, K\right\}$ of $F = 2K$ ECT features. A data set of $\hat{D}_N = \left\{\left(\underline{\Psi}^{(n)}; \underline{p}^{(n)}\right); n = 1, ...,\, N\right\}$ has been obtained and PLS algorithm (See Sec. 3.3.2) is applied on $\hat{D}_N$ in order to reduce the higher dimensional ECT feature space. Therefore, the ECT signals embedded inside the matrix $\underline{\underline{\Psi}}$ are compressed into the reduced extracted feature matrix $\underline{\underline{T}}$, where $\underline{\underline{T}} = \left(\underline{T}^{(n)};\, n = 1, ...,\, N\right)$. Finally, construct the q-th training set $D_{N_q} = \left\{\left(\underline{T}^{(n)}; p_q^{(n)}\right);\, n = 1, ...,\, N\right\}$ for each defect parameter from Q.

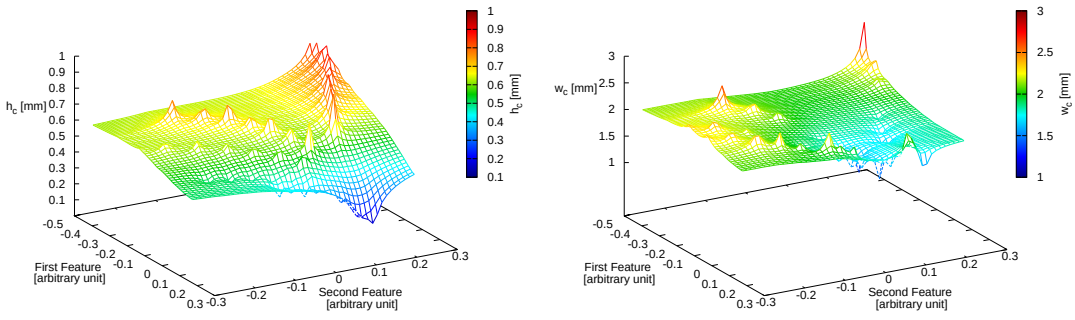
By reducing the ECT feature space, GRID-PLS mitigates the curse of dimensionality problem, due to the higher dimensional ECT feature space. PLS is also robust to deal with noisy signals/data. Therefore, GRID-PLS also may reduce the impact of noisy ECT features (i.e., corrupted ECT signals due to the presence of noise) during feature extraction. The robustness of GRID-PLS sampling over GRID sampling approach has been shown in [26], with SVR regression technique for crack characterization and localization problems within a tube structure. Even though GRID-PLS has shown improved prediction accuracy than GRID or LHS approaches, it also suffers from dealing with higher parametric inversion. As like GRID sampling, with increasing the number of parameters (i.e., Q) to be inverted, the number of total samples N increases exponentially which requires higher training set generation time and increases the complexity of learning model. Figure 3.2 shows the sampling distribution of groove parameters in $\aleph_p$ and the resultant extracted features in $\aleph_T$. The parameter space is homogeneously sampled while the extracted features associated to it is oddly sampled. $\aleph_T$ also exhibits patterns of different clusters (Fig. 3.2 (b)), where some

zones are under sampled and some zones are over sampled. On the other hand, PLS extracted the most suitable features from the ECT feature space by considering the contribution of defect parameters. ECT signals do not vary uniformly due to the variation of each defect parameter. For example, the impacts of ECT signals/data due to the variation of groove height is higher than the variation of groove width. With the increase of groove height h_c, eddy current signals are impacted with strong impedance variation of the inspecting coils. In contrast, for the considered parameters range, ECT signals do not change significantly due to the variation of crack width w_c. Since, the inspecting probes axial extension can be longer than the groove width, the prediction accuracy on crack width estimation is somehow bounded by the spatial resolution of the inspecting probes for this problem.



(a) (b)

Figure 3.2: Sample locations mapped on (a) parameter space $\aleph_p$ and (b) the resultant extracted feature space through GRID-PLS sampling.

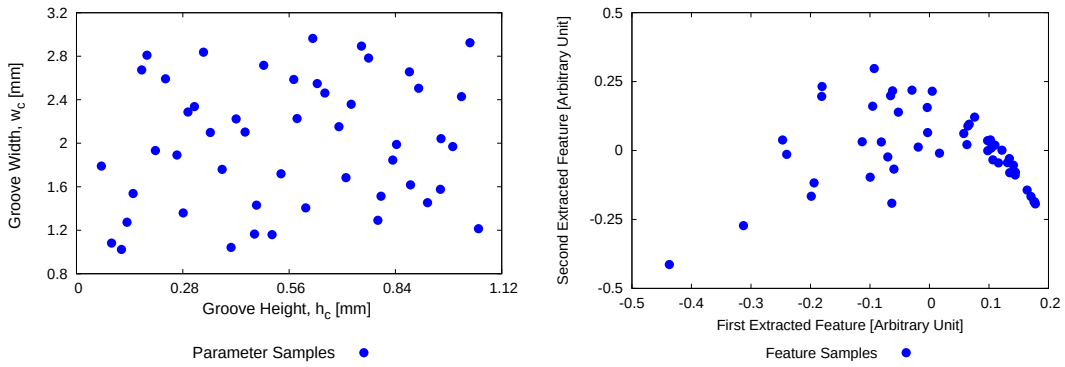


(a) (b)

Figure 3.3: Training set representation for groove (a) height h_c and (b) width w_c generated by GRID-PLS sampling.

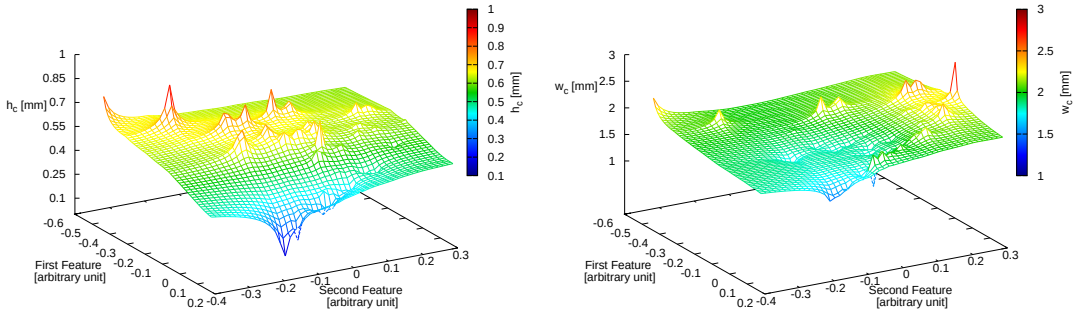
We can observe these phenomenon from Fig. 3.3 that with the variation of groove height h_c, ECT signals and consequently extracted features vary significantly com-

pared to groove width w_c. That means, the extracted features are significantly varied by changing groove height h_c (Fig. 3.3 (a)), while by varying w_c, the extracted features representation is almost flat Fig. 3.3 (b). This degrades the learning ability for the learning algorithm and degrades prediction accuracy for groove width estimation. That means GRID-PLS requires redundant training samples which can be avoided by intelligently and uniformly sampling the extracted feature space. As a consequence, more robust training model can be built with higher prediction accuracy during inversion. One of the ways to do this is to sample the crack parameter space in a way such that the resultant extracted feature space is uniformly sampled. In Sec. 3.3.5, we have shown an adaptive sampling strategy that samples the parameter space in a way that the extracted feature space is uniformly sampled.



(a) (b)

Figure 3.4: Sample locations mapped on (a) parameter space $\aleph_p$ and (b) the resultant extracted feature space through LHS-PLS sampling.



(a) (b)

Figure 3.5: Training set representation for groove (a) height h_c and (b) width w_c generated by LHS-PLS sampling.

3.3.4 LHS-PLS sampling

Similar to GRID-PLS sampling strategy, LHS-PLS sampling strategy can be applied by adopting LHS sampling for generating N number of samples in the parameter space instead of GRID sampling approach. LHS-PLS overcomes the limitation of curse of dimensionality problem for dealing with higher parametric inversion. Due to use of PLS feature extraction, higher dimensional ECT feature space can be reduced to lower dimension. However, due to the use of random sampling in parametric space, LHS-PLS can not assure higher prediction accuracy. Similar to GRID-PLS strategy, LHS-PLS sampling can not assure higher prediction accuracy for certain defect parameter (e.g., groove width w_c). From Fig. 3.4 we can see that the parameter space for the groove is randomly sampled while the resultant extracted feature space is not uniformly sampled. Similar to GRID-PLS sampling, Fig. 3.5 (a) also demonstrates that by sampling the parameter space randomly can not assure the contribution of groove width for significant changing of ECT signals and the resultant extracted features (Fig. 3.5 (b)). However, for only few higher values of w_c, extracted features are varied. Which also demonstrates, lower values of w_c do not contribute significantly for changing ECT signals that results lower variation between extracted features. This also degrades the learning ability of the learning model. One of the ways to improve this phenomenon is to sample groove parameter space so as to represent enough impedance variation information due to the variation of w_c. This can also help the learning algorithm to build optimal model for predicting w_c with higher accuracy.

3.3.5 PLS combined with output space filling (PLS-OSF) sampling

The main goal of this advanced sampling is to apply PLS feature extraction on the ECT feature space to reduce the dimension of actual ECT feature space and perform adaptive sampling directly in the extracted feature space in order to retrieve the lowest number of training samples during training phase. This provides an exhaustive representation of the I/O relationship for optimal and almost real time inverse solution during testing phase. The following steps describe the iterative procedure and Fig. 3.6 represents the corresponding flow chart in order to build suitable I/O pairs for building optimal training model by PLS-OSF approach.

1. **Initialization**- Generate N_0 number of initial samples by using a uniform GRID sampling within the defect parameter space $\aleph_p$. Thus, a matrix of defect parameters $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N_0)$ having $(N \times Q)$ dimension is formed, where $\underline{p}^{(n)}$ is the n-th row of $\underline{\underline{p}}$. By using $\Phi\{.\}$ generate ECT coil signals and fill the $(N \times F)$ feature matrix $\underline{\underline{\Psi}}$, whose n-th row is represented by $\underline{\Psi}^{(n)}$. Due to the complex nature of the ECT signals, the $n-th$ ECT signal vector is represented by $\underline{\Psi}^{(n)} = \left\{ \left(\Re \left\{ \Psi_k^{(n)} \right\}; \Im \left\{ \Psi_k^{(n)} \right\} \right); k = 1, ..., K \right\}$ of $F = 2K$ ECT features.

2. **Feature Extraction**- Build the $(N_0 \times F)$ matrix $\underline{\underline{\Psi}}'$ by subtracting to each f-th column of $\underline{\underline{\Psi}}$ $(f = 1, ..., F)$ its mean value, μ_f and compute the $(N_0 \times Q)$ ma-

trix $\underline{\underline{p}}'$ by subtracting to each q-th column of $\underline{\underline{p}}$ its mean value μ_q $(q = 1, ..., Q)$. Apply the PLS algorithm to linearly decompose $\underline{\underline{\Psi}}'$ and $\underline{\underline{p}}'$ by recalling Eq. 3.3 and obtain the reduced extracted feature matrix $\underline{\underline{T}}$ such that all the information about the ECT signals embedded inside $\underline{\underline{\Psi}}'$ (i.e., inside $\underline{\underline{\Psi}}$) are compressed into $\underline{\underline{T}}$. Assign the number of training sample $N_{iterative} = N_0$ and construct an initial training set $\hat{D}_{N_{iterative}} = \left\{ \left(\underline{T}^{(n)}; \underline{p}^{(n)} \right); n = 1, ..., N \right\}$ for the adaptive step.

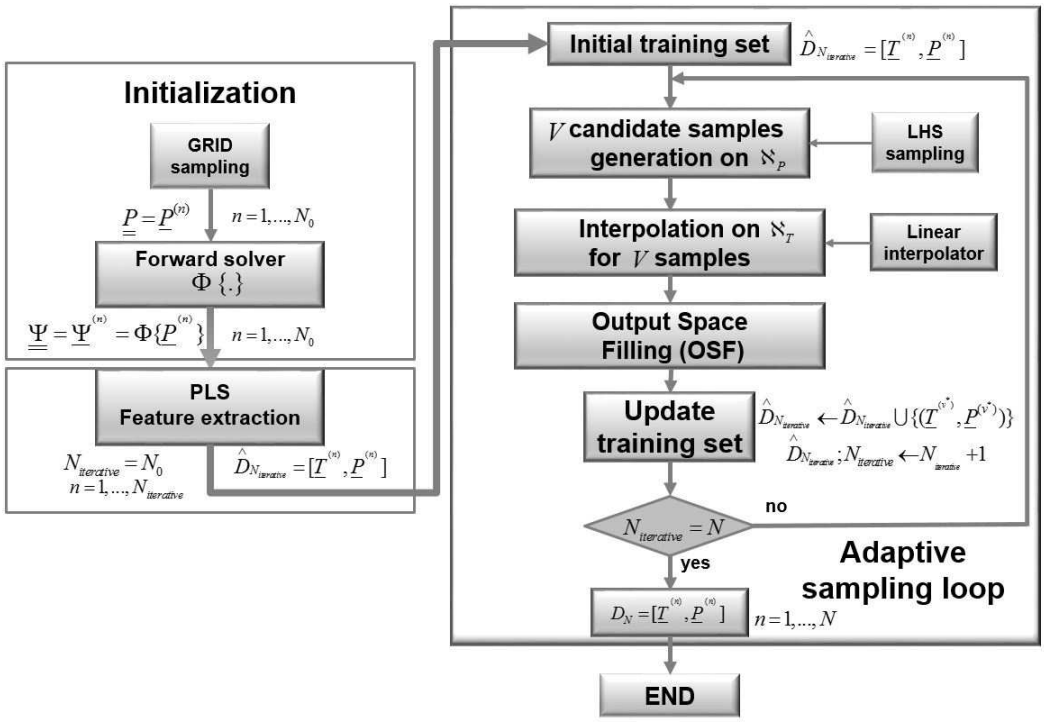


Figure 3.6: Flow chart representation of PLS-OSF sampling.

3. **Adaptive Sampling-** By using LHS, generate V candidate samples within the parameter space (i.e., $\aleph_p$) by $\underline{p}_{cand}^{(v)} = \left(p_{cand,q}^{(v)}; q = 1, ..., Q \right)$, where $v = 1, ..., V$. An estimation of the J-dimensional set of extracted features corresponding to each v-th candidate, $\underline{\tilde{T}}_{cand}^{(v)}$ is retrieved by applying a multi-dimensional linear interpolator on $\hat{D}_{N_{iterative}}$. Select the optimal $v = v_{opt}$ candidate (i.e., $\underline{p}_{cand}^{(v_{opt})}$) from V such that the minimum distance between the obtained extracted features $\underline{\tilde{T}}_{cand}^{(v_{opt})}$ and all the available extracted features $\underline{T}^{(n)}$ $(n = 1, ..., N_{iterative})$ within $\hat{D}_{N_{iterative}}$ is maximized [i.e., $v_{opt} = \arg\left(\max_{v=1,...,V} \left\{\min_{n=1,...,N} [d_{vn}]\right\}\right)$]. d_{vn} is the Euclidean distance between $\underline{\tilde{T}}_{cand}^{(v)}$ $(v = 1, ..., V)$ and $\underline{T}^{(n)}(n = 1, ..., N_{iterative})$, which can be described by $d_{vn} = \sqrt{\sum_{j=1}^{J} \left(\tilde{T}_{cand,j}^{(v)} - T_j^{(n)} \right)^2}$.

The ECT features $\underline{\Psi}_{cand}^{(v_{opt})} = \left\{ \left(\Re \left\{ \Psi_{cand,k}^{(v_{opt})} \right\}; \Im \left\{ \Psi_{cand,k}^{(v_{opt})} \right\} \right); k = 1, ..., K \right\}$ associated to the selected candidate sample is computed by utilizing $\Phi\{.\}$. Finally, the set of extracted features is obtained

$$\underline{T}_{cand}^{(v_{opt})} = \left(\underline{\Psi}_{cand}^{(v_{opt})} \right)' \times \underline{\underline{W}} \tag{3.4}$$

where $\left(\underline{\Psi}_{cand}^{(v_{opt})} \right)'$ is obtained by subtracting μ_f to each f-th element of $\underline{\Psi}_{cand}^{(v_{opt})}$ $(f = 1, ..., F)$. Finally, updated the training set with $\hat{D}_{N_{iterative}+1} = \hat{D}_{N_{iterative}} \cup \left\{ \underline{T}_{cand}^{(v_{opt})}; \underline{p}_{cand}^{(v_{opt})} \right\}$ and update $N_{iterative} = N_{iterative} + 1$. This is also known as the customized version of OSF (i.e., candidate parameters are chosen such that features are uniformly distributed).

4. **Stop Criterion**- The adaptive sampling step adds new sample iteratively until $N_{iterative} = N$ (N is desired/feasible training size).

Finally training sets are generated by exploiting the corresponding q-th set of I/O pairs $\hat{D}_{N,q} = \left\{ \left(\underline{T}^{(n)}; p_q^{(n)} \right); n = 1, ..., N \right\}$ of each defect parameter. By uniformly distributing the extracted feature space, redundant and unnecessary samples for different configuration of defect parameters can be avoided. This improves the prediction accuracy of the learning algorithm.

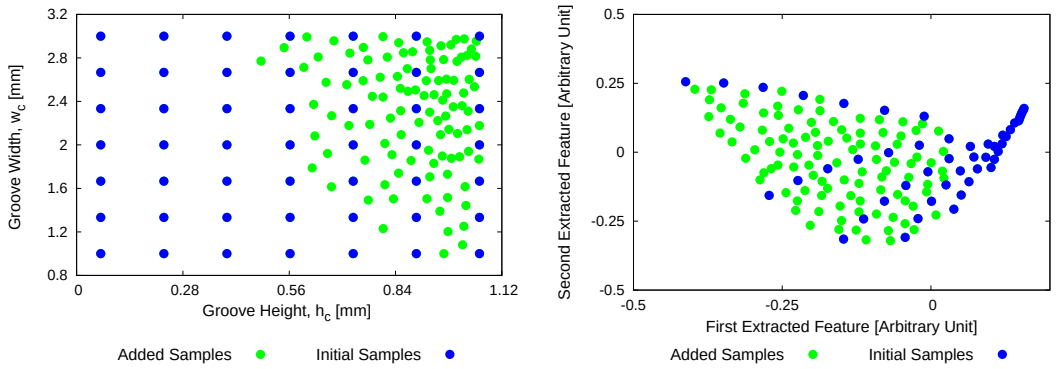
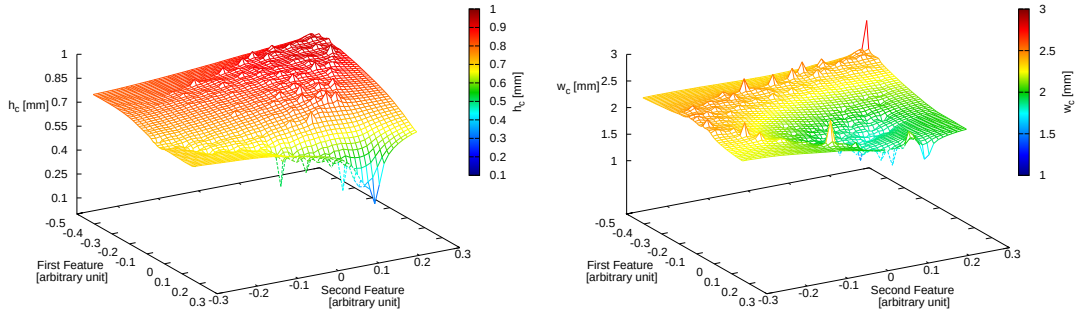


Figure 3.7: Sample locations mapped on $a)$ parameter space $\aleph_p$ and $b)$ the resultant extracted feature space through PLS-OSF sampling.



(a) (b)

Figure 3.8: Training set representation for groove $a)$ height h_c and $b)$ width w_c generated by PLS-OSF sampling.

In [27, 28, 31], the robustness of PLS-OSF is shown on simulated as well as experimental data by addressing different NDE problems. Considering the groove problem at hand, Fig. 3.7 shows that the groove parameter space is sampled in a way that the resultant extracted feature space is evenly sampled. Similarly, Fig. 3.8 illustrates that, PLS-OSF sampling strategy provides the extracted features which are varied according to the contribution of both groove parameter (i.e., h_c and w_c). By performing adaptive sampling, the extracted feature space is uniformly sampled. As most significant, ECT information are extracted by PLS feature extraction, the parameter space is filled in such a way that almost all adaptive samples are added for the higher values of h_c and different values of w_c parameter. As higher w_c contributes mostly for changing ECT signals, extracted features (Fig. 3.8 (b)) are varied within the zone having higher w_c values than the zone having lower w_c values. This improves the learning ability of the algorithm, and as a consequence, a more accurate model can be obtained from smaller training sets [27]. In the next Chapter, inverse model generation procedures have been discussed by applying different sampling strategies.

Chapter 4

Inverse model generation through LBE

This chapter describes inverse model generation procedures based on the sampling strategies described in Chapter 3. To build accurate training model, it is crucial to obtain optimal hyper parameters of both the prediction and feature extraction techniques. Therefore, different calibration strategies have been discussed in order to retrieve optimal parameters which will be followed by error metric definitions.

4.1 Problem definition

By recalling the same problem described in Chapter 2, lets consider a conductive tube is affected by a single axisymmetric groove (i.e., defect) having $Q = 2$ parameters $\underline{p} = (h_c,\, w_c)$ (Fig. 2.1). h_c and w_c are the height and width of the groove, respectively. The ECT signal collected by the coils at the k-th ($k = 1, ...,\, K$) scanning position is represented by $\underline{\Psi} = \{(\Re\{\Psi_k\}\,; \Im\{\Psi_k\})\,; k = 1, ...,\, K\}$, where $\underline{\Psi} \in \mathbb{R}^{1\times F}$. Hence, the ECT signal space $\aleph_\Psi$ having $dim\,(\aleph_\Psi) = 2K$, contains $F = 2K$ features.

4.2 Inverse model definition and testing

In this section we have defined different training models by applying various sampling strategies that have already been shown in Chapter 3.

4.2.1 Inversion through high dimensional ECT signals

Lets start from the basic GRID sampling strategy described in Sec. 3.2.1. N number of samples have been generated through varying defect parameters within the parameter space $\aleph_p$. A matrix of defect parameters, $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N)$ having $(N \times Q)$ dimension is obtained by full factorial GRID sampling, where $\underline{p}^{(n)}$ is the n-th row of $\underline{\underline{p}}$. By using $\Phi\{.\}$ (i.e., CIVA), the ECT signals are generated and filled a $(N \times F)$ matrix $\underline{\underline{\Psi}} = (\underline{\Psi}^{(n)}; n = 1, ..., N)$ of $F = 2K$ features, whose n-th row is represented by $\underline{\Psi}^{(n)} = \left\{ \left(\Re \left\{ \Psi_k^{(n)} \right\}; \Im \left\{ \Psi_k^{(n)} \right\} \right); k = 1, ..., K \right\}$. Finally, the q-th training set $(q = 1, ..., Q)$ is build for each defect parameter $D_{N_q} = \left\{ \left(\underline{\Psi}^{(n)}; p_q^{(n)} \right); n = 1, ..., N \right\}$ to have the q-th separate training model for each parameter. As we have discussed in Sec. 2.2.4, once we have built training sets, different prediction techniques can be fitted for building training (i.e., inversion) models. However, we have selected Support Vector Regression (SVR) as one of the suitable learning algorithms, hence the mathematical formulation of SVR is elaborated in the next section. More details of some other prediction techniques are also available in [25].

Having a q-th trained model by SVR we are interested in evaluating the inversion performance of the models. Hence, a m-th $(m = 1, ..., M)$ test sample $\underline{\Psi}^{(m)}$ of F ECT features associated to a previously-unseen crack parameters configuration $\underline{p}^{(m)} = p_q^{(m)}; q = 1, ..., Q$ is generated. Finally, $\underline{\Psi}^{(m)}$ is given as input to the built SVR training model in order to estimate the q-th parameter of the crack, $\tilde{p}_q^{(m)}$.

4.2.1.1 Training model by Support Vector Regression (SVR)

Support Vector (SV) is the algorithm of a non-linear generalization of the generalized portrait algorithm developed in Russia [60]. It is firmly developed on basis of the statistical learning theory which has been developed by Vapnik, Chervonenkis and others (known as VC theory). VC theory characterizes the properties of the learning machines which enables them to generalize to unseen data.

4.2.1.2 Linear model

To be general, let the input vector $\underline{\Psi}^{(n)}$ of n-th sample (containing ECT signals) of F dimension (i.e., $\underline{\Psi} \in \mathbb{R}^{1 \times F}$), and the corresponding real scalar output $p_q^{(n)}$ for the q-th defect parameter. Therefore, the corresponding training set can be defined as

$$D_{N_q} = \left\{ \left(\underline{\Psi}^{(n)}; p_q^{(n)} \right); n = 1, ..., N; q = 1, ..., Q \right\}, \quad \underline{\Psi}^{(n)} \in \aleph_\Psi; p_q^{(n)} \in \aleph_p \quad (4.1)$$

where, $\aleph_\Psi \subset \mathbb{R}^{N\times F}$ is the input ECT signal space and $\aleph_p \subset \mathbb{R}^{N\times Q}$ is the parameter space. SVR searches a function $\Theta_q(\underline{\Psi}^{(n)})$ that reflects the relationship between the input $\underline{\Psi}^{(n)}$ and the output $p_q^{(n)}$ in the best way. Lets start with the linear function

$$\Theta_q(\underline{\Psi}^{(n)}) = \underline{\omega} \cdot (\underline{\Psi}^{(n)}) + b,\ \underline{\omega} \in \mathbb{R}^F, b \in \mathbb{R}^N \tag{4.2}$$

where $\underline{\omega}$, b are the coefficients of the developed model. The difference between the actual output $p_q^{(n)}$ for the input data $\underline{\Psi}^{(n)}$ and the predicted output $\Theta_q(\underline{\Psi}^{(n)})$ is evaluated by means of ε-insensitive loss function (Eq. 4.3). That means SVR allow us to have a margin ε within which errors are accepted within the sample data without affecting prediction as [61, 62]

$$\mathcal{L}\left(p_q^{(n)},\ \Theta_q(\underline{\Psi}^{(n)})\right) = \mathcal{L}\left(\left|p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)})\right|_\varepsilon\right), \tag{4.3}$$

where

$$\left|p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)})\right|_\varepsilon = \begin{cases} 0, & if\ \left|p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)})\right| \leq \varepsilon, \\ \left|p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)})\right| - \varepsilon, & \text{otherwise} \end{cases}$$

Flatness of (4.2) means we seek for a smaller $\underline{\omega}$ value. Through minimizing the Euclidean norm $||\underline{\omega}||^2$ we can minimize the model complexity (the flatter the function the simpler it is and also it becomes more general). To fit the model to the available training set D_{N_q} we have to find the optimal values of $\underline{\omega}$ and b by using Empirical Risk Minimization (ERM) principle,

$$(\underline{\omega}^{opt},\ b^{opt}) = \arg\min_{\underline{\omega},b} \sum_{n=1}^{N} |p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)})|_\varepsilon. \tag{4.4}$$

This type of optimization refers the existence of a function which approximates all $p_q^{(n)}$ within a precision ε. This is also to mention that the minimization is also a trade off between model complexity and the error on D_{N_q}. SVR approach consists in minimizing the trade off between the complexity of the model and the degree to which errors larger than ε can be tolerated by tuning a user defined constant parameter C ($C \geq 0$).

$$(\underline{\omega}^{opt},\ b^{opt}) = \arg\min_{\underline{\omega},b} \left[\frac{1}{2}||\underline{\omega}||^2 + C\sum_{n=1}^{N} |p_q^{(n)} - \Theta_q(\underline{\Psi}^{(n)})|_\varepsilon\right], \tag{4.5}$$

$$(\underline{\omega}^{opt},\ b^{opt}) = \arg\min_{\underline{\omega},b} \left[\frac{1}{2}||\underline{\omega}||^2 + C\sum_{n=1}^{N} |p_q^{(n)} - \underline{\omega} \cdot \underline{\Psi}^{(n)} - b|_\varepsilon\right].$$

Sometimes it is also possible to have better prediction while we allow the possibility of outliers. In [46] the authors adopted soft margin loss function by introducing slack

variable to cope with the in-feasible constrains of the optimization problem. Slack variables are added in inequality optimization problem to transform it to equality. Hence Eq. (4.5) can be rewritten as

$$(\underline{\omega}^{opt},\, b^{opt}) = \arg\min_{\underline{\omega},b,\xi_i,\xi_i^*} \left[\frac{1}{2}||\underline{\omega}||^2 + C\sum_{n=1}^{N}(\xi_n + \xi_n^*)_\varepsilon\right], \tag{4.6}$$

subject to

$$y_n - \underline{\omega}\cdot\underline{\Psi}^{(n)} - b \leq \varepsilon + \xi_n,$$

$$\underline{\omega}\cdot\underline{\Psi}^{(n)} + b - y_n \leq \varepsilon + \xi_n^*,$$

$$\xi_n,\, \xi_n^* \geq 0 \quad n = 1, ..., N,$$

where ξ_n, ξ_n^* are slack variables which are used to solve the above constraint quadratic problem. This problem is also called the primal which can be solved by using Lagrange multipliers theory.

4.2.1.2.1 Lagrange multipliers Lagrange multipliers are used to find maxima / minima of multivariate functions that are subject to constraint. For example, if $F(x_1,\, x_2)$ is a function of two variables and $g(x_1,\, x_2)$ is another function of two variables, and we define a Lagrangian L

$$L(x_1,\, x_2,\, z) = F(x_1,\, x_2) + zg(x_1,\, x_2).$$

$(x_1,\, x_2)$ is a relative extremum of F subject to $g(x_1,\, x_2) = 0$, then there is some value that $z = \tau$ such that

$$\frac{\delta L}{\delta x_1}|_{(x_1,x_2,\tau)} = \frac{\delta L}{\delta x_2}|_{(x_1,x_2,\tau)} = \frac{\delta L}{\delta z}|_{(x_1,x_2,\tau)} = 0.$$

Similarly we can proceed from Eq. (4.6) by introducing α_n, α_n^*, η_n, η_n^* as Lagrange multipliers to have the corresponding Lagrangian (L),

$$L = \frac{1}{2}||\underline{\omega}||^2 + C\sum_{n=1}^{N}(\xi_n + \xi_n^*) - \sum_{n=1}^{N}\alpha_n(\varepsilon + \xi_n - p_q^{(n)} + \omega\cdot\underline{\Psi}^{(n)} + b) \tag{4.7}$$

$$-\sum_{n=1}^{N}\alpha_n^*(\varepsilon + \xi_n^* + p_q^{(n)} - \underline{\omega}\cdot\underline{\Psi}^{(n)} - b) - \sum_{n=1}^{N}(\eta_n\xi_n + \eta_n^*\xi_n^*).$$

By introducing non-negative Lagrange multipliers we will try to get the optimal value of $\underline{\omega}$, b, ξ_n, ξ_n^* (are also known as primal) that minimize the Lagrangian. This approach leads to Lagrange dual problem and the mentioned solution gives the primal variable as the function of dual variables (Lagrange multipliers α_n, α_n^*, η_n, η_n^*). In

optimization problem, solving one constrained problem often leads to solve another constrained problem, in general this is called duality. So the dual variables at Eq. (4.7) have to satisfy the positivity constraint

$$\alpha_n,\, \alpha_n^*,\, \eta_n,\, \eta_n^* \geq 0.$$

It follows from the saddle point condition that the partial derivatives of L with respect to the primal variables $(\underline{\omega},\, b,\, \xi_n,\, \xi_n^*)$ and equating to zero, we can obtain

$$\partial_b L = \sum_{n=1}^{N}(\alpha_n^* - \alpha_n) = 0. \tag{4.8}$$

$$\partial_{\underline{\omega}} L = \underline{\omega} - \sum_{n=1}^{N}(\alpha_n - \alpha_n^*)\underline{\Psi}^{(n)} = 0. \tag{4.9}$$

$$\partial_{\xi_n} L = C - \alpha_n - \eta_n = 0, \tag{4.10}$$

$$\eta_n = C - \alpha_n.$$

$$\partial_{\xi_n^*} L = C - \alpha_n^* - \eta_n^* = 0, \tag{4.11}$$

$$\eta_n^* = C - \alpha_n^*.$$

and the corresponding dual can be obtained from Eq. (4.7) as

$$\max\ L = -\frac{1}{2}\sum_{n,o=1}^{N}(\alpha_n - \alpha_n^*)(\alpha_o - \alpha_o^*)\underline{\Psi}^{(n)} \cdot \underline{\Psi}^{(o)} - \varepsilon\sum_{n=1}^{N}(\alpha_n + \alpha_n^*) + \sum_{n=1}^{N} p_q^{(n)}(\alpha_n - \alpha_n^*), \tag{4.12}$$

which is subjected to

$$0 \leq \alpha_n,\, \alpha_n^* \leq C,$$

$$\sum_{n=1}^{N}(\alpha_n - \alpha_n^*) = 0.$$

From Eq. (4.12), we can observe that the dual variables $\eta_n,\, \eta_n^*$ are eliminated through conditions at Eq. (4.10) and Eq. (4.11), respectively. These variables did not appear in the dual objective function anymore and were present only in the dual feasibility conditions. Therefore, from Eq. (4.9), $\underline{\omega}^{opt}$ can be mentioned as linear combination of transformed input vector

$$\underline{\omega}^{opt} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\underline{\Psi}^{(n)}. \tag{4.13}$$

α_n^{opt} and α_n^{*opt} being the optimal α_n and α_n^* obtained from Eq. (4.12). Thus, a result by utilizing Eq. (4.2) the dual formulation can be used to express the predicted output $\tilde{p}_q^{(m)}$ of an input $\underline{\Psi}^{(m)}$ in terms of dual variables

$$\tilde{p}_q^{(m)} = \Theta_q(\underline{\Psi}^{(m)}) = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\underline{\Psi}^{(n)} \cdot \underline{\Psi}^{(m)} + b \tag{4.14}$$

This is also known as so called Support vector expansion, where $\underline{\omega}$ can be described as the linear combination of training samples $\underline{\Psi}^{(n)}$. Hence (4.14) can be described by the dot products between the training and test data, without the explicit computation of $\underline{\omega}$.

4.2.1.2.2 b **computation** b can be computed by exploiting Karush-Khun-Tucker (KKT) condition [46]. According to KKT, at the product between dual variables and the constraints must vanish at the solution point.

$$\alpha_n^{opt}(\varepsilon + \xi_n - p_q^{(n)} + \underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} + b) = 0, \tag{4.15}$$

$$\alpha_n^{*opt}(\varepsilon + \xi_n^* + p_q^{(n)} - \underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} - b) = 0, \tag{4.16}$$

also $\xi_n \xi_n^* = 0$, $\alpha_n^{opt} \alpha_n^{*opt} = 0$ (for every constrain, at-least one variable must be zero). Therefore,

$$(C - \alpha_n^{opt})\xi_n = \eta_n \xi_n = 0$$

$$(C - \alpha_n^{*opt})\xi_n^* = \eta_n^* \xi_n^* = 0$$

which leads to

$$\alpha_n^{opt}(\varepsilon - p_q^{(n)} + \underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} + b^{opt}) + \alpha_n^{opt}\xi_n = 0, \;\; with\, \xi_n = 0,\; \alpha_n^{opt} \in (0, C),$$

$$b^{opt} = p_q^{(n)} - \underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} - \varepsilon,\; \xi_n = 0, \;\; with\, \alpha_n^{opt} \in (0, C). \tag{4.17}$$

$$\alpha_n^{*opt}(\varepsilon + p_q^{(n)} - \underline{\omega} \cdot \underline{\Psi}^{(n)} - b) + \alpha_n^{*opt}\xi_n^* = 0, \;\; with\, \xi_n^* = 0,\; \alpha_n^{*opt} \in (0, C),$$

$$b^{opt} = p_q^{(n)} - \underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} + \varepsilon,\; \xi_n^* = 0, \;\; with\, \alpha_n^{*opt} \in (0, C). \tag{4.18}$$

From Eq. (4.15), for $|\underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} + b^{opt} + \xi_n - p_q^{(n)}| \geq \varepsilon$ Lagrange multipliers may be nonzero. Therefore, for all the samples within range of ε and the α_n^{opt}, α_n^{*opt} vanish. Similarly for $|\underline{\omega}^{opt} \cdot \underline{\Psi}^{(n)} + b^{opt} + \xi_n^* - p_q^{(n)}| < \varepsilon$ from Eq. (4.16), α_n^{opt}, α_n^{*opt} have to be zero for fulfilling KKT conditions. In this way a sparse expansion of $\underline{\omega}^{opt}$ is available in terms of $\underline{\Psi}^{(n)}$. That means we do not need all $\underline{\Psi}^{(n)}$ to described $\underline{\omega}^{opt}$. Finally the input samples comes with non vanishing coefficients are called Support vectors.

4.2.1.3 Nonlinear model

SV algorithm can be used for finding non-linear function. The training input data $\underline{\underline{\Psi}}$ is simply mapped into some feature space. Lets have a non-linear transformation $\phi(\underline{\underline{\Psi}}) : R^{N \times F} \longrightarrow \aleph_{\mathcal{H}}$, which maps the inputs $\underline{\underline{\Psi}}$ into high dimensional feature space, $\aleph_{\mathcal{H}}$. As we know SVR algorithm depends on the dot products between samples $\underline{\underline{\Psi}}$, the dot product of non-linear transformation can be represent by fulfilling the condition at Eq. (4.12)

$$\max\ L = -\frac{1}{2}\sum_{n,o=1}^{N}(\alpha_n - \alpha_n^*)(\alpha_o - \alpha_o^*)\phi(\underline{\Psi}^{(n)})\cdot\phi(\underline{\Psi}^{(o)}) \tag{4.19}$$

$$-\varepsilon\sum_{n=1}^{N}(\alpha_n + \alpha_n^*) + \sum_{n=1}^{N}p_q^{(n)}(\alpha_n - \alpha_n^*).$$

which is subjected to

$$0 \le \alpha_n, \alpha_n^* \le C,$$

$$\sum_{n=1}^{N}(\alpha_n - \alpha_n^*) = 0.$$

Similar to Eq. (4.13) and Eq. (4.14) we can find

$$\underline{\omega}^{opt} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\phi(\underline{\Psi}^{(n)}), \tag{4.20}$$

$$\tilde{p}_q^{(m)} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\phi(\underline{\Psi}^{(n)})\cdot\phi(\underline{\Psi}^{(m)}) + b^{opt}. \tag{4.21}$$

where b^{opt} can be calculated

$$b^{opt} = p_q^{(n)} - \underline{\omega}^{opt}.\phi(\underline{\Psi}^{(n)}) - \varepsilon,\ \xi_n = 0,\ \ with\,\alpha_n^{opt} \in (0, C)$$

$$b^{opt} = p_q^{(n)} - \underline{\omega}^{opt}.\phi(\underline{\Psi}^{(n)}) + \varepsilon,\ \xi_n^* = 0,\ \ with\,\alpha_n^{*opt} \in (0, C).$$

4.2.1.3.1 Kernel method Kernel methods are very popular as learning algorithms where it uses instance-based learning. Let $D_{N_q} = \left(\underline{\Psi}^{(n)}; p_q^{(n)}\right),\ n = 1, ..., N;\ q = 1, ..., Q$ is a training data set, where $p_q^{(n)}$ is the n-th output of the corresponding of n-th vector $\underline{\Psi}^{(n)}$. The available response estimate $\tilde{p}(\underline{\Psi}^{(m)})$ for a given value $\underline{\Psi}^{(m)}$ can be written as the weighted average of the training outputs $\{p_q^{(n)}\}\ \ (n = 1, ..., N)$

$$\tilde{p}_q(\underline{\Psi}^{(m)}) = \mathcal{F}(\underline{\Psi}^{(m)}) = \frac{\sum_{n=1}^{N}\kappa(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})p_q^{(n)}}{\sum_{n=1}^{N}\kappa(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})}. \tag{4.22}$$

The weight function $\kappa(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})$ is associated to each output value $p_q^{(n)}$ that depends on the location $\underline{\Psi}^{(n)}$ in the input space by the known variables and the location $\underline{\Psi}^{(m)}$ in the same space where prediction is to be done. This function $\kappa(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})$ is called kernel function and the method is known as kernel method. Often kernel function can be represented as like the following form

$$\kappa(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)}) = g(-\sigma * dist(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})),$$

where $dist(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})$ is the distance between $\underline{\Psi}^{(m)}$ and $\underline{\Psi}^{(n)}$, σ is the smoothing parameter. The estimated output $\tilde{p}_q(\underline{\Psi}^{(m)})$ defined in Eq. (4.22) can be obtained by the weighted average of $p_q^{(n)}$ $(n = 1, ..., N)$. With increasing number of samples, $dist(\underline{\Psi}^{(m)}, \underline{\Psi}^{(n)})$ is becoming small. Whereas, σ defines how fast the kernel value drops with the distance between points. In Tab. 4.1 different kernel functions formulation are shown.

4.2.1.3.2 Mercer theorem From Mercer theorem [61], we know that to guarantee a symmetric function $\kappa(x,\, y)$ in input space, we can represent an inner product in the feature space if

$$\int_x \int_y \kappa(x, y) g(x) g(y) dx dy \quad \geq 0,$$

is valid for all $g \neq 0$ for which $\int g^2(s) ds \ \leq \propto$. Then the kernel function κ can be extended in terms of ϕ_n. Hence we can get

$$\kappa(x,\, y) = \sum_{n=1}^{\infty} \lambda_n \phi_n(x) \phi_n(y) \quad \lambda_n \geq 0,$$

So the mapping from the input space to feature space should satisfy the kernel function κ as the inner product

$$\phi(x).\phi(y) = \sum_{n=1}^{\infty} \lambda_n \phi_n(x) \phi_n(y) = \kappa(x,\, y).$$

In our case, original ECT signals/data $\underline{\Psi}^{(n)}$ $(n = 1, 2, ..., N)$ belongs to the input space. Where feature space is a Hilbert space [63] which extends the vector algebraic methods from two or three dimensional space to spaces with finite or infinite dimension. The dimension is normally larger than the number of data points (N) and in some cases it can be infinity. Let us to have a non-linear transformation of the input sample $\underline{\Psi}^{(n)}$, $\phi(\underline{\Psi}) : R^N \longrightarrow \aleph_{\mathcal{H}}$, which maps the inputs $\underline{\Psi}$ into high dimensional space.

Now we can find that the transformed input vectors appear in the form of dot product in Eq. (4.12) and Eq. (4.14). So we can define an another function.

$$\kappa(\underline{\Psi}^{(n)},\, \underline{\Psi}^{(m)}) = \phi(\underline{\Psi}^{(n)}).\phi(\underline{\Psi}^{(m)}).$$

where one can avoid explicit handling $\phi(\underline{\Psi})$ in feature space by using kernel function which is independent of its dimension. This approach is also called kernel trick [63]. As long the kernel satisfies Mercer's theorem that is stated above, Eq. (4.12) can be efficiently solved. Some of the well known kernels are shown in Table 4.1.

Kernel Type	Kernel definition
Linear	$\kappa(\underline{\Psi}^{(n)}, \underline{\Psi}^{(m)}) = \underline{\Psi}^{(n)}.\underline{\Psi}^{(m)}$
Gaussian	$\kappa(\underline{\Psi}^{(n)}, \underline{\Psi}^{(m)}) = exp(-\gamma\|\|\underline{\Psi}^{(n)} - \underline{\Psi}^{(m)}\|\|^2)$
Polynomial	$\kappa(\underline{\Psi}^{(n)}, \underline{\Psi}^{(m)}) = (\delta + \gamma.\underline{\Psi}^{(n)}.\underline{\Psi}^{(m)})^d$

Table 4.1: Different kernel formulations.

Once the kernel $\kappa(\underline{\Psi}^{(n)}, \underline{\Psi}^{(m)})$ value is found then the predicted output $\tilde{p}_q^{(m)}$ can be found from Eq. (4.21) as

$$\tilde{p}_q^{(m)} = \sum_{n=1}^{N} (\alpha_n^{opt} - \alpha_n^{*opt}) \kappa(\underline{\Psi}^{(n)}, \underline{\Psi}^{(m)}) + b^{opt} \tag{4.23}$$

This approach is named as $\varepsilon-$based SVM for regression and it is used when the desired accuracy of the estimation is a priory known parameter. SVR generalization performance or the estimation accuracy depends on a good setting of hyper parameters C, ε and the kernel parameters (e.g., γ as for Gaussian kernel). The problem of optimal parameter selection is further complicated by the fact that SVR model complexity (and hence its generalization performance) depends on all three parameters.

- Parameter C determines the trade off between the model complexity (flatness) and the degree to which deviations larger than ϵ are tolerated in optimization formulation. For example, if C is too large (infinity), then the objective is to minimize the empirical risk only, without regard to model complexity part in the optimization formulation.

- Parameter ε controls the width of the ε-insensitive zone, used to fit the training data. The value ε can affect the number of support vectors used to construct the regression function. The bigger, the fewer support vectors are selected. On the other hand, bigger-values results in more flat estimates, hence, both C and ε-values affect model complexity.

- The training model behavior is very sensitive to gamma parameter. Larger γ increases the model tendency to be over fitted. When γ is very small, the model is too constrained and cannot capture the complexity of the data.

Therefore, selection of these parameters should be done by following some guidelines which can provide us optimal and accurate training models. In Sec. 4.4, a couple of strategies are discussed for selecting SVR model parameters for different training strategies.

4.2.2 Inversion through low dimensional extracted features

Applying feature extraction technique (i.e., PLS) for reducing high dimensional ECT feature space to a low dimensional space, requires to update the formulation of SVR.

By applying GRID-PLS, LHS-PLS and PLS-OSF sampling strategies described in Sec. 3.3, ECT features matrix $\underline{\underline{\Psi}} = \left(\underline{\Psi}^{(n)};\, n = 1, ..., N\right)$ of $(N \times F)$ dimension from $\aleph_\Psi$ is transformed to a lower dimensional space, $\aleph_T$ by PLS feature extraction. $\aleph_T$ contains the extracted features, $\underline{\underline{T}} = \left(\underline{T}^{(n)};\, n = 1, ..., N\right)$ of dimension $(N \times J)$ for the associated defect parameters, $\underline{\underline{p}} = (\underline{p}^{(n)}; n = 1, ..., N)$ having $(N \times Q)$ dimension. The q-th training set $(q = 1, ..., Q)$ is built for each defect parameter $D_{N_q} = \left\{\left(\underline{T}^{(n)}; p_q^{(n)}\right);\, n = 1, ..., N\right\}$ to have separate training model for each parameter.

At this stage, an ε-SVR model is trained for each q-th defect parameter $(q = 1, ..., Q)$, by exploiting the corresponding q-th set of I/O pairs on the generated training set (i.e., $D_{N,q} = \left\{\left(\underline{T}^{(n)}; p_q^{(n)}\right); n = 1, ..., N\right\}$) . A m-th test sample $\underline{\Psi}^{(m)}$ $(m = 1, ..., M)$ of F ECT features associated to a previously-unseen crack parameter configuration $\underline{p}^{(m)} = (p_q^{(m)};\, q = 1, ..., Q)$ is projected through PLS wight matrix $\underline{\underline{W}}$ ($\underline{\underline{W}}$ is obtained during training set generation see Sec. 3.3) into the J-dimensional PLS-extracted features space [i.e., $\underline{T}^{(m)} = \underline{\Psi}^{(m)\prime} \times \underline{\underline{W}}$] . $\underline{\Psi}^{(m)\prime}$ is obtained by subtracting each f-th element of $\underline{\Psi}^{(m)}$ $(f = 1, ..., F)$ to its mean value μ_f]. Finally, $\underline{T}^{(m)}$ is given as input to the q-th SVR model in order to estimate the q-th defect parameter, $\tilde{p}_q^{(m)}$ $(q = 1, ..., Q)$.

Training model by Support Vector Regression (SVR)

Similar to previous section, let the input vector $\underline{T}^{(n)}$ of n-th sample of J dimension (i.e., $\underline{T}^{(n)} \in \mathbb{R}^{1\times J}$), and the corresponding real scalar output $p_q^{(n)}$ for the q-th defect parameter. Having the training set $D_{N,q} = \left\{\left(\underline{T}^{(n)}; p_q^{(n)}\right); n = 1, ..., N\right\}$, the $\varepsilon - SVR$ the goal is to find a function $\Theta_q(\underline{T}^{(n)})$ that has maximum deviation ε from the actual output $p_q^{(n)}$ for the input data $\underline{T}^{(n)}$ and as flat as possible. Therefore from Eq. (4.2) we get

$$\Theta_q(\underline{T}^{(n)}) = \underline{\omega} \cdot (\underline{T}^{(n)}) + b,\ \underline{\omega} \in \mathbb{R}^J, b \in \mathbb{R}^N$$

where $\underline{\omega}$, b are the coefficients of the developed model and the corresponding loss function

$$|p_q^{(n)} - \Theta_q(\underline{T}^{(n)})|_\varepsilon = \{0, |p_q^{(n)} - \Theta_q(\underline{T}^{(n)})| - \varepsilon\} \quad n = 1, ..., N$$

By solving the constraint quadratic problem through dual problem and Lagrange multipliers we can update the condition at (4.12) by

$$\max\ L = -\frac{1}{2}\sum_{n,o=1}^{N}(\alpha_n - \alpha_n^*)(\alpha_o - \alpha_o^*)\underline{T}^{(n)} \cdot \underline{T}^{(o)} \tag{4.24}$$

$$-\varepsilon \sum_{n=1}^{N}(\alpha_n + \alpha_n^*) + \sum_{n=1}^{N} p_q^{(n)}(\alpha_n - \alpha_n^*).$$

which is subjected to

$$0 \leq \alpha_n, \alpha_n^* \leq C,$$
$$\sum_{n=1}^{N}(\alpha_n - \alpha_n^*) = 0.$$

where, α_n,α_n^* are Lagrange multipliers. Similar to Eq. (4.13) and Eq. (4.14) we can find

$$\omega^{opt} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\underline{T}^{(n)} \tag{4.25}$$

$$\tilde{p}_q^{(m)} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\underline{T}^{(n)} \cdot \underline{T}^{(m)} + b^{opt} \tag{4.26}$$

where b^{opt} can be calculated with the slack variables ξ_n, ξ_n^* and user defined tuning parameter C ($C \geq 0$)

$$b^{opt} = p_q^{(n)} - \underline{\omega}.\underline{T}^{(n)} - \varepsilon,\ \xi_n = 0,\ \ with\, \alpha_n^{opt} \in (0, C)$$

$$b^{opt} = p_q^{(n)} - \underline{\omega}.\underline{T}^{(n)} + \varepsilon,\ \xi_n^* = 0,\ \ with\, \alpha_n^{*opt} \in (0, C).$$

In case of nonlinear model, Eq. (4.25) and Eq. (4.26) can be expressed by Eq. (4.27) and Eq. (4.28), respectively.

$$\omega^{opt} = \sum_{n=1}^{N}(\alpha_n - \alpha_n^*)\phi(\underline{T}^{(n)}) \tag{4.27}$$

$$\tilde{p}_q^{(m)} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\phi(\underline{T}^{(n)}) \cdot \phi(\underline{T}^{(m)}) + b^{opt} \tag{4.28}$$

where b^{opt} can be represented by

$$b^{opt} = p_q^{(n)} - \underline{\omega}^{opt}.\phi(\underline{T}^{(n)}) - \varepsilon,\ \xi_n = 0,\ \ with\, \alpha_n^{opt} \in (0, C)$$

$$b^{opt} = p_q^{(n)} - \underline{\omega}^{opt}.\phi(\underline{T}^{(n)}) + \varepsilon,\ \xi_n^* = 0,\ \ with\, \alpha_n^{*opt} \in (0, C).$$

Finally, by fulfilling Mercer condition the kernel version of the predicted model shown in Eq. (4.23) can be updated by

$$\tilde{p}_q^{(m)} = \sum_{n=1}^{N}(\alpha_n^{opt} - \alpha_n^{*opt})\kappa(\underline{T}^{(n)}, \underline{T}^{(m)}) + b^{opt}. \tag{4.29}$$

4.3 Definition of error and signal to noise ratio

For assessing the inversion performance of the developed LBE strategies, we are considering the robustness of the LBE approach on noisy data by means of prediction accuracy. In this section, the mathematical formulation of the applied noise and the prediction error metric are described which will be applied through out this research work.

4.3.1 Signal-to-noise ratio (SNR)

In real life experiment, defect investigation suffers for various types of noise such as set up (i.e., cables and circuitry) noise, man made noise etc. All of these noise can corrupt experiment data. Therefore, it is a good practise to keep into account the noise effect in the simulation too. To partially consider noise effects, Additive White Gaussian Noise (AWGN) with a Signal-to-Noise Ratio (SNR) is applied for blurring NDE signals. Depending on the different NDE applications the definition of imposing noise can varied. In this thesis we have mostly addressed ECT signals. However, at Chapter 7 we have described the application of LBE in the context of Ultrasound Testing (UT) and multiphysics approach through data fusion.

Signal-to-noise ratio (SNR) applied on complex valued signals

$$SNR = 10log_{10}\left\{\frac{\Sigma_{k=1}^{K}|\Psi_k|^2}{\Sigma_{k=1}^{K}|\zeta_k|^2}\right\}.$$

where,

- $\Psi_k = (\Re\{\Psi_k\}; \Im\{\Psi_k\})$ is the complex value (containing both real and imaginary terms) of the impedance variation for k-th point ($k = 1, ..., K$).
- $\zeta_k = (\Re\{\zeta_k\}; \Im\{\zeta_k\})$ denotes the complex valued (containing both real and imaginary terms) additive Gaussian noise which can corrupt the k-th ECT signal ($k = 1, ..., K$).

Signal-to-noise ratio (SNR) applied on real valued

$$SNR = 10log_{10}\left\{\frac{\Sigma_{k=1}^{K}|\Psi_k|^2}{\Sigma_{k=1}^{K}|\zeta_k|^2}\right\}.$$

where,

- $\Psi_k = (\Re\{\Psi_k\})$ is the real valued signal at k-th point (e.g., unlike ECT signal, UT reflection signal provides only real value) ($k = 1, ..., K$).
- $\zeta_k = (\Re\{\zeta_k\})$ denotes the real valued additive Gaussian noise which can corrupt the k-th UT signal.

4.3.2 Error metrics

For the performance evaluation of the proposed strategies, the estimation accuracy has been quantitatively evaluated through the following error metrics:

1. Relative error of the q-th parameter

$$\xi(p_q) = \frac{|p_q^{(m)} - \tilde{p}_q^{(m)}|}{|p_q^{(m)}|}.$$

2. Mean square error (MSE) of the q-th parameter

$$MSE(p_q) = \frac{1}{M}\sum_{m=1}^{M} |p_q^{(m)} - \tilde{p}_q^{(m)}|^2.$$

3. Normalized mean error (NME) of the q-th parameter

$$NME(p_q) = \frac{1}{M}\sum_{m=1}^{M} \frac{|p_q^{(m)} - \tilde{p}_q^{(m)}|}{|p_q^{(m)}|}.$$

Here,

- $p_q^{(m)}$ is the actual value of the q-th parameter of m-th sample from M samples
- $\tilde{p}_q^{(m)}$ is the predicted value of the q-th parameter of m-th sample from M samples.
- M is the total number of samples.

4.4 Calibration strategies

We need to select some parameters for developing an inversion model by fitting the training set. In that case we need a model selection (i.e., hyper-parameters selection) process which can validate the trained model. One of the wider used tool for model selection is the so-called cross validation procedure [33]. It consists in removing some data from training set and the learning model is trained on this new training set. Then the removed data can be used as a validation set to evaluate the performance of the trained model. This procedure is repeated until when all the training set data have been used for training and validation purpose.

4.4.1 SVR parameters estimation through cross validation

Cross-validation analysis can be used to find the best combination of parameters to train learning models using SVR technique too. Gaussian Radial Basis Function (RBF) kernel for ε-SVR model has been utilized for this purpose. For simplicity, we have fixed $\varepsilon = 10^{-1}$ for all the analysis. Hence, SVR is characterized by two main parameters: C and γ. C is called as penalty factor and γ represents the exponent in the Gaussian kernel. To find the best (C, γ) pair, classical $\bigvee$-fold cross-validation approach is utilized. In this approach N samples form a training set $D_{N_q} = \left\{ \left(\underline{\Psi}^{(n)}; p_q^{(n)} \right); n = 1, ..., N; q = 1, ..., Q \right\}$ for the q-th parameter $(q = 1, ..., Q)$, which is divided into $\bigvee$ subsets of approximately equal size. After that, each ϑ-th subset ($\vartheta = 1, ..., \bigvee$) is used for validation set and a SVR model is trained by utilizing remaining $\bigvee - 1$ subsets. Finally, the result is used for evaluating prediction accuracy on the ϑ-th subset by defining Mean Squared Error (MSE) as the estimation error.

$$MSE_{\vartheta,\,q}(C,\gamma) = \frac{1}{N_\vartheta} \Sigma_{n=1}^{N_\vartheta} \left\{ p_q^{(n)} - \tilde{p}_q^{(n)}\,(C,\gamma) \right\}^2$$

Here,

- N_ϑ denotes the number of samples inside the ϑ-th subset
- $p_q^{(n)}$ represents actual q-th parameter of n-th sample
- $\tilde{p}_q^{(n)}\,(C,\gamma)$ is the predicted q-th parameter of n-th sample for the given pair (C,γ)

Cross-validation error, MSE for (C,γ) pair is computed as the average MSE obtained for all the $\bigvee$ subsets as

$$MSE_q(C,\gamma) = \frac{1}{\bigvee} \Sigma_{\vartheta=1}^{\bigvee} MSE_{\vartheta,\,q}(C,\gamma)$$

finally the best configuration is calculated from different combination of (C,γ) pairs for the q-th parameter and can be defined as

$$(C_q^{opt}, \gamma_q^{opt}) = \arg \left[\min_{(C,\gamma)} \left\{ MSE_q\,(C,\gamma) \right\} \right].$$

Generally, the optimal $(C_q^{opt}, \gamma_q^{opt})$ pairs have been calculated from the fixed value range of C and γ which are given as

$$C_q = \left[10^{-1},\, 1,\, 10,\, 10^2,\, 10^3,\, 10^4,\, 10^5,\, 10^6 \right]. \tag{4.30}$$

$$\gamma_q = \left[10^{-6},\, 10^{-5},\, 10^{-4},\, 10^{-3},\, 10^{-2},\, 10^{-1},\, 1,\, 10 \right]. \tag{4.31}$$

4.4.2 PLS and SVR parameters estimation through cross validation error

The choice of optimal number of features (i.e., J) to be extracted from actual ECT signal space is a crucial issue. There is no particular criteria available for selecting optimal J value. A common rule of thumb is that the number of extracted features should be kept as low as possible to minimize the complexity of the learning algorithm in training phase. However, it should be large enough to ensure that all the essential information from ECT data is compressed within the extracted features. Within this guideline and for giving a working suggestion, it is worth pointing out that the PLS technique ranks the features from the most to the less important ones according to the corresponding amount of information on the unknown quantity to be predicted [64]. Therefore, an effective and easy-to-implement strategy for choosing the optimal value of J is to build a training set with a maximum number of extracted features J_{max}. Then iteratively removing the higher-order (i.e., less informative) ones such that an optimum error is found. We have utilized cross validation error MSE provided by SVR for this optimization. Therefore, the optimal triplet of (J, C, γ) is obtained by performing cross validation error on the training set of N samples performed by SVR. The steps are given below:

1. A training set $D_{N_q} = \left\{ \left(\underline{T}^{(n)}; p_q^{(n)} \right) ; n = 1, ..., N; q = 1, ..., Q \right\}$ is defined, where $\underline{T}^{(n)} \in \aleph_\Psi \subset \mathbb{R}^{N \times J}$ of J number of extracted features ($J = 1, ..., J_{max}$) for the q-th parameter. Sub-optimal C and γ parameters ($C_{J,q}^{sub}$, $\gamma_{J,q}^{sub}$) are obtained by performing $\bigvee$-fold cross-validation approach with $\bigvee = 5$ through cross-validation error (i.e., MSE) on the D_{N_q} training set. The fixed value ranges of C and γ are applied as mentioned in Eq. (4.30) and Eq. (4.31), respectively.

2. Optimal J value (J_q^{opt}) for each q-th parameter has been chosen for the lowest cross-validation error, MSE_q value from all triplet (J_q, $C_{J,q}^{sub}$, $\gamma_{J,q}^{sub}$) combination, where $J = 1, ..., J_{max}$. That can be defined as

$$J_q^{opt} = \arg \left\{ \min_{(J, C_{J,q}^{sub}, \gamma_{J,q}^{sub})} MSE_q \right\} \quad q = 1, ..., Q; \; J = 1, ..., J_{max}$$

3. The mean value of J_q^{opt} has been chosen for the optimal J^{opt}.

$$J^{opt} = \frac{1}{Q} \sum_{q=1}^{Q} J_q^{opt}.$$

4. Finally, optimal ($C_q^{opt} = C_{J^{opt},q}^{sub}$, $\gamma_q^{opt} = \gamma_{J^{opt},q}^{sub}$) pairs are obtained for J^{opt} number of optimal features.

4.4.3 PLS and SVR parameters estimation based on cross validation and inversion error

In this strategy, we are also considering the performance evaluation through inversion error for choosing PLS and SVR parameters. Hence, the optimal triplet of (J, C, γ) is obtained by performing cross validation error on the training set of N samples performed by SVR and inversion error on noisy as well as noiseless test set by the SVR trained model. The steps are given below:

1. For a training set $D_{N_q} = \left\{ \left(\underline{T}^{(n)}; p_q^{(n)} \right) ;\, n = 1, ..., N;\, q = 1, ..., Q \right\}$, $\underline{T}^{(n)} \in \aleph_\Psi \subset \mathbb{R}^{1 \times J}$ consider J number of extracted features ($J = 1, ..., J_{max}$) for the q-th parameter. Sub-optimal C and γ parameters ($C_{J,q}^{sub}$, $\gamma_{J,q}^{sub}$) are obtained by performing $\bigvee$-fold cross-validation approach with $\bigvee = 5$ through cross-validation error (i.e., MSE) on the D_{N_q}training set. Similar to previous cases, C and γ values are applied from the ranges mentioned in Eq. (4.30) and Eq. (4.31), respectively.

2. Build inversion models for the q-th parameter by applying SVR using ($C_{J,q}^{sub}$, $\gamma_{J,q}^{sub}$) pairs.

3. Predict the q-th parameter through inversion by the trained model on $Noiseless$ as well as noisy test sets (i.e., $SNR = 20,\ 30,\ 40$ [dB]).

4. Sub optimal J value ($J_{q,\,SNR}^{sub}$) for each q-th parameter has been chosen for the lowest prediction error, $NME_{q,SNR}$ among all the test set of different SNR from the corresponding triplet (J_q, $C_{J,q}^{sub}$, $\gamma_{J,q}^{sub}$) combinations, where $J = 1, ..., J_{max}$; $q = 1, ..., Q$ and $SNR = 20,\ 30,\ 40,\ \infty$ [dB] . That can be defined as

$$J_{q,\,SNR}^{sub} = \arg \left\{ \min_{(J, C_{J,q}^{sub}, \gamma_{J,q}^{sub})} NME_{q,SNR} \right\}$$

5. Optimal J^{opt} has been obtained by the mean value of $J_{q,\,SNR}^{sub}$.

$$J^{opt} = \left\{ \frac{1}{4Q} \sum_{SNR} \sum_{q=1}^{Q} J_{q,\,SNR}^{sub} \right\}.$$

6. Finally, optimal ($C_q^{opt} = C_{J^{opt},q}^{sub}$, $\gamma_q^{opt} = \gamma_{J^{opt},q}^{sub}$) pairs are obtained for J^{opt} number of optimal features.

In the next chapters, LBE strategies will be applied for inverse problem solution by addressing different NDT problems. Chapter 5 and Chapter 6 will be devoted for performing inversion within the context of ECT signals. An innovative multiphysics data fusion approach by combing ECT and ultrasound testing signals has been shown in Chapter 7 for solving a complex NDT problem.

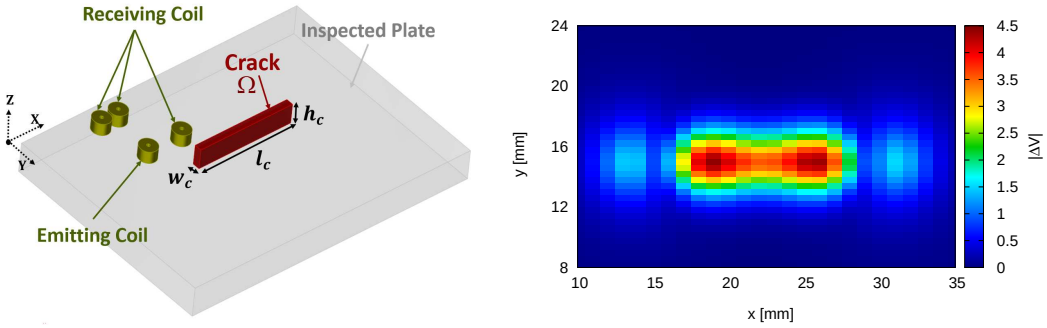
Chapter 5

Crack(s) characterization in conductive plate(s)

This chapter describes the real-time crack(s) characterization problem by exploiting Learning by Examples (LBE) strategy in the context of Eddy Current Testing (ECT). An optimal training set has been generated in offline phase by adopting Partial Least Squares (PLS) feature extraction combined with a customized version of output space filling (OSF) within the framework of LBE. Support Vector Regression (SVR) algorithm is utilized for developing an accurate model based on the training set and subsequently real-time inversion (online phase) has been performed on unknown test set. The robustness of the refereed hereafter as PLS-OSF/SVR approach is numerically assessed in presence of synthetic noisy test set and on experimental data.

5.1 Problem definitions

Let us consider a conductive plate of thickness 1.27 mm, having conductivity 1.02 MS/m, relative permeability 1 and permeability 1. The plate is affected by a single rectangular subsurface crack (e.g, the crack is attached with the bottom side of the plate) having geometrical region Ω within the structure under test (SUT) (i.e., Fig. 5.1). The crack is characterized by total $Q = 3$ descriptors $\underline{p} = (l_c,\, h_c,\, w_c)$, where l_c, h_c, w_c represent the crack length, height and width, respectively. The plate is inspected by 4 coils in absolute mode working at a frequency of 200 kHz. For this test configuration one coil is active as emitting and a second one as receiving (e.g., the emitting receiving mode). More details on coil parameters can be found in [65]. The receiving coil collects the ECT signals from 28 and 23 positions with a step size of 1 mm along X and with a step size of 0.4 mm along Y directions, respectively through a raster scan from a constant lift-off height 1.00 mm. Therefore, ECT signals (i.e., voltage variation signal ΔV) are collected from $K = 28 \times 23 = 644$ number of inspected points.



(a) (b)

Figure 5.1: Examples of (a) studied plate geometry and (b) ECT signal map in terms of absolute value of coil voltage variation.

5.2 Training and test set configuration

This section describes the numerical and experimental configuration of the treated problem for performing inversion. Based on the ECT model described in [66, 67], ECT signals have been built by utilizing CIVA [15] software.

Parameters	Min	Max
Length l_c [mm]	4.00	13.00
Height h_c [mm]	0.189	0.931
Width w_c [mm]	0.05	0.4

Table 5.1: Cracks parameters ranges.

5.2.1 Numerical set up

In this case, three training sets for three different trained models have been created by changing the crack dimensions within the range mentioned in Table 5.1 by PLS-OSF sampling approach. A set of N samples has been built by utilizing PLS-OSF sampling (see Sec. 3.3.5) method within the selected ranges. More in detail, starting from $N_0 = 27$ initial number of samples we added training samples adaptively until $N = 512$, which represents the maximum number of samples.

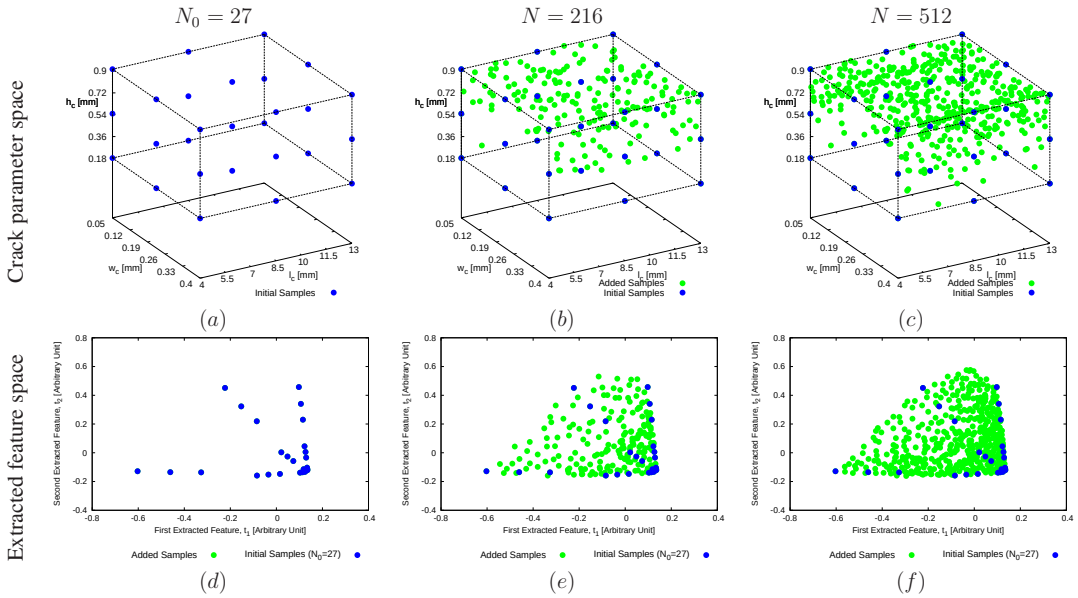


Figure 5.2: Numerical assessment: (Length $l_c = [4,\ 13]$ mm, height $h_c = [0.189,\ 0.931]$ mm and width $w_c = [0.05, 0.4]$ mm) - Disposition of the training samples generated by PLS-OSF in (a)-(c) the crack parameter space and (d)-(f) the extracted feature space for $(a)(d)$ $N_0 = 27$ $(b)(e)$ $N = 216$ and $(c)(f)$ $N = 512$.

The optimum PLS parameter (i.e., J) and SVR hyper-parameters (e.g., values of C and γ parameters) are chosen by adopting the calibration strategy mentioned in Sec. 4.4.2. The optimal (C, γ) pairs are obtained through the calibration step on $J \in [2,\ 20]$ and performing cross validation by SVR on the training set having $N = 512$ for each parameter of the crack. Through performing the calibration, $J^{opt} = 5$ number of features are extracted from $F = 2 \times K = 1288$ number of ECT features. Figure 5.2 represents the exploration of adaptive training samples generation in the parameter space as well as in extracted feature space. In Fig. 5.2 $(d) - (f)$ feature spaces are represented for the first two extracted features of training set having $J = 5$ extracted features for imaging purpose.

Similar to PLS-OSF sampling, three training sets for three different trained models have been built by applying GRID sampling approach through changing the crack dimensions within the range mentioned in Table 5.1. The details of the GRID sampling strategy can also be retrieved from Sec. 3.2.1. The optimum values of (C, γ) pairs of SVR are tuned by exploiting cross-validation strategy defined in Sec. 4.4.1 for GRID sampling and during the training phase on the different training sets. In all cases, for applying ε-SVR, ε is kept fixed at 10^{-1}. Table 5.2 shows the optimal SVR calibration parameters obtained for both GRID and PLS-OSF sampling approaches. Aftermath, different training models (offline phase) are obtained for q-th crack parameter ($q = 1, ..., Q$) separately for different sampling strategies (i.e., PLS-OSF and GRID sampling) by the corresponding SVR parameters.

	PLS-OSF			GRID		
Parameters	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE
l_c	3.00	0	2.17×10^{-02}	4.00	-2.00	4.54×10^{-02}
h_c	0	-1.00	3.64×10^{-03}	5.00	-6.00	2.85×10^{-03}
w_c	3.00	0	5.39×10^{-03}	5.00	-4.00	3.43×10^{-03}

Table 5.2: Set of optimal calibration parameters found for PLS-OSF sampling for the training sets at $N = 512$, $F = 1288$ and $J^{opt} = 5$.

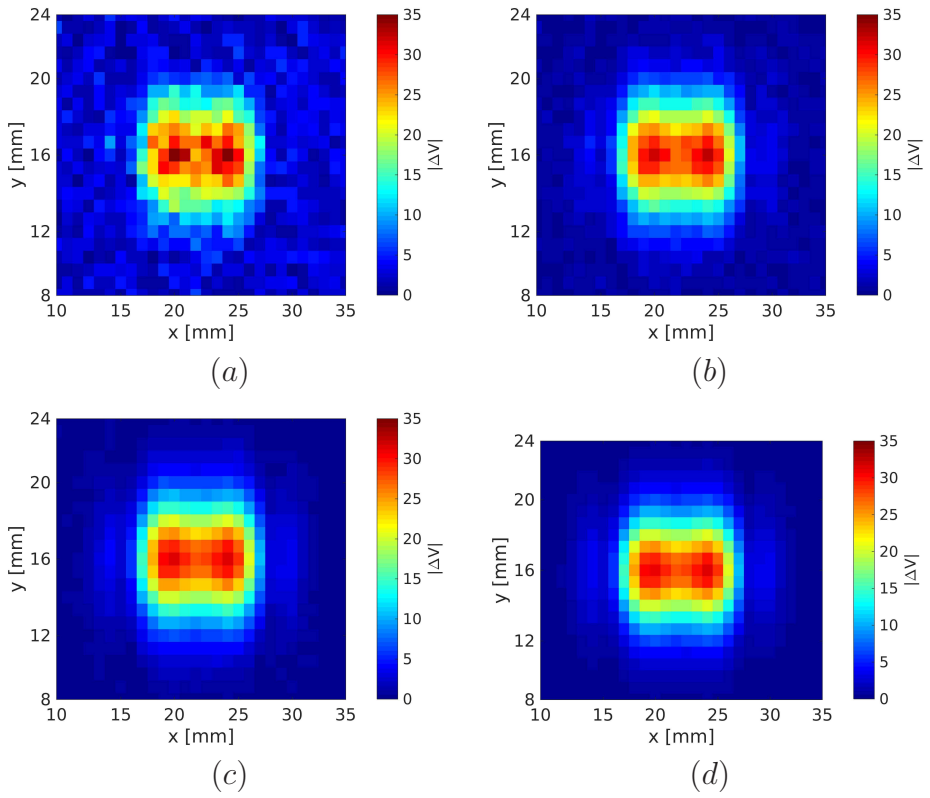


Figure 5.3: Numerical assessment: Examples of absolute value of coil voltage variation corrupted by AWGN through (a) $SNR = 10$ [dB], (b) $SNR = 20$ [dB], (c) $SNR = 30$ [dB] and (d) $SNR = 40$ [dB] for a crack having length l_c = 11.7 mm, $h_c = 0.895$ mm, $w_c = 0.364$ mm.

Training set parameters	
Number of ECT features, $F = 2K$	1288
Number of extracted features, J	5
Sampling strategies	PLS-OSF, GRID
Number of initial samples, N_0	27
Maximum number of samples, N	512
Test set parameters	
Number of ECT features, F	1288
Number of extracted features, J	5
Sampling strategies	LHS
Number samples, M	1000
Signal to noise ratio, SNR	10, 20, 30, 40[dB]

Table 5.3: Training and test set parameters.

An unknown test set of $M = 1000$ samples for 3 crack parameters has been generated by using LHS design (see Sec. 3.2.2). Like training set, the same number of ECT features (i.e., $F = 2K = 1288$) are considered for each test sample for treating trained models obtained by GRID/SVR. To partially consider noise effects, Additive White Gaussian Noise (AWGN) with a Signal-to-Noise Ratio (SNR) is applied for blurring ECT. The impact of noise on ECT signal map for a given crack configuration (arbitrarily chosen from the Test set) on different SNR are shown Fig. 5.3. For applying PLS-OSF sampling, $J = 5$ features are extracted by projecting the test set into extracted feature space through the PLS weight matrix (e.g., the one obtained from training). Table 5.3 summarizes the training and test set parameters that are used for this analysis. Finally, the estimation of crack parameters on the unknown test set (during online phase) is evaluated by the trained models.

5.2.2 Experimental set up

A homogeneous conductive plate having the same properties described Sec. 5.1 is affected by a sub surface crack. Four measurements have been conducted for four different crack configurations for building the experimental test set. Fig. 5.4 shows the corresponding absolute ECT signal maps of the four experimental configurations. The actual dimensions of these four cracks are indicated in Tab. 5.4. All other inspection parameters remain unchanged. ECT response signals are collected from $K = 644$ measurement points which are represented $F = 2K = 1288$ ECT features for treating with GRID/SVR model. In case of PLS-OSF/SVR, $F = 1288$ ECT features are projected to the PLS feature space and $J = 5$ extracted features are retrieved through the PLS weight matrix (obtained during training phase). Training models obtained by PLS-OSF/SVR and GRID/SVR approaches (described in Sec. 5.2.1) have been utilized for predicting the actual parameters of the experimental test set. The predic-

tion accuracy will be assessed by computing the relative prediction error, ξ for each parameter of a single crack.

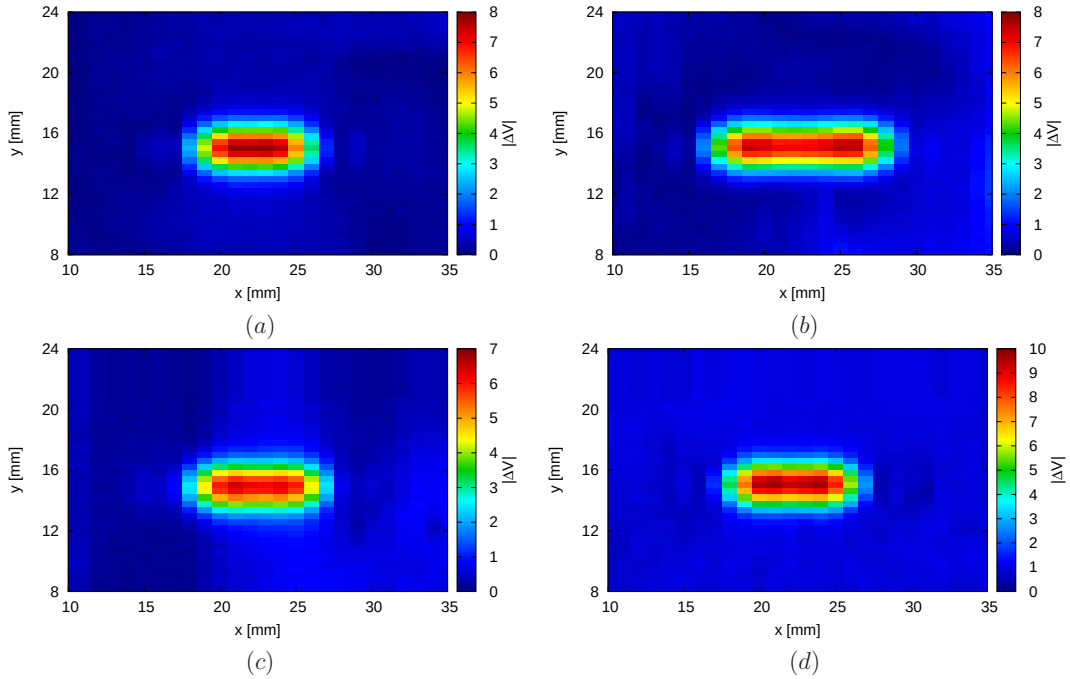


Figure 5.4: Experimental assessment: ECT signal map representation at (a) crack 1 (l_c = 8 mm, $h_c = 0.508$ mm, $w_c = 0.2$ mm), (b) crack 2 (l_c = 12 mm, $h_c = 0.508$ mm, $w_c = 0.2$ mm), (c) crack 3 (l_c = 9 mm, $h_c = 0.508$ mm, $w_c = 0.1$ mm) and (d) crack 4 (l_c = 9 mm, $h_c = 0.508$ mm, $w_c = 0.4$ mm).

Crack No.	Length l_c [mm]	Height h_c [mm]	Width w_c [mm]
1	8.00	0.508	0.20
2	12.00	0.508	0.20
3	9.00	0.508	0.10
4	9.00	0.508	0.40

Table 5.4: Experimental assessment: crack dimension in experimental test set.

5.3 Results and discussion

In this Section, the predictive performance of SVR by means of inversion is elaborated by applying both GRID and PLS-OSF sampling approaches. Numerical and experimental validation of different learning algorithms by using GRID sampling strategy were described in [25]. In this analysis the inversion accuracy of SVR using PLS-OSF sampling method is numerically and experimentally assessed and compared with GRID sampling approach. Dealing with higher dimensional problem (e.g., $F = 1288$ ECT features), GRID sampling requires as to have higher number of training samples (i.e., N) to get enough ECT information that can build accurate model [25]. This phenomenon is known as curse of dimensionality. This also makes it difficult for SVR model to deal with noisy data for lower number of training samples N.

PLS-OSF/SVR shows improved prediction accuracy for lower N values, which is also robust in case of noisy test set. By applying PLS feature extraction, training and test sets can be represented in lower dimension ($J \ll F$) with most significant features. This enables the prediction technique (e.g., SVR) to be more accurate in training model creation and inversion. By uniformly exploring the feature space and the associated crack parameter space can lead PLS-OSF sampling strategy to have optimal (i.e., complete and non-redundant) representation of the extracted feature space. This results lower prediction error on crack dimension estimation for lower N as well as in case of noisy test data. Fig. 5.5 - Fig. 5.7 summarize the robustness of inversion accuracy in presence of noisy test set by utilizing PLS-OSF/SVR approach and compare with the performance of GRID/SVR approach. We can observe that with increasing N, prediction error, NME decreases slightly for higher SNR by using GRID/SVR for l_c , h_c and w_c estimation. That indicates that dealing a problem of $F = 1288$ dimension, number of applied training samples $N = 512$ is not enough for accurate training model development by GRID/SVR. Therefore, with decreasing SNR, GRID/SVR method fails to predict any crack parameters. Whereas, PLS-OSF/SVR approach deals with only $J = 5$ extracted features which required lower number of training samples to have accurate training model. As a result, the prediction accuracy of PLS-OSF/SVR approach increases highly while N is increasing through adaptive sampling.

Though PLS-OSF can provide optimal training set, prediction accuracy of SVR is different for different crack parameters. We can observe this behavior from Fig. 5.5, ECT signal strongly changes with the variation of crack height h_c, hence higher prediction accuracy is obtained for lower N. Contrary, l_c estimation accuracy can be significantly improved by adaptive increasing N. Due to the narrower crack gap (i.e., crack width w_c) ranges with respect to the probe diameter, ECT measurement for the variation of crack width is suffered for spatial resolution problem. Moreover, eddy current does not significantly change due to the variation of crack width. This impacts on ECT measurement and consequently w_c estimation (Fig. 5.6) that leads to higher NME (Fig. 5.7).

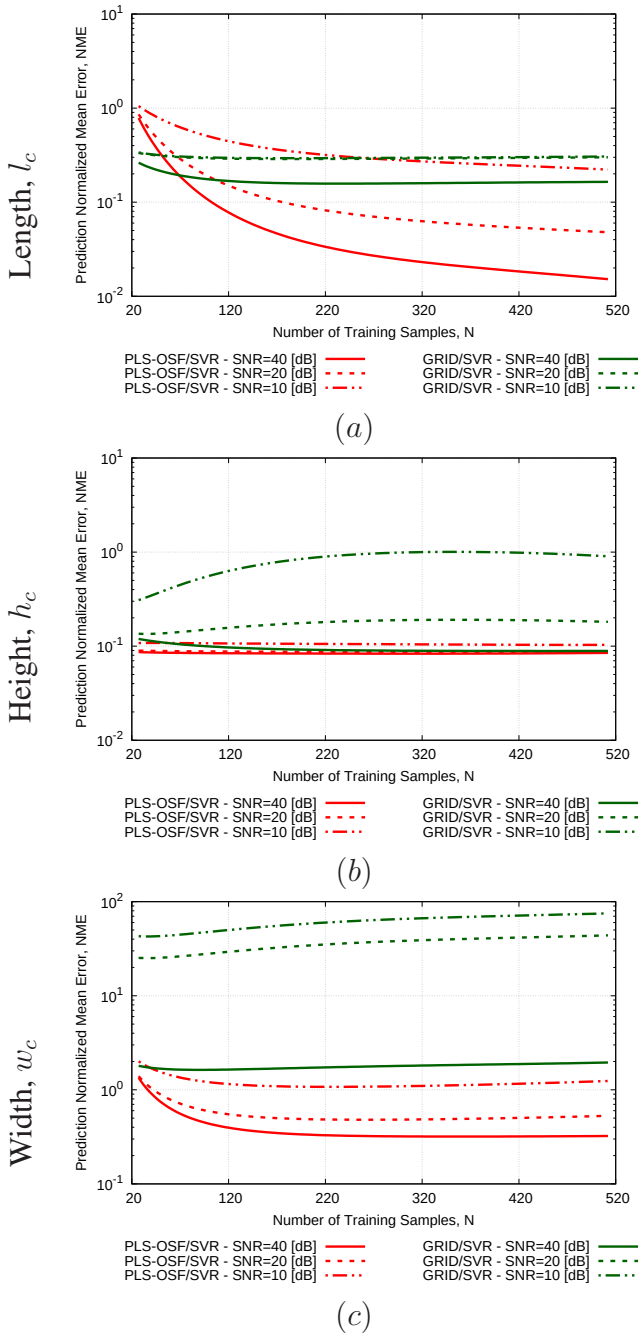


Figure 5.5: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparisons in terms of N vs. NME are shown for crack (a) length l_c, (b) height h_c and (c) width w_c estimation for test set at $SNR = 10,\ 20,\ 40$ [dB] while $F = 1288,\ J = 5,\ M = 1000$.

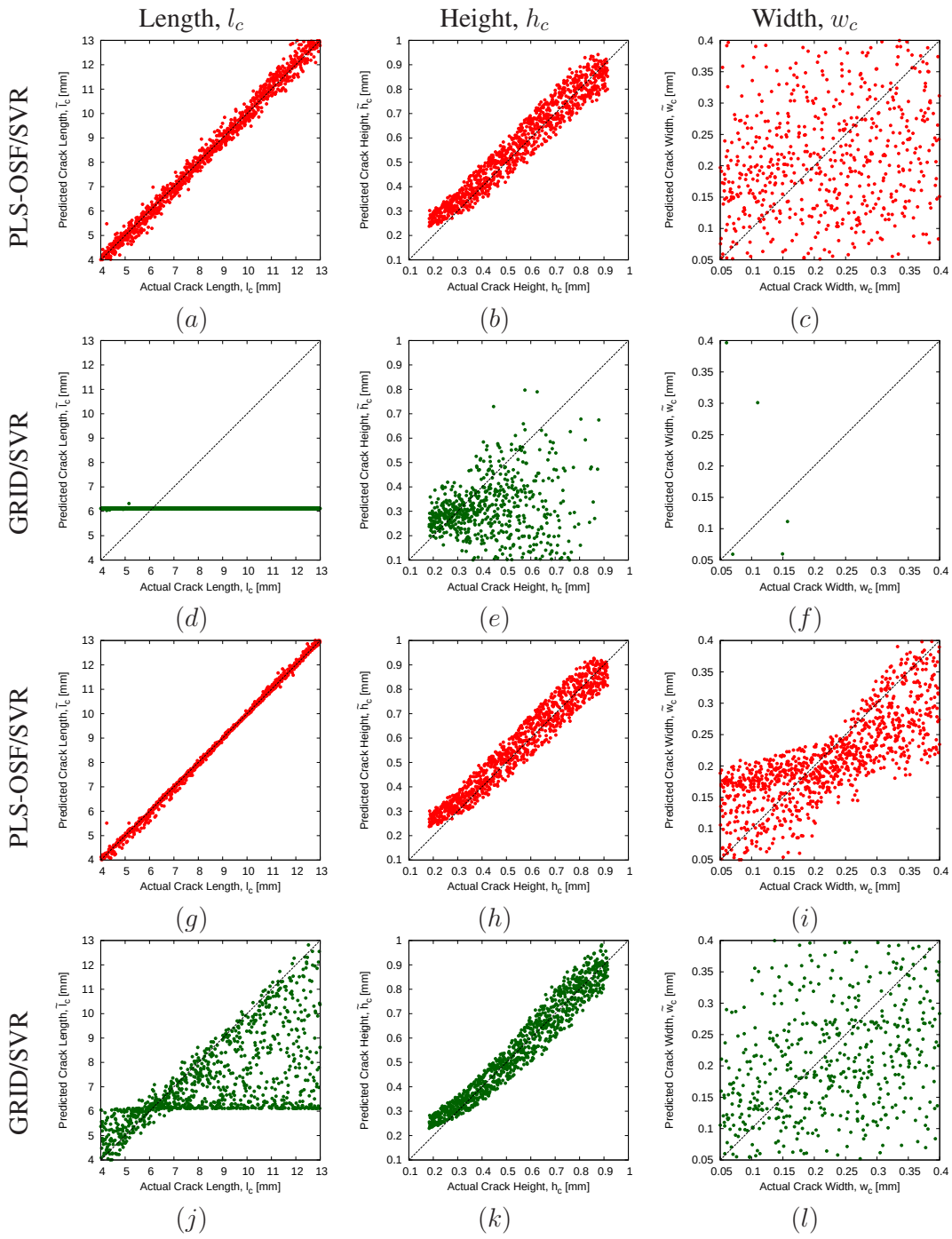


Figure 5.6: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparison are shown in terms of true vs. predicted (crack length l_c, height h_c, width w_c) regression plots on the test set at $(a)-(f)$ $SNR = 10$ [dB] and at $(g)-(l)$ $SNR = 40$ [dB], while $N = 512,\ F = 1288,\ J = 5,\ M = 1000$.

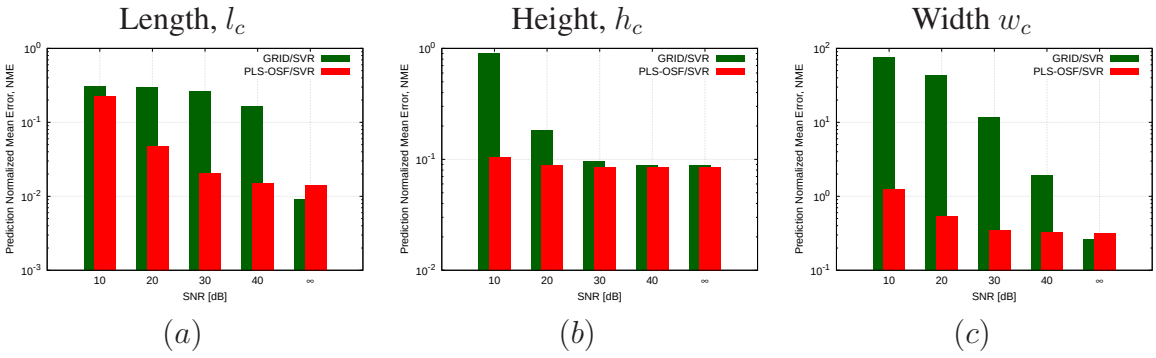


Figure 5.7: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparisons in terms of SNR vs. NME are shown crack (a) length l_c, (b) height h_c and (c) width w_c estimation for test set at noiseless and $SNR = 10,\ 20,\ 30,\ 40$ [dB] while $N = 512,\ F = 1288,\ J = 5,\ M = 1000$.

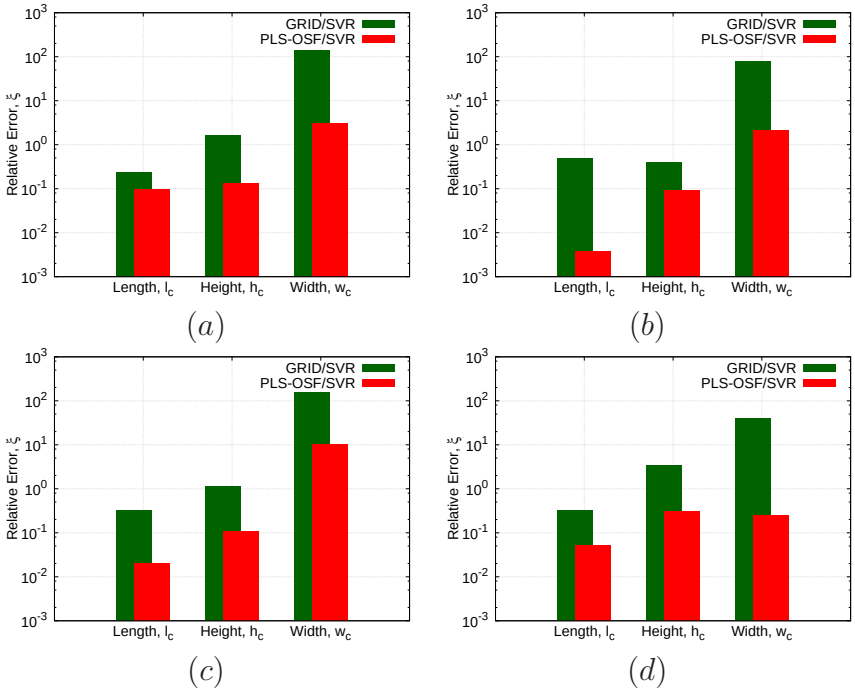


Figure 5.8: Experimental assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparisons in terms relative error, ξ on crack parameters estimation of (a) crack no. 1 (l_c = 8 mm, $h_c = 0.508$ mm, $w_c = 0.2$ mm), (b) crack no. 2 (l_c = 12 mm, $h_c = 0.508$ mm, $w_c = 0.2$ mm), (c) crack no. 3 (l_c = 9 mm, $h_c = 0.508$ mm, $w_c = 0.1$ mm) and (d) crack no. 4 (l_c = 9 mm, $h_c = 0.508$ mm, $w_c = 0.4$ mm) for $N = 512$, $F = 1288$ and $J = 5$.

Crack No.	Crack 1			Crack 2			Crack 3			Crack 4		
Parameters	Actual	Predicted	$\xi(\%)$	Actual	Predicted	$\xi(\%)$	Actual	Predicted	$\xi(\%)$	Actual	Predicted	$\xi(\%)$
Length l_c [mm]	8.00	8.79	9.87	12.00	12.04	**0.33**	9.00	8.82	2.00	9.00	9.13	1.44
Height h_c [mm]	0.508	0.574	12.99	0.508	0.544	**7.08**	0.508	0.562	10.62	0.508	0.665	30.90
Width w_c [mm]	0.20	0.80	300	0.20	0.63	215	0.10	1.1	1000	0.40	0.50	**25.00**

Table 5.5: Experimental assessment: performance of PLS-OSF/SVR for estimating the crack parameters (l_c, h_c and w_c) in terms of relative error, ξ for $N = 512$, $F = 1288$ and $J = 5$.

Crack No.	Crack 1			Crack 2			Crack 3			Crack 4		
Parameters	Actual	Predicted	$\xi(\%)$	Actual	Predicted	$\xi(\%)$	Actual	Predicted	$\xi(\%)$	Actual	Predicted	$\xi(\%)$
Length l_c [mm]	8.00	6.11	23.62	12.00	6.11	49.08	9.00	6.11	32.11	9.00	6.11	32.11
Height h_c [mm]	0.508	1.32	160	0.508	0.303	40.35	0.508	1.10	116.53	0.508	2.20	333
Width w_c [mm]	0.20	-28.4	14309	0.20	-1.57	7985	0.10	-1.57	15871	0.40	-1.57	4042

Table 5.6: Experimental assessment: performance of GRID/SVR for estimating the crack parameters (l_c, h_c and w_c) in terms of relative error, ξ for $N = 512$, $F = 1288$.

Figure 5.8 depicts the experimental assessment of crack parameters (l_c, h_c and w_c) estimation through SVR techniques using PLS-OSF and GRID sampling strategies by means of relative error, ξ_q $(q = 1, ..., Q)$. PLS-OSF/SVR shows highest prediction accuracy for l_c estimation with lowest relative error ($\xi_{l_c} \approx 0.33\%$) and highest relative error ($\xi_{l_c} \approx 9.87\%$) for crack no. 2 and crack no. 1, respectively (Tab. 5.6). As crack no. 2 and crack no. 1 is the largest and smallest cracks, respectively within all the four cracks, these results are somehow expected. Similarly, having same length (i.e., $l_c = 9$), in case of crack no. 3 and crack no. 4 prediction accuracy by PLS-OSF/SVR shows similar prediction accuracy for l_c estimation. h_c estimation accuracy (ξ_{h_c}) for all the cracks ranges from $\sim 7\%$ to 30% by PLS-OSF/SVR. Among all the cracks, w_c estimation of the crack no. 4 (the widest crack having width, $w_c = 4.00$ mm), shows lowest relative error ($\xi_{w_c} \approx 25\%$) for applying PLS-OSF/SVR strategy. It is also visible that GRID/SVR has higher inversion error (Tab. 5.6) compare to PLS-OSF/SVR for estimating all the crack parameters among all the cracks.

5.4 Conclusion

In this research, we have assessed the performance of PLS-OSF/SVR strategy within the framework of LBE for real-time flaw characterization based on ECT signal. PLS-OSF/SVR and GRID/SVR have been illustrated for different number of training samples, N on noiseless as well as on noisy test set. The robustness of PLS-OSF/SVR on noisy test set has been analyzed in terms of prediction accuracy and compared with more standard GRID/SVR approach (Fig. 5.5 - Fig. 5.7). The prediction accuracy for crack dimension estimation by applying PLS-OSF/SVR has also been compared with GRID/SVR on experimental data. From the computational point of view, the

average prediction time was only 0.03 s for predicting $M = 1000$ test samples during online phase by PLS-OSF/SVR on standard laptop. This also demonstrates that inversion is quasi-real time. By applying PLS-OSF sampling strategy, an optimal training set can be obtained which implies higher prediction accuracy in presence of noisy data. Moreover, PLS-OSF sampling strategy increases the learning capability of employed algorithm (i.e., SVR). This strategy also yields higher prediction accuracy for crack dimension estimation on real experimental data. Due to the generalization capability of SVR, PLS-OSF/SVR also exhibits higher prediction accuracy during experimental assessment (Fig. 5.8). Next Chapter will be devoted to the analysis of PLS-OSF/SVR performance on more complex test cases involving higher dimensional parameter space.

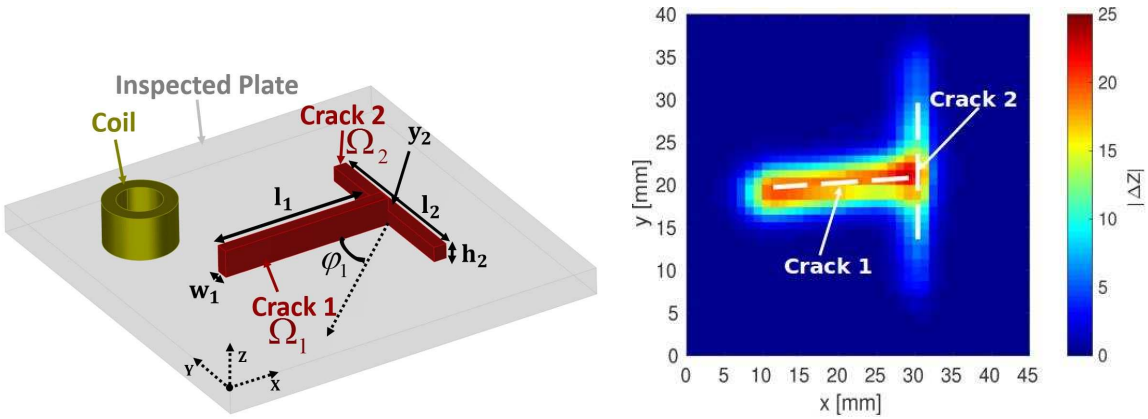
Chapter 6

Multiple crack characterization in conductive plate

This chapter describes the real-time crack characterization and localization inside a Structure Under Test (SUT) by exploiting Learning by Examples (LBE) strategy in the context of Eddy Current Testing (ECT). Within the framework of LBE, an optimal training set has been generated in offline phase by adopting Partial Least Squares (PLS) feature extraction combined with a customized version of output space filling (OSF). The applicability of PLS-OSF sampling approach combining with Support Vector Regression (SVR) has been utilized for multi-crack characterization and localization problem having high dimensional parametric space. The robustness of PLS-OSF/SVR is numerically assessed in presence of synthetic noisy test data and compared with standard approach (e.g., GRID/SVR).

6.1 Problem definitions

A homogeneous plate of thickness 1.27 mm, and conductivity 1.02 MS/m , made by Inconel alloy with relative permeability 1 has been investigated. The plate has been affected by two cracks (Fig. 6.1). Crack 1 is a surface breaking crack, Ω_1 of 1.27 mm height, while crack 2, Ω_2 is a subsurface crack having fixed width 0.19 mm. Ω_1 and Ω_2 are characterized by total $Q = 6$ descriptors $\underline{p} = (l_1,\ w_1,\ \varphi_1,\ l_2,\ h_2,\ y_2)$. l_1, w_1, φ_1 represent the length, width and angular displacement of crack 1 respectively and crack 2 is characterized by length l_2 , height h_2 and axial position y_2 along y axis. The plate is inspected by a single coil working in absolute mode of frequency 500 kHz with lift off 0.27 mm. More details on coil parameters can be found in [68]. The coil collects the ECT signals from 46 and 41 positions along X and Y directions, respectively with a step size of 0.5 mm through a raster scan. Therefore, ECT signals (i.e., impedance variation signal) are collected from $K = 46 \times 41 = 1886$ number of inspected points.



(a) (b)

Figure 6.1: Examples of (a) studied plate geometry and (b) ECT signal map in terms of absolute value of coil impedance variation.

Parameters	Min	Max
Crack 1		
Length l_1 [mm]	7.00	13.00
Angular position φ_1 [deg]	-15	15
Width w_1 [mm]	0.01	0.35
Crack 2		
Length l_2[mm]	7.00	13.00
Height h_2[mm]	0.245	1.016
Axial position y_2[mm]	24	31

Table 6.1: Cracks parameters ranges.

6.2 Training and test set configuration

This section describes the numerical configuration of the treated problem for performing inversion. ECT signals have been generated by CIVA [15] based on the model described in [66, 67]. Six training sets for six different trained models have been created by changing the parameters of both cracks within the parameter ranges mentioned in Tab. 6.1. Different training sets have been built by PLS-OSF and GRID sampling methods and the corresponding ECT signals (i.e., impedance variation signal) are collected from $K = 46 \times 41 = 1886$ number of inspected points.

The initial number of samples $N_0 = 729$ and maximum number of $N = 4096$ are chosen for the appropriate performance comparison between PLS-OSF/SVR and GRID/SVR (i.e., complete $3^6 = 729$ and $4^6 = 4096$ factorial design) strategies. Each

of these samples is associated to $F = 2K = 3772$ ECT features, which have been reduced to J extracted features by PLS feature extraction for PLS-OSF/SVR strategy (see Sec. 3.3.5).

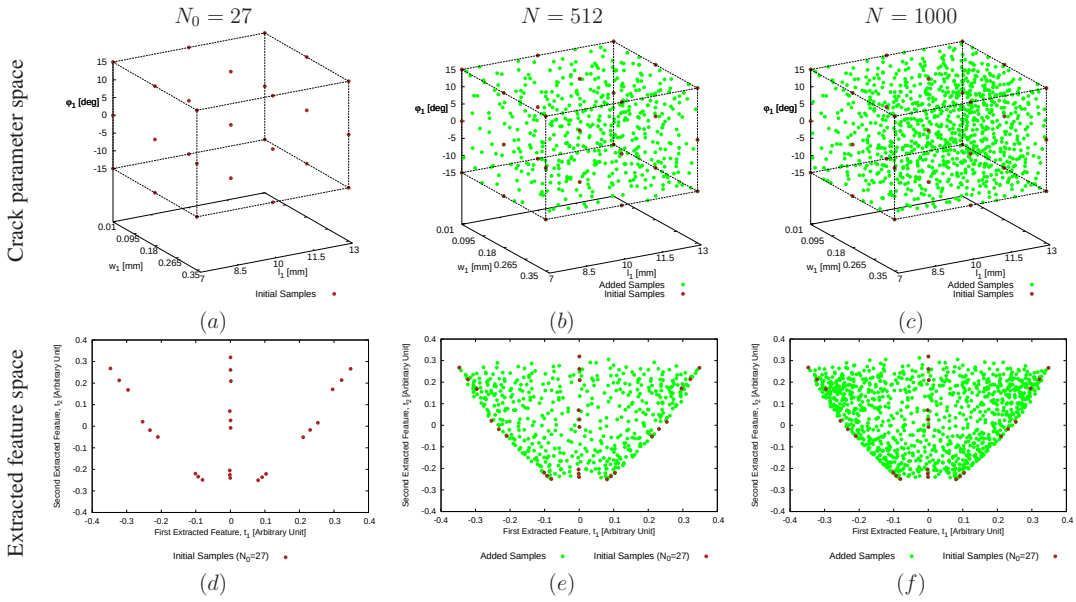


Figure 6.2: Numerical assessment: (Length $l_1 = [7,\ 13]$ mm, width $w_1 = [0.01,\ 0.35]$ mm, angular position $\varphi_1 = [-15,\ 15]$ deg.; length $l_2 = 10$ mm, $h_2 = 0.7$ mm and $y_2 = 27$ mm) - Disposition of the training samples generated by PLS-OSF in (a)-(c) the crack parameter space and (d)-(f) the extracted feature space for $(a)(d)$ $N_0 = 27$, $(b)(e)$ $N = 512$ and $(c)(f)$ $N = 1000$.

	PLS-OSF			GRID		
Parameters	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE
l_1	5.00	-4.00	2.29×10^{-03}	2.00	-4.00	5.84×10^{-03}
w_1	5.00	-6.00	2.50×10^{-03}	2.00	-6.00	4.71×10^{-03}
φ_1	5.00	-2.00	6.81×10^{-03}	2.00	-3.00	1.13×10^{-01}
l_2	4.00	-1.00	7.05×10^{-02}	1.00	-3.00	1.44
h_2	5.00	-4.00	2.14×10^{-03}	1.00	-3.00	7.15×10^{-03}
y_2	3.00	-1.00	1.32×10^{-02}	2.00	-3.00	2.98×10^{-01}

Table 6.2: Set of optimal calibration parameters found for both PLS-OSF and GRID sampling approaches for the training sets at $N = 4096$, $F = 3772$, and $J^{opt} = 15$.

Figure 6.2 represents the exploration of adaptive training samples generation in the parameter space as well as in the extracted feature space. For visualization purpose, the training set of three dimensional parameter space with the corresponding two dimensional extracted features space are shown by varying the parameters of crack 1 and fixing crack 2 parameters.

Optimum PLS parameter (i.e., J) and SVR hyper parameters (e.g., values of C and γ parameters) are chosen by adopting the calibration strategy mentioned in Sec. 4.4.2. The optimal (C, γ) pairs are obtained through the calibration step on $J \in [2,\ 20]$ and performing cross validation by SVR on the training set having $N = 4096$ for each parameters of the cracks. By performing the calibration, optimal number of extracted features is found for $J = 15$. Hence, 15 features are extracted from F number of ECT features.

The details of the GRID sampling strategy can be retrieved from Sec. 3.2.1. The optimum values of (C, γ) pairs of SVR are tuned by exploiting cross-validation strategy defined in Sec. 4.4.1 for GRID sampling and during the training phase on the different training sets. In all cases, for applying ε-SVR, ε is kept fixed at 10^{-1}. In Tab. 6.2, the obtained optimal values of (C, γ) pairs are shown for both PLS-OSF and GRID sampling strategies for all the cracks parameters. After which, different training models (offline phase) are obtained for q-th crack parameter ($q = 1, ..., Q$) separately by different sampling strategies with the corresponding SVR parameters.

Training set parameters	
Number of ECT features, $F = 2K$	3772
Number of extracted features, J	15
Sampling strategies	PLS-OSF, GRID
Number of initial samples, N_0	729
Maximum number of samples, N	4096
Test set parameters	
Number of ECT features, F	3772
Number of extracted features, J	15
Sampling strategies	LHS
Number samples, M	1000
Signal to noise ratio, SNR	10, 20, 30, 40[dB]

Table 6.3: Training and test set parameters.

An unknown test set of $M = 1000$ samples for 6 crack parameters has been generated by using LHS design. Similar to the training sets, the same number of ECT features (i.e., $F = 2K = 3772$) is also considered for each test sample for treating training model obtained by GRID/SVR. Whereas, for PLS-OSF/SVR trained models, $J = 15$ features are extracted by projecting the test set into extracted feature space through the PLS weight matrix (e.g., the one obtained from training). As mentioned in previous chapter, Additive White Gaussian Noise (AWGN) with a signal-to-noise ratio (SNR) is applied for blurring ECT signals. The noise effects on ECT signal map for different SNR are shown in Fig. 6.3 for a single test sample (arbitrarily chosen from the test set). The performance of the inversion scheme is analyzed by computing the normalized mean error (NME). Table 6.3 summarizes all the training and test set

parameters that are used for this analysis. Finally, the estimation of crack parameters on the unknown test set is evaluated (online phase) by the trained models.

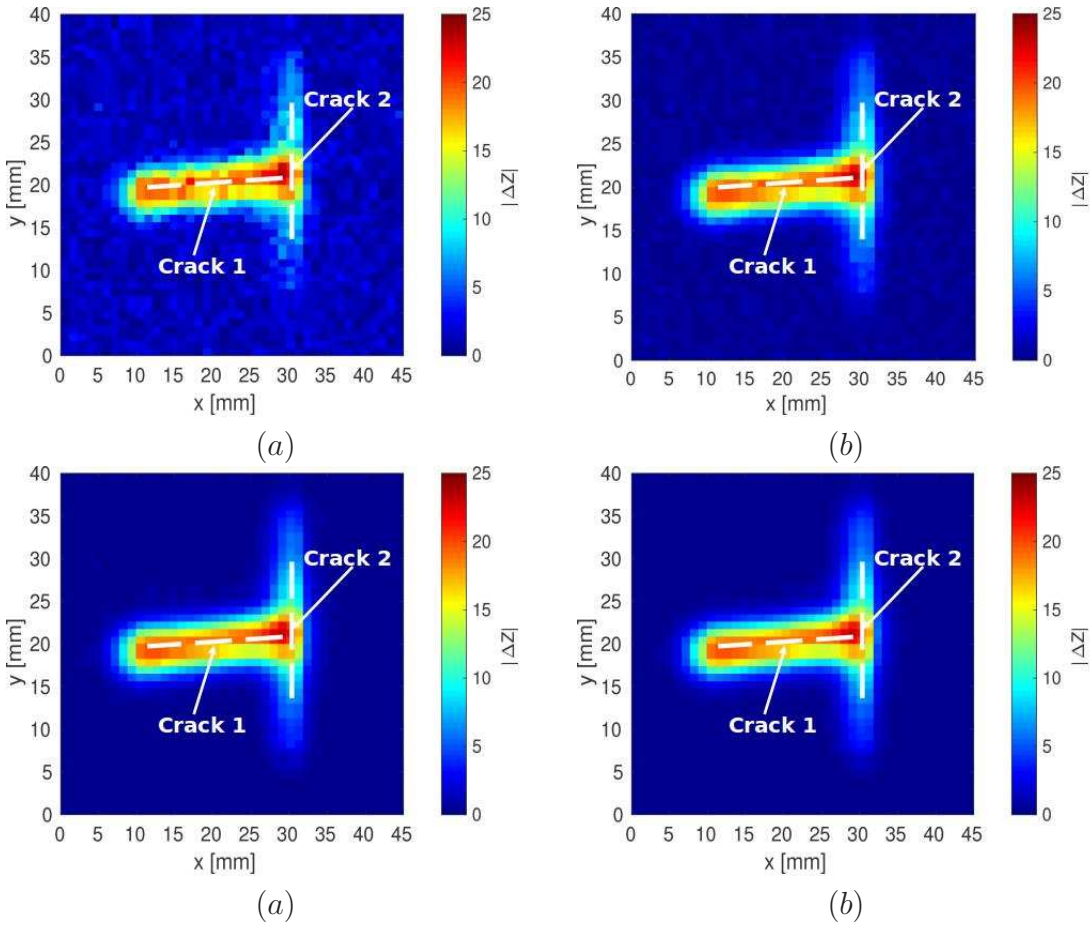


Figure 6.3: Numerical assessment: Examples of absolute value of coil impedance variation corrupted by AWGN through (a) $SNR = 10$ [dB], (b) $SNR = 20$ [dB], (c) $SNR = 30$ [dB] and (d) $SNR = 40$ [dB] for the cracks having length $l_1 = 10.08$ mm, width $w_1 = 0.186$ mm, angular position $\varphi_1 = 6.20$ deg, length $l_2 = 12.26$ mm, height $h_2 = 0.955$ and axial position $y_2 = 25.93$ mm.

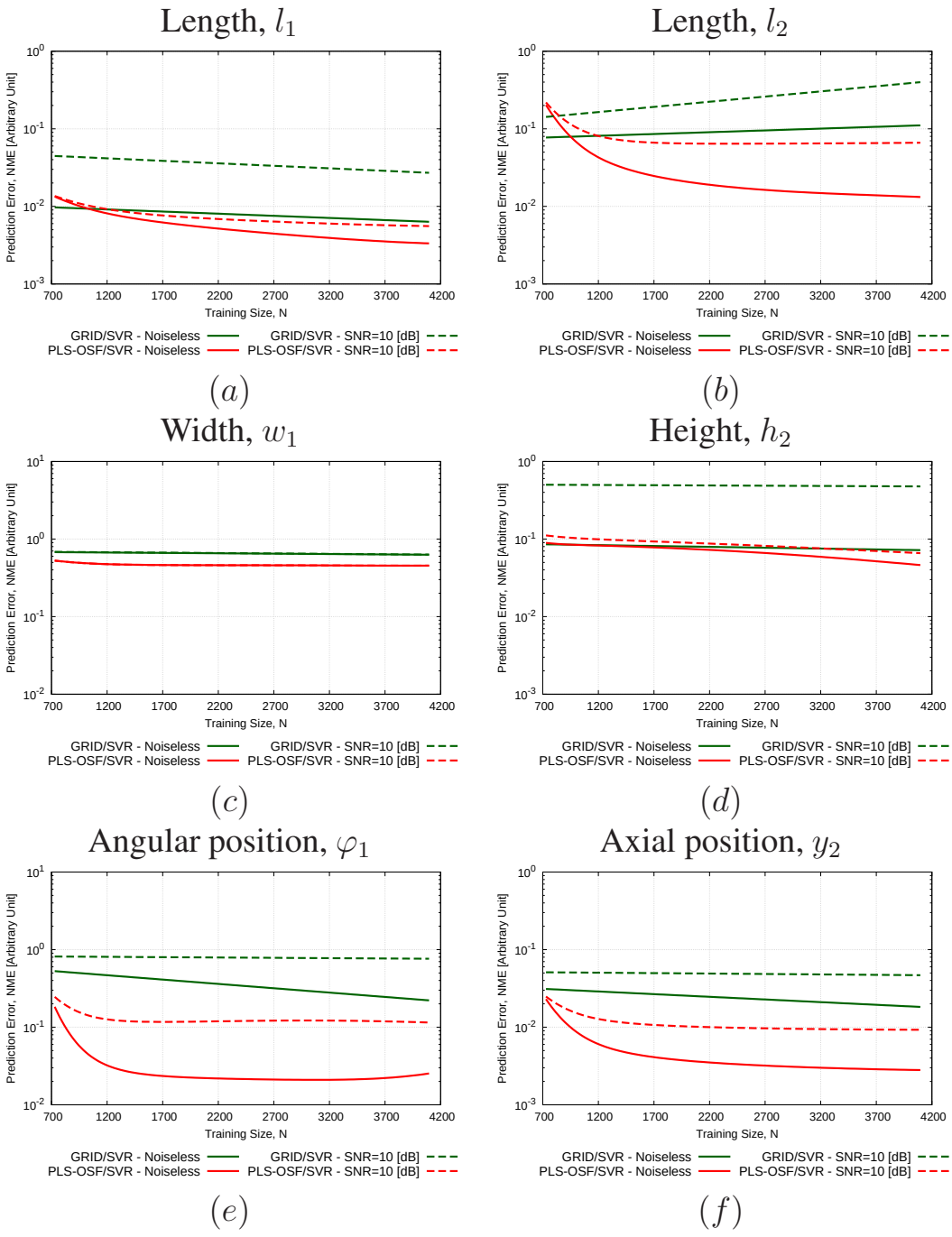


Figure 6.4: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparisons in terms of N vs NME are shown at $(a)(c)(e)$ crack 1 parameters (crack length l_1, width w_1, angular position φ_1) estimation and at $(b)(d)(f)$ crack 2 parameters (length l_2, height h_2, axial position y_2) estimation for test set at noiseless and $SNR = 10$ [dB] while $F = 3772$, $J = 15$, $M = 1000$.

6.3 Results and discussion

Having a six dimensional parametric problem at hand requires us to generate high number of training samples, which is cumbersome when using standard GRID approach. Applying PLS features extraction, ensures most significant information are projected to a much reduced extracted feature space from actual ECT features. By uniformly distributing the extracted feature space, redundant and unnecessary samples for different configurations of cracks parameters are avoided. As a consequence, more accurate model are obtained from smaller training sets. Figure 6.4 depicts the inversion performance comparison between PLS-OSF/SVR and GRID/SVR approaches for noiseless and noisy (i.e., $SNR = 10$ dB) test sets. Whereas, GRID/SVR shows higher prediction error for estimating almost all the parameters of crack 1 and crack 2 than PLS-OSF/SVR. PLS-OSF/SVR has shown much improved prediction accuracy on noiseless as well as noisy test sets for l_1, φ_1 and w_1 estimation once N is larger than 1000 samples.

The illustrations of the prediction performance are shown in terms of true vs. predicted plots for crack 1 and crack 2 in Fig. 6.5 and Fig. 6.6, respectively. Being a surface breaking crack (crack 1), length l_1 estimation is one of the easiest parameters to be estimated. Hence, at $N = 4096$ training samples, the prediction accuracy for both PLS-OSF/SVR and GRID/SVR are almost similar for noiseless test set. However, in case of noisy data (e.g., $SNR = 10$ [dB]), PLS-OSF/SVR outperforms GRID/SVR. Whereas, PLS-OSF/SVR shows higher prediction accuracy than GRID/SVR for angular position, φ_1 estimation (Fig. 6.4 and Fig. 6.5) for both noiseless and noisy test sets. ECT signal does not change significantly due to the variation of crack width w_1 (the probe diameter is larger than the gap size on this test case). Thus, crack width estimation shows higher prediction error among all the cracks parameters for both GRID/SVR and PLS-OSF/SVR approaches. From Fig. 6.1(b), we can see that impedance variation signals due to the presence of crack 1 (surface breaking crack) is much stronger than the one due to crack 2 (subsurface crack). Hence, ECT signals due to the presence of crack 2 are more prone to be affected by the noise (Fig. 6.3). Moreover, due to the crack parametric definition, we are varying the axial position y_2 while varying the length l_2, this makes more complex to estimate l_2. Even when increasing N, the prediction accuracy of l_2 estimation by GRID/SVR approach is not improved (Fig. 6.4 (b)). By adopting PLS-OSF, the extracted feature space is filled in such a way that the contribution of each parameter is retained and this phenomenon reflects higher prediction accuracy for l_2, h_2 and y_2 estimation than GRID/SVR approach.

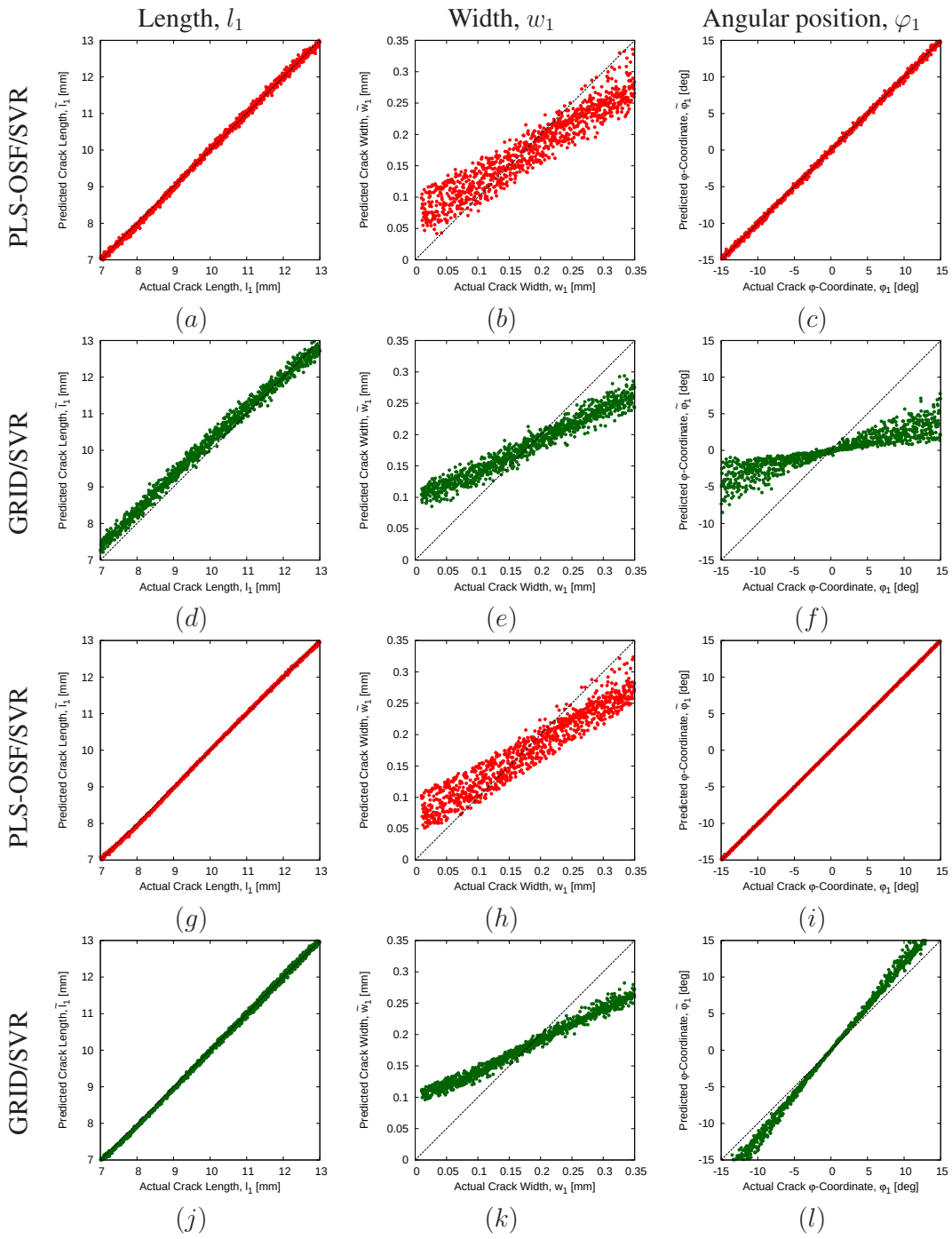


Figure 6.5: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparison are shown in terms of true vs. predicted (crack length l_1, width w_1, angular position φ_1) regression plots at $(a) - (f)$ on the test set at $SNR = 10$ [dB] and at $(g) - (l)$ on noiseless test set, while $N = 4096$, $F = 3772$, $J = 15$, $M = 1000$.

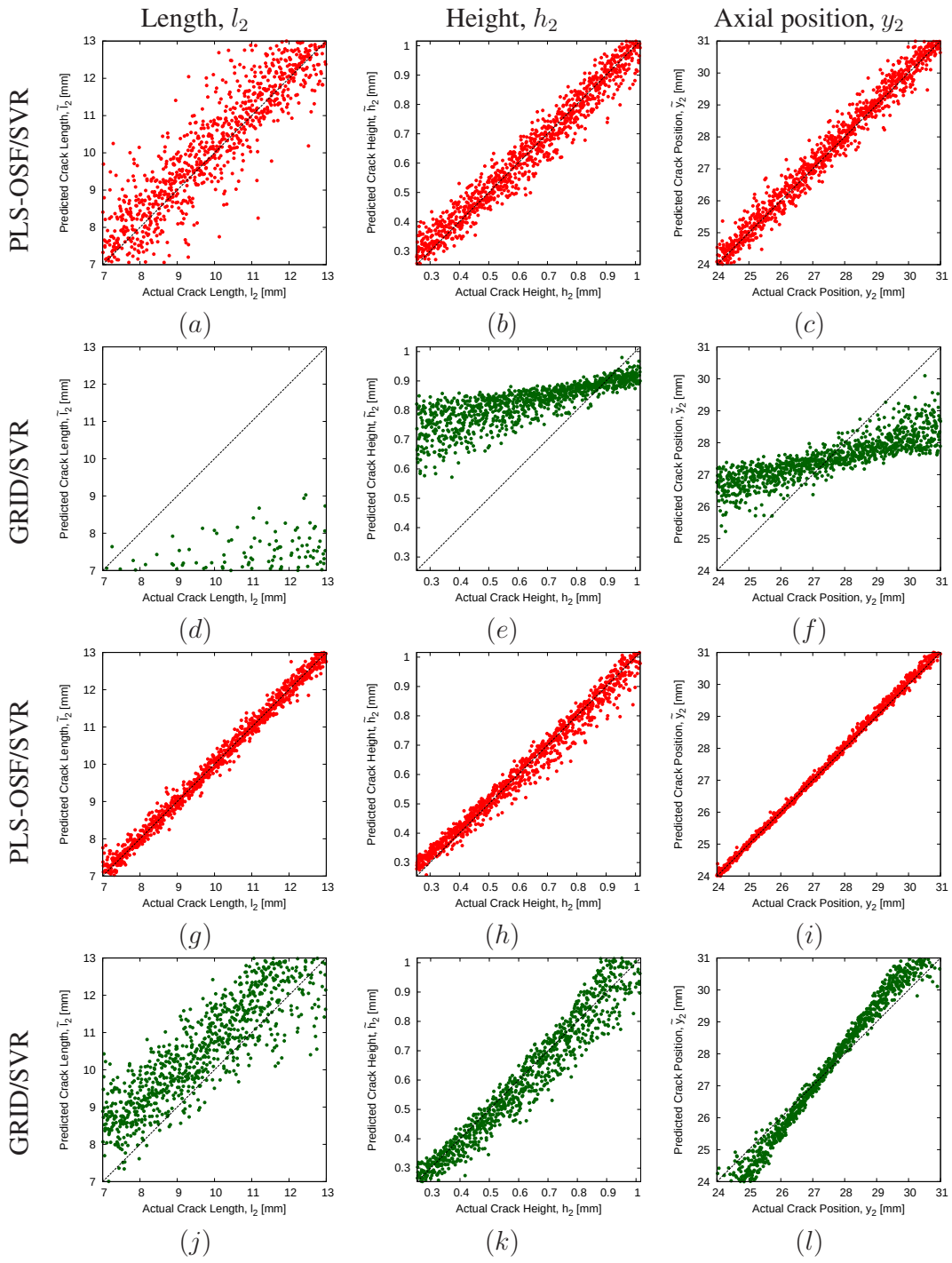


Figure 6.6: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparison are shown in terms of true vs. predicted (crack length l_2, height h_2, axial position y_2) regression plots at $(a) - (f)$ on the test set at $SNR = 10$ [dB] and at $(g) - (l)$ on noiseless test set, while $N = 4096$, $F = 3772$, $J = 15$, $M = 1000$.

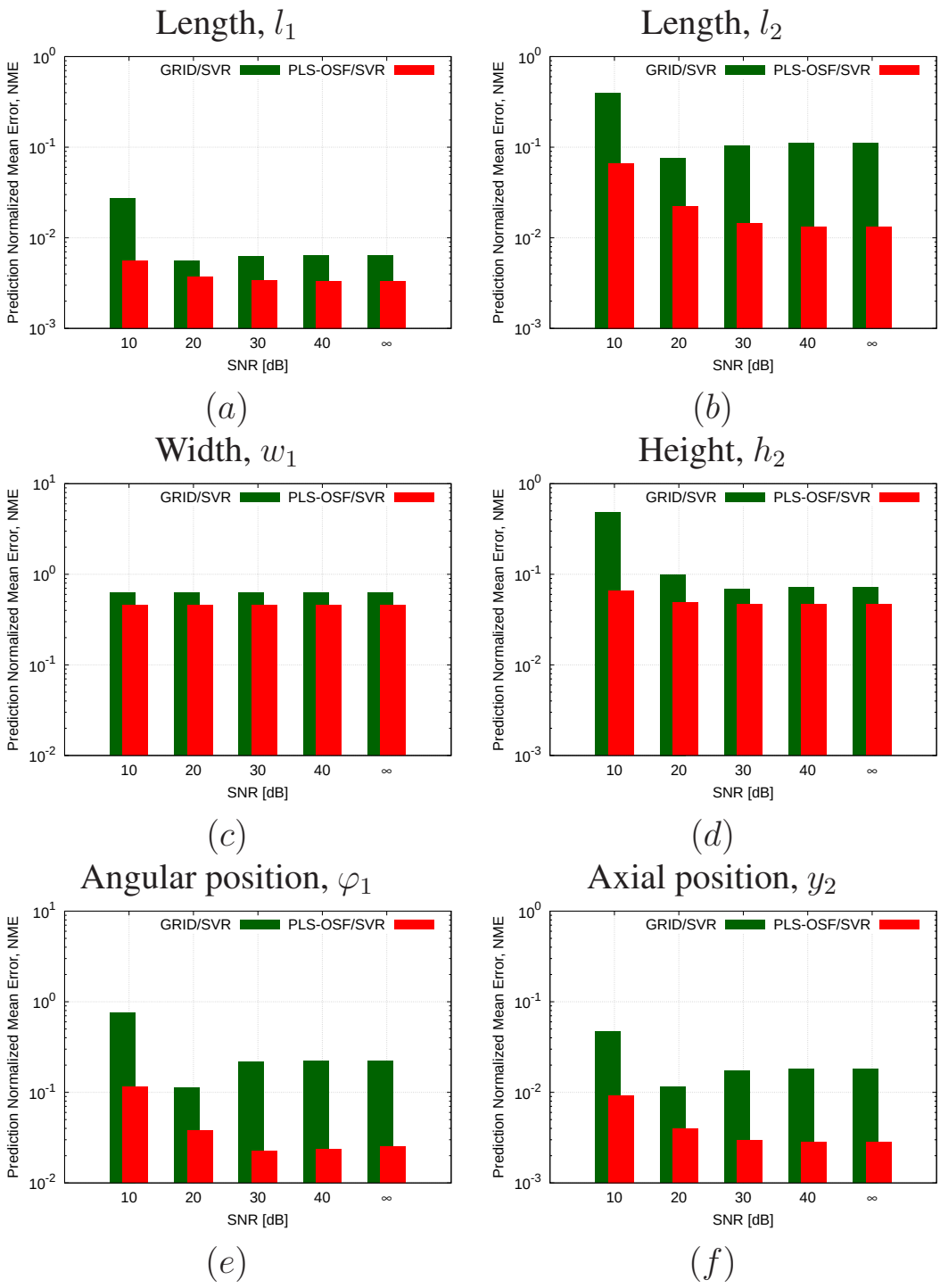


Figure 6.7: Numerical assessment: PLS-OSF/SVR vs. GRID/SVR prediction comparisons in terms of SNR vs. NME are shown at $(a)(c)(e)$ crack 1 parameters (crack length l_1, width w_1, angular position φ_1) estimation and at $(b)(d)(f)$ crack 2 parameters (length l_2, height h_2, axial position y_2) estimation for test set at noiseless and $SNR = 10,\ 20,\ 30,\ 40$ [dB] while $N = 4096,\ F = 3772,\ J = 15,\ M = 1000$.

Prediction accuracy of different crack parameters, in terms of prediction error (NME), are also shown in Fig. 6.7 for different SNR at $N = 4096$. It is also visible that PLS-OSF/SVR has higher prediction accuracy than GRID/SVR for predicting all crack parameters on both noisy and noiseless test sets. Moreover, PLS-OSF/SVR takes ~ 0.05 s for predicting $M = 1000$ test samples where GRID/SVR takes ~ 6 s on standard laptop.

6.4 Conclusion

In this test case, we have shown quasi real-time multi-crack characterization and localization by combining an adaptive sampling strategy through PLS feature extraction with OSF and SVR. The number of parameters to be inverted depends on the complexity of the problem, suitable ECT features and the training set size N. Standard high dimensional problem implies the generation of very huge training set which is also time demanding. Providing a certain complexity and with a good set of extracted features in the training set, the increment of Q impacts on the suitable training size N. By applying PLS-OSF/SVR, higher prediction accuracy has been obtained for performing inversion on the studied problem having six parameters for a training set of $N \geq 1000$ than GRID/SVR. One of the significant advantages of PLS-OSF strategy is to project the high dimensional ECT signals to a different and much lower dimensional space by PLS. This approach can be adopted to the other NDT application where GRID sampling may not work for the curse of dimensionality problem. In Chapter 7, PLS-OSF/SVR has been applied for both ECT and Ultrasound Testing for enhancing the diagnosis performance. In order to address such kinds of problems, an innovative data fusion strategy combining eddy current testing and ultrasound testing signals has been developed. The preliminary studies have shown promising results which will be discussed in the next Chapter.

Chapter 7

A preliminary study on crack characterization based on multiphysics NDT data fusion

This chapter describes an inversion methodology by exploiting Learning by examples (LBE) technique by using multi-physics simulation and data fusion approach. Within the framework of LBE, we have applied Eddy Current Testing (ECT) and Ultrasonic Testing (UT) NDT methods to solve a complex crack localization and characterization problem by developing accurate training models for each method. After which, an innovative multi-physics data fusion approach (ECT-UT) has been adopted for developing a robust inversion schema. Partial Least Squares (PLS) feature extraction combined with a customized version of output space filling (OSF) (i.e., PLS-OSF sampling) has been applied with the Support Vector Regression (SVR) algorithm for developing an accurate training model based on the training set and subsequently real-time inversion (online phase) has been performed on unknown test data. The robustness of the proposed ECT-UT approach is numerically assessed in presence of synthetic noisy test data through PLS-OSF/SVR strategy and compared with each single NDT method (i.e., ECT and UT). Furthermore, inversion performance of PLS-OSF/SVR strategy on ECT-UT data has been compared briefly with GRID-PLS/SVR strategy.

7.1 Problem definition: ECT

A homogeneous plate made by aluminium 2024 alloy of thickness 6 mm has been investigated by Eddy Current Testing (ECT) approach. The plate consists of a fastener (bore hole) of radius 3.75 mm and 6.00 mm height. The plate is affected by a single notch (e.g., narrow crack) of volume Ω having fixed width 0.01 mm and height 2 mm (Fig. 7.1) which is attached with the fastener. The narrow crack is characterized by total $Q = 3$ descriptors of length (l_c), ligament (δ_c), and angular distance (φ_c) (i.e., $\underline{p} = (l_c, \delta_c, \varphi_c)$). The plate is inspected by a single coil working in absolute mode

of frequency 1 kHz with lift off 1 mm. The coil collects the ECT signals from 81 positions along X directions with a step size of 0.5 mm and from 41 positions along Y directions with a step size of 1 mm, respectively through a raster scan. Therefore, ECT signals (i.e., impedance variation signal) are collected from $K = 81 \times 41 = 3321$ number of inspected points.

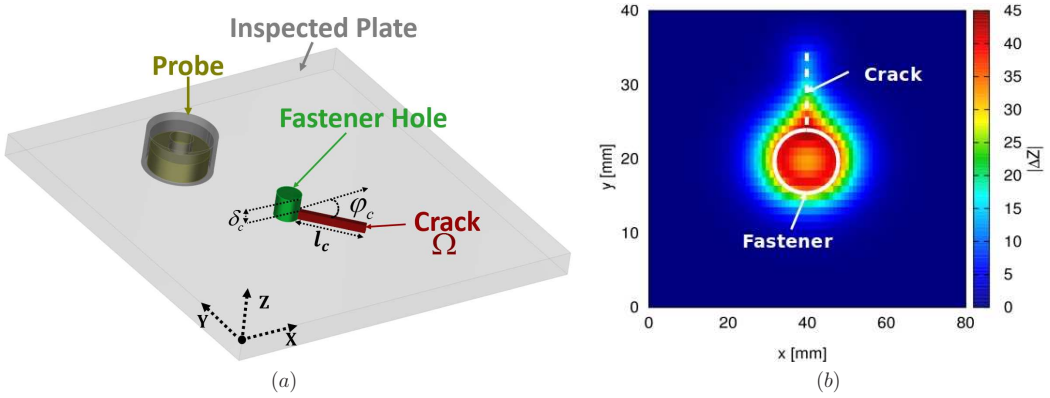
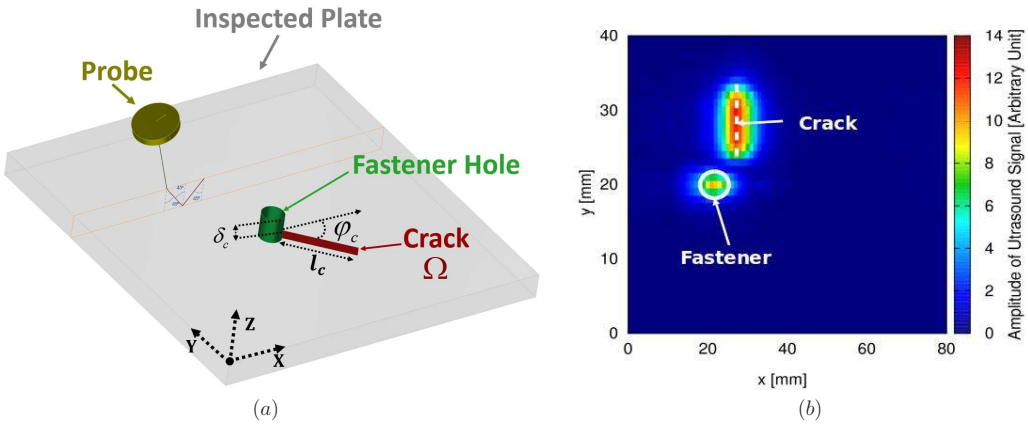


Figure 7.1: Examples of (a) studied plate geometry and (b) ECT signal map in terms of the absolute value of coil impedance variation.

7.2 Ultrasound testing treatment

In this case, the same problem described in Sec. 7.1 will be treated through Ultrasonic Testing (UT) NDT method. The plate made by aluminum 2024 alloy, having thickness 6 mm and density 2.77 g.cm^{-3} consists of a fastener (bore hole). The plate is affected by a single notch (e.g., narrow crack) Ω having fixed width 0.01 mm and height 2 mm (Fig. 7.2) which is attached with the fastener. The crack Ω is characterized by total $Q = 3$ descriptors by of length (l_c), ligament (δ_c), and angular distance (φ_c) (i.e., $\underline{p} = (l_c, \delta_c, \varphi_c)$). The plate has been investigated by a ray probe by using water coupling medium (i.e., density 1 g.cm^{-3}). The probe is acting both for transmitting and receiving UT signals. The probe is positioned in a way such that the incident angle (i.e., incident angle $19.381°$) of the transmitting UT ray generate $45°$ shear waves (i.e., secondary waves/S-waves) in the aluminum plate. More details of the treated problem and probe definition are available in [69]. The inspecting probe collects the UT signals from 81 positions along the X direction with step size 0.5 mm and 41 position along Y direction with step size 1 mm through a raster scan. Therefore, UT signals (i.e., reflected signals) are collected from $K = 81 \times 41 = 3321$ number of inspected points. CIVA simulator uses a hybrid model based on Kirchhoff approximation and high-frequency Geometrical Theory of Diffraction (GTD) for generating diffraction/scattering waves from planner-like defects [70]. Brief description of UT signal formulation is also available at Chapter A.1. The amplitude of the reflections/scatters wave have been collected through a C-Scan (e.g., maximum ray

amplitude available at each inspection point).



(a) (b)

Figure 7.2: Examples of (a) studied plate geometry and (b) UT signal map in terms of amplitude of ultrasound signal.

7.3 Data fusion using ECT and UT

In the previous sections, we have defined the (same) test case from the ECT and UT perspective. Each NDT method has some pros and cons according to their own physics. That is, for the mentioned NDT problem at hand, ECT signal is mostly affected for the presence of fastener. Due to the significant probe impedance variation, the area of the fastener (i.e., bore hole) is acting as a circular defect within a plate. The impedance variation due to the fastener is much stronger than narrow crack, thus, when the crack is placed deeper inside the plate, the ECT signals contribution due to the presence of narrow crack is becoming weaker. From Fig. 7.3 (c), we can observe that the crack located at ligament distance $\delta_c = 4.00$ mm deeper from the surface of the plate is not visible. However, the same crack placed at the same position can be easily observed while UT inspection is performed (Fig. 7.3 (d)). Conversely, in Fig. 7.3 (b), UT signals are weak while the crack is placed at the top surface of the plate (i.e., $\delta = 0$ mm), but ECT inspection highlights the presence of crack (Fig. 7.3 (a)). From the perspective of the UT inspection, this phenomena is due to the fact that, only the back scattered waves that are reflected from the crack surface are collected as UT signals. Whereas, the crack is moving towards the bottom surface of the plates the UT probe collects the reflected waves coming from the crack surface and the back end surface of the plate. Considering all these reflected waves increase the amplitude of the UT signal due to the presence of the crack. This behavior obviously depends on the crack position and on the NDT method (i.e., ECT and UT) employed. As a consequence, different impacts on the crack characterization and localization performance are expected based on ECT and UT methods.

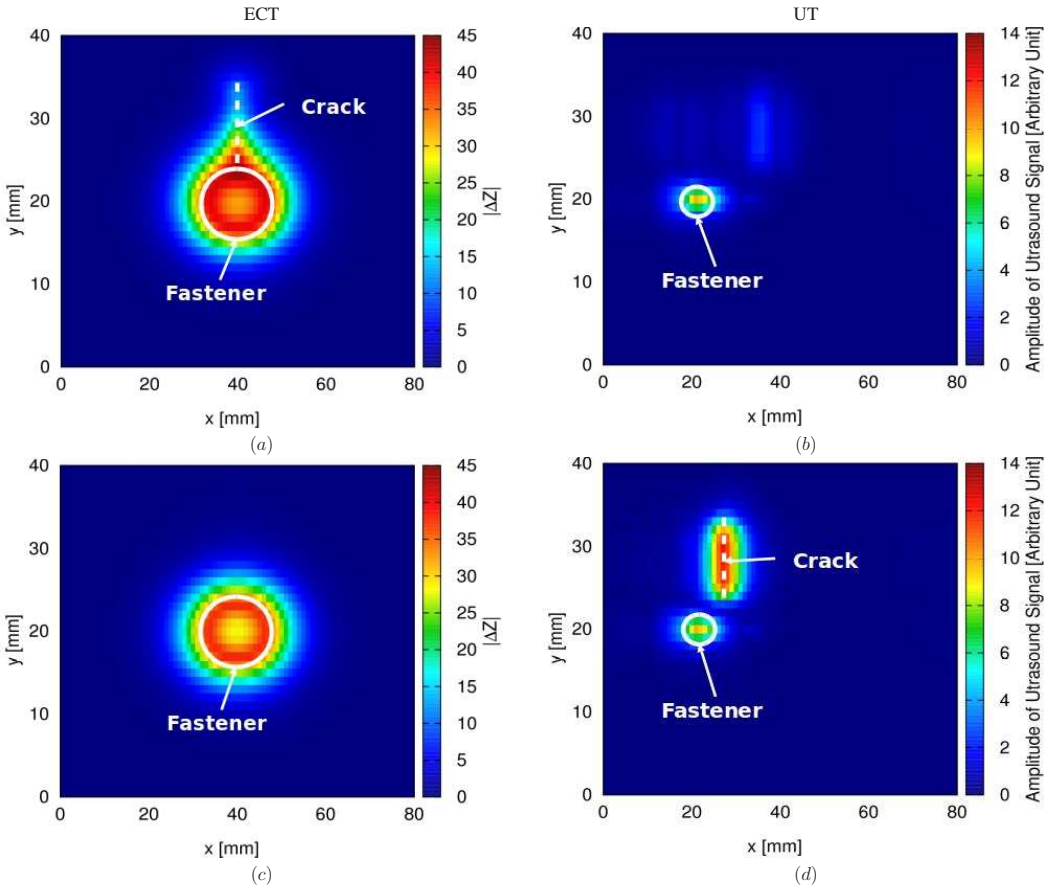


Figure 7.3: ECT and UT signal map representation for the crack configuration at $(a) - (b)$ length $l_c = 10$ mm, ligament position $\delta_c = 0$ mm, angular position $\varphi_c = 90$ deg; and at $(c) - (d)$ length $l_c = 10$ mm, ligament position $\delta_c = 4$ mm, angular position $\varphi_c = 90$ deg.

One of the ways to deal with this problem is to develop an inversion schema that allows us to consider both of the NDT methods in order to maximize the crack characterization and localization capability. We have applied multi-physics data fusion strategy to enhance the inversion performance. In this data fusion approach, the ECT signals and UT signals are generated separately by their own forward solver (i.e., CIVA [15]) and the training sets of both these data sets are fused by concatenating ECT and UT data which is represented by ECT-UT data.

7.4 Training model generation

This section describes the numerical configuration of the addressed problem for performing inversion. ECT and UT signals are separately generated by using two different forward solver (i.e., CIVA [15] modules based on ECT [71] and UT [72] models).

Six training sets (corresponding to three different crack parameters) for six different training models have been created by changing the crack dimensions within the range mentioned in Table 7.1 accordingly by PLS-OSF and GRID-PLS sampling approaches. For the evaluation purpose, the developed training models are treated for three different cases: (1) ECT signals, (2) UT signals and (3) multi-physics by combination of ECT and UT signals, denoted by ECT-UT (i.e., data fusion). ECT-UT data are generated by concatenating ECT and UT signals. It is worth to be mentioned that, both ECT and UT signals are different and compatibility of each signal with other while concatenating is the vital issue for this data fusion. For example, ECT signals are complex signals (having both real and imaginary parts) while UT signals provide real values only. Therefore, the dimension of ECT signal space is twice than UT signal space. Thanks to PLS feature extraction, the original NDT signals are transformed to a different reduced feature space, which is totally different from the original NDT signal space. Similarly, actual ECT-UT signals are transformed to the reduced feature space. Therefore, same number of extracted features can be used for the performance evaluation of each NDT method.

Crack parameters	**Min**	**Max**
Length l_c [mm]	3.00	10.00
Ligament position δ_c [mm]	0	4.00
Angular position φ_c [deg]	0	90

Table 7.1: Crack parameters ranges.

The initial number of samples $N_0 = 27$ and $N = 216$ maximum samples are chosen for the appropriate performance comparison between ECT, UT and ECT-UT data fusion for using PLS-OSF sampling strategy (see Sec. 3.3.5). Each of these samples is associated to the number of features $F_{ECT} = 2K = 6642$, $F_{UT} = K = 3321$ and $F_{ECT-UT} = 9963$ for ECT, UT and ECT-UT signals, respectively. The corresponding features of each data set have been reduced to J number of extracted features by PLS feature extraction strategy.

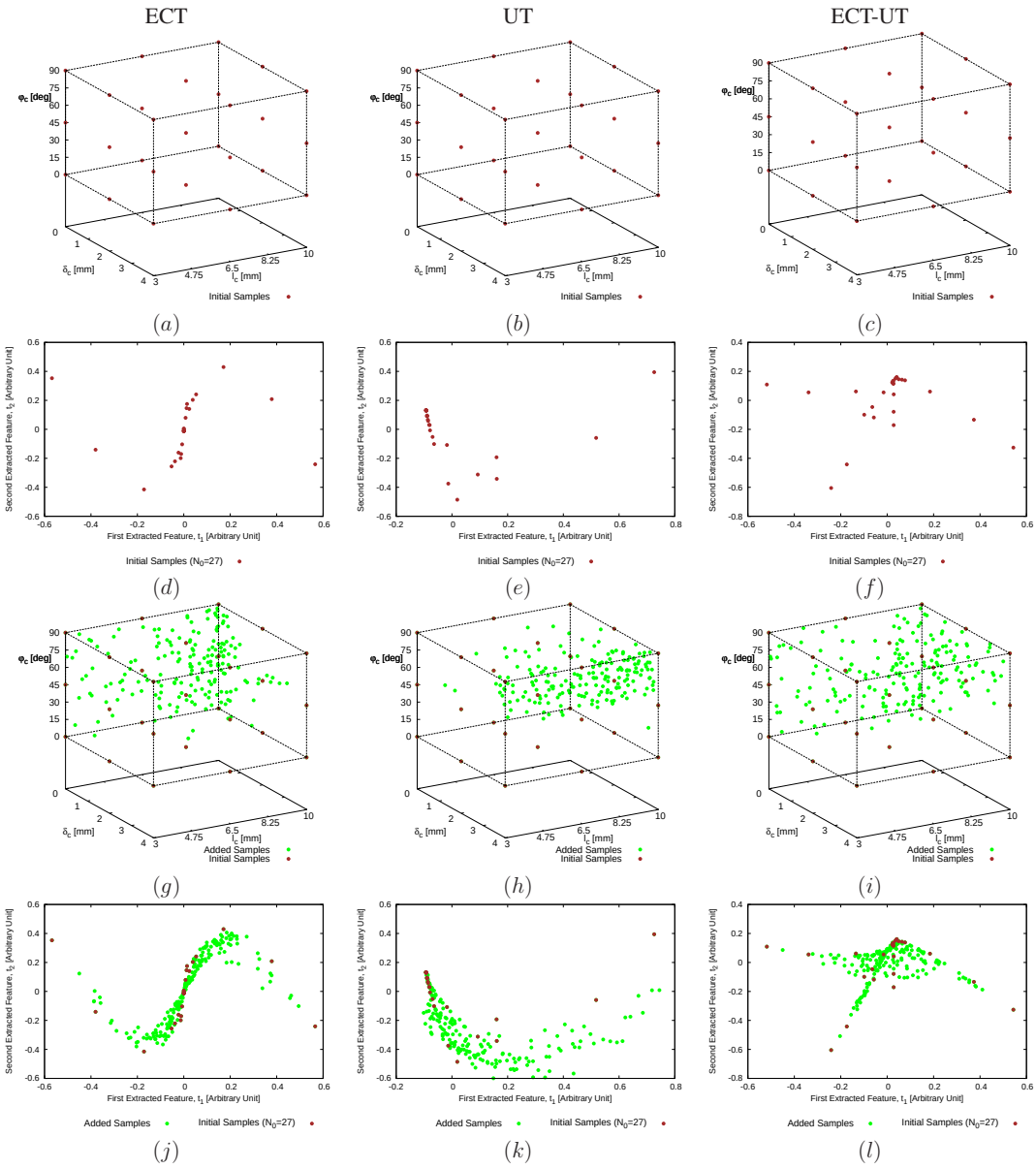


Figure 7.4: Disposition of the training samples generated by PLS-OSF on ECT, UT and ECT-UT data in (a)-(c), (g)-(i) the crack parameter space and (d)-(f), (j)-(l) the extracted feature space for $(a)(f)$ $N_0 = 27$ and $(g)(l)$ $N = 216$

Figure 7.4 represents adaptive training samples generation in the parameter space as well as in the extracted feature space for all the data sets (i.e., ECT, UT and ECT-UT). Due to the different perspectives of NDT methods, the new samples are adaptively added for lower and higher values of δ_c by performing ECT (Fig. 7.4 (g)) and UT (Fig. 7.4 (h)) inspections, respectively. Whereas, by applying ECT-UT data fusion, new samples are added in a way so as to consider both higher and lower values of δ_c (Fig. 7.4 (i)). Dealing with higher number of ECT-UT data (i.e., $F_{ECT-UT} = 9963$)

increases curse of dimensionality problem, which needs high number of training samples, N for developing accurate training models by standard sampling strategies (e.g., GRID). Therefore, unlike previous test cases, instead of GRID sampling, GRID-PLS sampling (see Sec. 3.3.3) strategy has been adopted for the performance evaluation comparison with PLS-OSF sampling strategy.

7.4.1 Parameter selection for ECT-UT data fusion

As discussed before, ECT and UT signals need proper treatment before performing training set generation and as well as training model creation by using data fusion. The behavior of each extracted data set obtained by ECT and UT is different. For example, imposing synthetic noise on ECT is not same as imposing synthetic noise in UT signals. Having stronger impact for the presence of fastener, ECT signal is more prone to be corrupted than UT signal data. From Fig. 7.6 and Fig. 7.7 we can see that how Additive White Gaussian Noise (AWGN) corrupts both ECT and UT signals. Therefore, we need to consider this issue while choosing optimum PLS parameter (i.e., J) and SVR hyper parameters (e.g., values of C and γ parameters). Thence, in case of ECT-UT data fusion, optimal J value has been chosen from a range $J \in [2,\ 25]$ by the calibration strategy mentioned in Sec. 4.4.3 for PLS-OSF/SVR strategy. In this case, cross-validation step has been done by SVR on the ECT-UT training set having $N = 216$ training samples. Based on the obtained sub optimal (C, γ) pairs on each J, inversion has been performed on an unknown test set of $M = 1000$ samples for each crack parameter. Finally, by performing the full calibration steps mentioned in Sec. 4.4.3, optimal number of extracted features is found at $J^{opt} = 20$ and the corresponding SVR parameters (C_q^{opt}, γ_q^{opt}) are obtained. Thus, $J^{opt} = 20$ number of features are extracted from $F = 3 \times K = 9963$ number of ECT-UT features. Figure 7.5 shows the prediction error analysis for each crack parameter on different J values and different SNR.

	PLS-OSF			GRID-PLS		
Parameters	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE
l_c	2.00	-1.00	2.13×10^{-01}	2.00	-1.00	4.67×10^{-01}
ρ_c	2.00	-1.00	6.71×10^{-02}	1.00	-1.00	8.20×10^{-02}
φ_c	5.00	-2.00	7.70	6.00	-3.00	1.40×10^{01}

Table 7.2: Set of optimal calibration parameters found for both PLS-OSF and GRID-PLS sampling approaches for the training sets of ECT-UT data fusion at $N = 216$ and $J^{opt} = 20$.

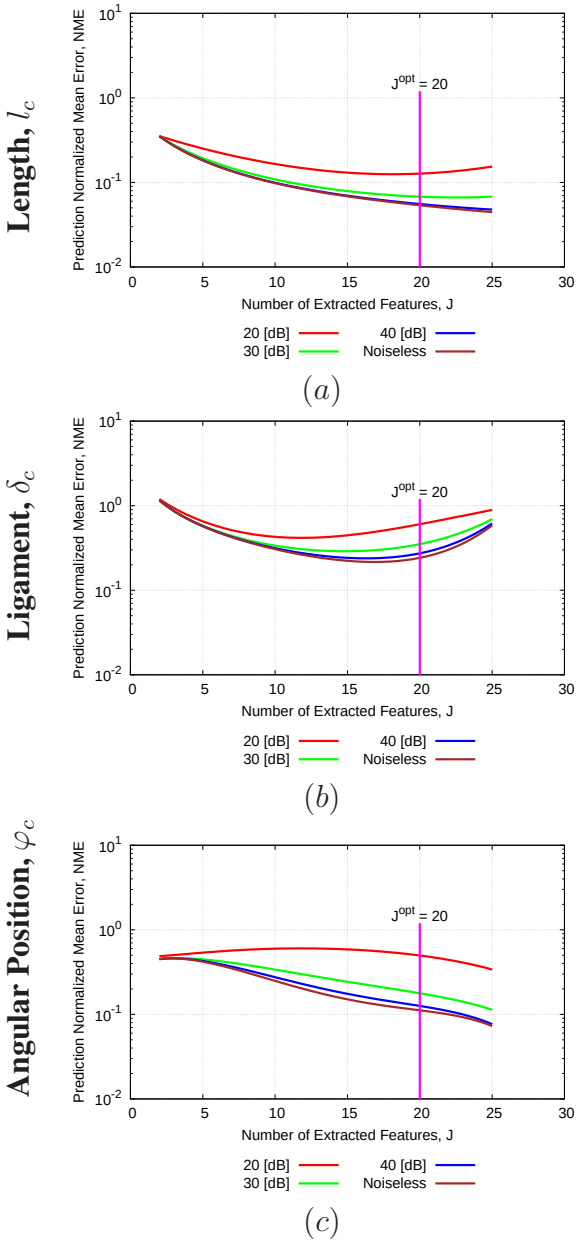


Figure 7.5: Prediction Normalized Mean Error, (NME) vs. J by inversion performed on the test set at $SNR = 20,\ 30,\ 40$ [dB] and $Noiseless$ for crack (a) length l_c, (b) ligament δ_c and (c) angular position φ_c estimation with, $N = 216$, $M = 1000$ through ECT-UT.

For the case of GRID-PLS/SVR sampling, (C_q^{opt}, γ_q^{opt}) pairs are obtained by performing cross-validation as described in Sec. 4.4.1 for $J^{opt} = 20$ and $N = 216$ training samples. Table 7.2 describes the (C_q^{opt}, γ_q^{opt}) pairs that have been used in the rest of the analysis for ECT-UT method for using both PLS-OSF/SVR and GRID-PLS/SVR strategies.

7.4.2 Parameter selection for ECT and UT signals

To be consistent and for valid comparison, the optimal number of extracted feature, $J^{opt} = 20$ is used for both ECT and UT signals separately. Therefore, ($C^{opt}_{q;\,J=20}$, $\gamma^{opt}_{q;\,J=20}$) parameters are obtained by performing cross-validation as described in Sec. 4.4.1 for $J = 20$, and $N = 216$ samples for ECT and UT training sets for each crack parameter by applying PLS-OSF/SVR and GRID-PLS/SVR strategies.

Tables 7.3 and 7.4 describe all the obtained ($C^{opt}_{q;\,J=20}$, $\gamma^{opt}_{q;\,J=20}$) parameters for ECT and UT signals, respectively. ε is kept fixed at 10^{-1} for all the cases.

	PLS-OSF			GRID-PLS		
Parameters	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE
l_c	2.00	-1.00	2.86×10^{-01}	2.00	-1.00	3.60×10^{-01}
ρ_c	3.00	-1.00	4.64×10^{-02}	1.00	0	5.84×10^{-02}
φ_c	3.00	-1.00	4.15×10^{01}	6.00	-3.00	2.92×10^{01}

Table 7.3: Set of optimal calibration parameters found for both PLS-OSF and GRID-PLS sampling approaches for the training sets of ECT data at $N = 216$ and $J^{opt} = 20$.

	PLS-OSF			GRID-PLS		
Parameters	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE	$log_{10}(C^{opt})$	$log_{10}(\gamma^{opt})$	MSE
l_c	3.00	-1.00	8.74×10^{-02}	2.00	-1.00	6.14×10^{-01}
ρ_c	3.00	-2.00	1.85×10^{-01}	2.00	-1.00	4.63×10^{-01}
φ_c	6.00	-2.00	2.70×10^{01}	3.00	-1.00	4.33×10^{01}

Table 7.4: Set of optimal calibration parameters found for both PLS-OSF and GRID-PLS sampling approaches for the training sets of UT data at $N = 216$ and $J^{opt} = 20$.

Afterwhich, different training models (offline phase) are obtained for q-th crack parameter ($q = 1, ..., Q$) separately for ECT, UT and ECT-UT data fusion methods by the corresponding SVR parameters.

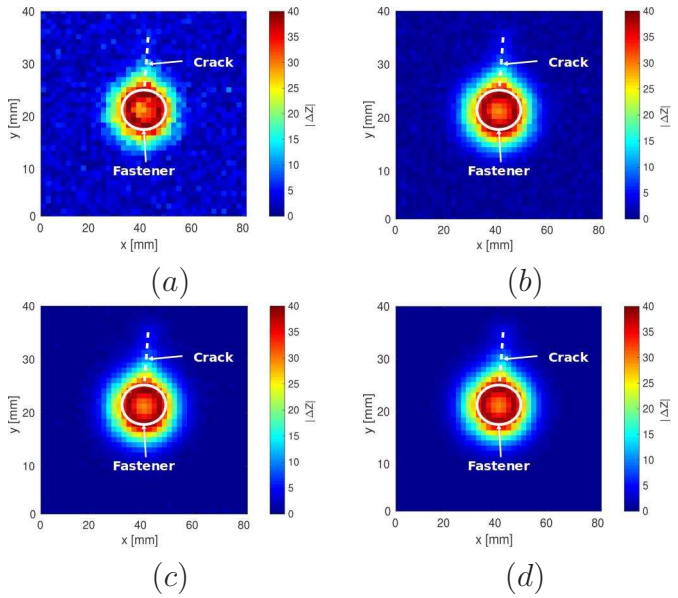


Figure 7.6: Examples of amplitude of ECT signals corrupted by AWGN through (a) $SNR = 10$ dB, (b) $SNR = 20$ dB, (c) $SNR = 30$ dB and (d) $SNR = 40$ dB for a crack having length l_c = 9.80 mm, ligament distance $\delta_c = 0.451$ mm and angular position $\varphi_c = 85.70$ deg.

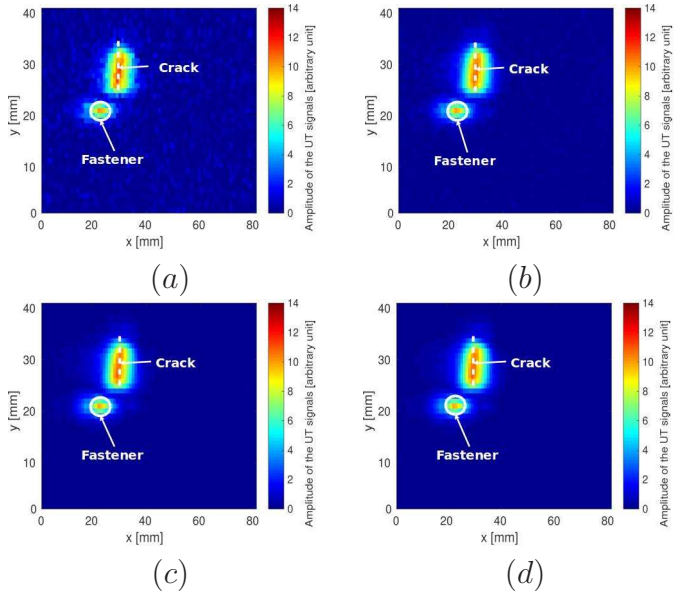


Figure 7.7: Examples of amplitude of UT signals corrupted by AWGN through (a) $SNR = 10$ dB, (b) $SNR = 20$ dB, (c) $SNR = 30$ dB and (d) $SNR = 40$ dB for a crack having length l_c = 9.36 mm, ligament distance $\delta_c = 3.62$ mm and angular position $\varphi_c = 87.77$ deg.

7.5 Test set generation

An unknown test set of $M = 1000$ samples for 3 crack parameters has been generated by using LHS design (see Sec. 3.2.2). Similar to the training sets, the same number of ECT features (i.e., $F_{ECT} = 2K = 6642$), UT features (i.e., $F_{UT} = 2K = 6642$) and ECT-UT features (i.e., $F_{ECT-UT} = 3K = 9963$) are considered for each test sample. For both sampling strategies (i.e., PLS-OSF and GRID-PLS), $J = 20$ features are extracted by projecting the test sets into extracted feature space through the PLS weight matrix (e.g., obtained from the corresponding ECT, UT and ECT-UT methods during training phase) for each test set. To partially consider noise effects, AWGN for different SNR is applied for blurring ECT and UT signals separately. The noise effects on ECT signal map for a given crack configuration (arbitrarily chosen from the test set) on different SNR are shown in Fig. 7.6. Similarly, UT signals are also corrupted by imposing AWGN through SNR (Fig. 7.7).

Table 7.5 shows the training and test set parameters that are used for this analysis. Finally, the estimation of crack parameters on the unknown test set is evaluated (online phase) by the developed trained models.

Training set parameters	
Number of ECT features, $F_{ECT} = 2K$	6642
Number of UT features, $F_{UT} = K$	3321
Number of ECT-UT features, $F_{ECT-UT} = 3K$	9963
Number of extracted features, J	20
Sampling strategies	PLS-OSF, GRID-PLS
Number of initial samples, N_0	27
Maximum number of samples, N	216
Test set parameters	
Number of ECT features, $F_{ECT} = 2K$	6642
Number of UT features, $F_{UT} = K$	3321
Number of ECT-UT features, $F_{ECT-UT} = 3K$	9963
Number of extracted features, J	20
Sampling strategies	LHS
Number samples, M	1000
Signal to noise ratio, SNR	10, 20, 30, 40 [dB]

Table 7.5: Training and test set parameters.

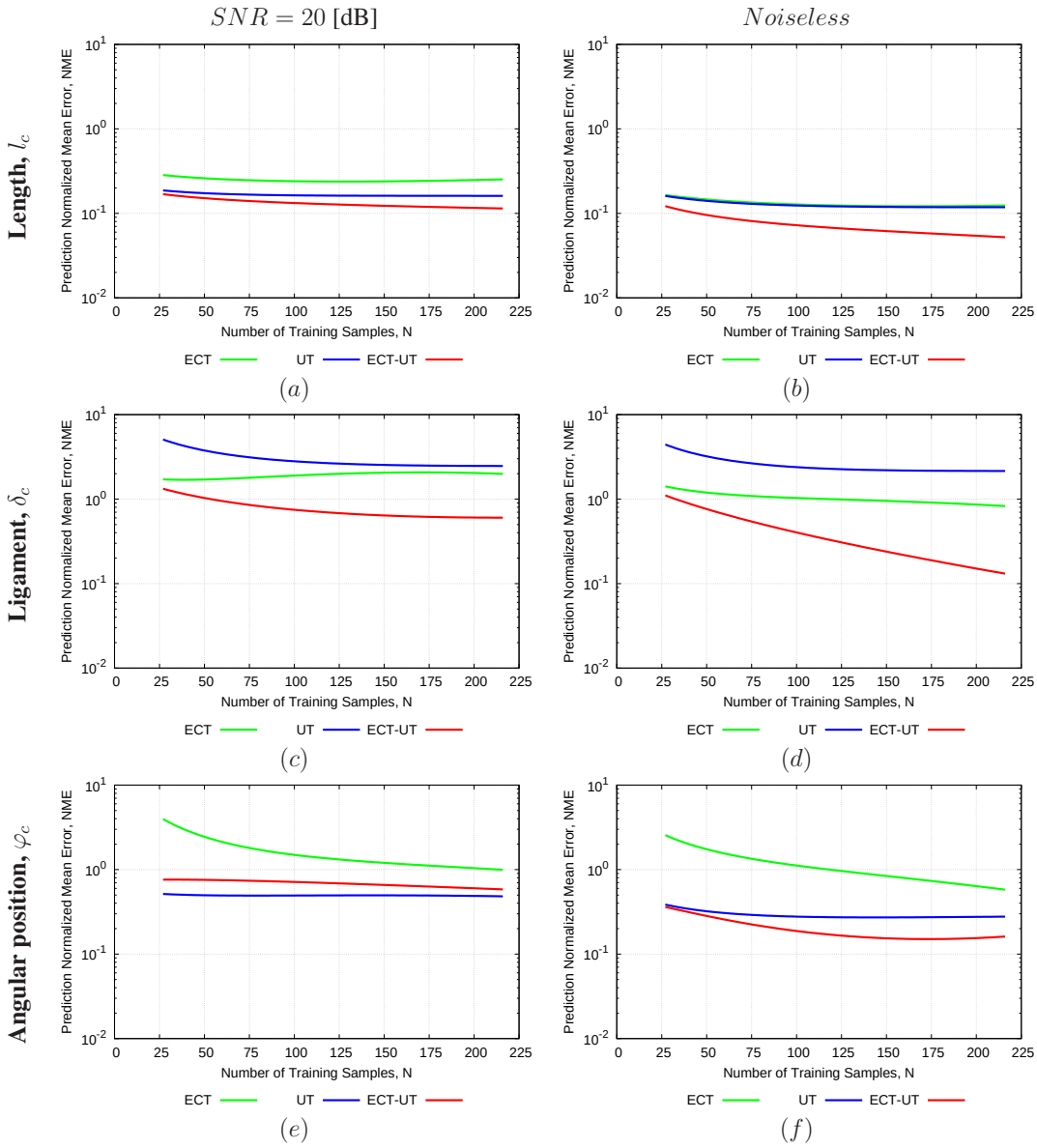


Figure 7.8: Numerical assessment: (crack length, $l_c = [3, 10]$ mm, ligament, $\delta_c = [0, 4]$ mm, and angular position, $\varphi_c = [0, 90]$ deg.) - PLS-OSF/SVR prediction results are shown in terms of prediction error (Normalized Mean Error, NME) vs. number of training samples, N for crack (a)-(b) length l_c, (c)-(d) ligament δ_c and (e)-(f) angular position φ_c estimation for $J = 20$, at $(a)(c)(e)$ test set at 20 dB, and at $(b)(d)(f)$ $Noiseless$ test set with $M = 1000$ through ECT, UT and ECT-UT.

7.6 Results and discussion

By combining both ECT and UT signals, ECT-UT data fusion contains both information from ECT and UT signals. This improves the learning ability of the SVR for training model development. Therefore, ECT-UT can enhance the prediction accuracy by using PLS-OSF/SVR approach. By applying PLS feature extraction, we can retrieve most significant information from the combined signals (i.e., ECT-UT). As we have found ECT signals are mostly corrupted for imposing noise, by combining ECT and UT signals, we can improve the overall inversion performance. Fig. 7.8 (a) shows that l_c estimation by using ECT suffers on noisy data than UT signals, while by using for both ECT and UT signals l_c estimation has similar prediction accuracy on $Noiseless$ test set (Fig. 7.8 (b)). On the other hand, crack ligament distance δ_c estimation is showing lower prediction error for adopting ECT signals than UT signals for both noisy and $Noiseless$ test set (Fig. 7.8 (c), (d)). Whereas, Fig. 7.8 (e) and Fig.7.8(f) depict that applying UT signals exhibit improved crack angular position φ_c estimation than ECT signals on noisy and $Noiseless$ test sets.

It is also visible that by combining both signals we can be benefited from both NDT methods and adaptively increasing training samples through PLS-OSF improves prediction accuracy for all the crack parameters. Hence, by increasing N, ECT-UT data fusion shows improved prediction accuracy on crack length l_c, ligament distance δ_c and angular position φ_c estimation through PLS-OSF/SVR strategy than ECT and UT methods separately. The inversion schema for using ECT-UT data fusion becomes more robust on $Noiseless$ and in presence of noisy test sets.

Fig. 7.9 and Fig. 7.10 show the scatter plots of true vs. predicted crack parameters obtained by PLS-OSF/SVR strategy for $N = 216$. Qualitatively speaking, ECT-UT data fusion improves the learning ability that provides better l_c, δ_c and φ_c estimation than ECT and UT signals. By employing ECT signals, l_c estimation becomes more difficult due to the presence of fastener than applying UT signals. However, ECT-UT provides higher prediction accuracy for l_c estimation on noisy as well as noiseless test set. At Fig. 7.8 (e) we found that applying UT method provides lower φ_c estimation error than ECT-UT on noisy data, however, Fig. 7.9(h) and Fig. 7.9(i) show, ECT-UT signals are showing better prediction than UT signals. This contradictory phenomenon may arise due to the definition of normalized mean error (see Sec. 4.3.2), where prediction error may increase if the actual parameter contains some values which is near to zero. In Fig. 7.11, the inversion performance accuracy of ECT-UT data fusion in terms of NME for different SNR at $N = 216$ are compared with ECT and UT methods. Here, it also highlights that ECT-UT data fusion improves the prediction accuracy which is robust on noisy data.

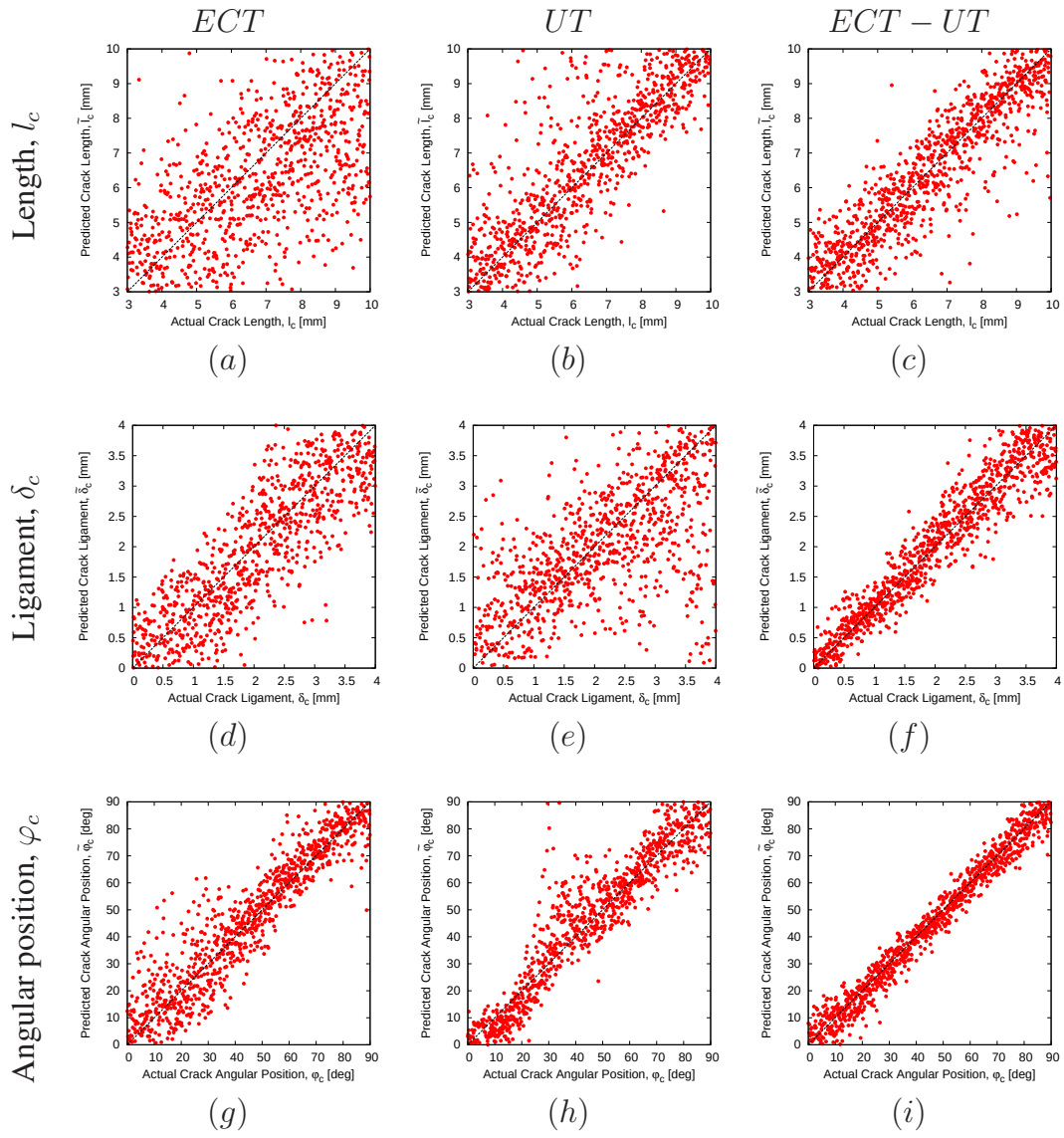


Figure 7.9: Numerical assessment: (crack length, $l_c = [3,\ 10]$ mm, ligament, $\delta_c = [0,\ 4]$ mm, and angular position, $\varphi_c = [0,\ 90]$ deg.) - PLS-OSF/SVR prediction results are shown in terms of true vs. predicted (Crack length l_c, ligament δ_c and angular position φ_c) regression plots on the test set at $SNR = 20$ [dB], for $N = 216$, $J = 20$, $M = 1000$ through ECT, UT and ECT-UT.

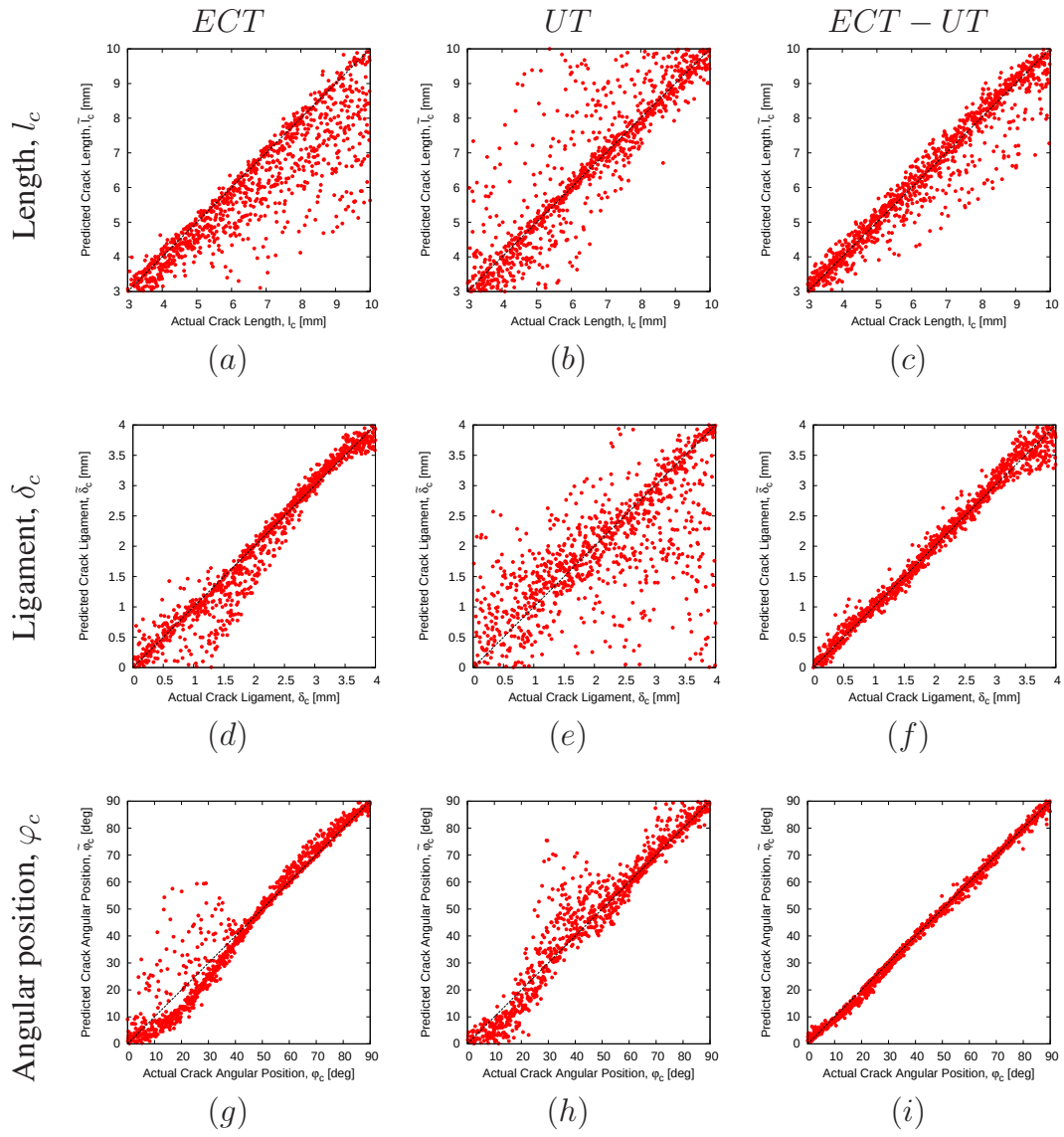


Figure 7.10: Numerical assessment: (crack length, $l_c = [3,\ 10]$ mm, ligament, $\delta_c = [0,\ 4]$ mm, and angular position, $\varphi_c = [0,\ 90]$ deg.) - PLS-OSF/SVR prediction results are shown in terms of true vs. predicted (Crack length l_c, ligament δ_c and angular position φ_c) regression plots on $Noiseless$ test set for $N = 216$, $J = 20$, $M = 1000$ through ECT, UT and ECT-UT.

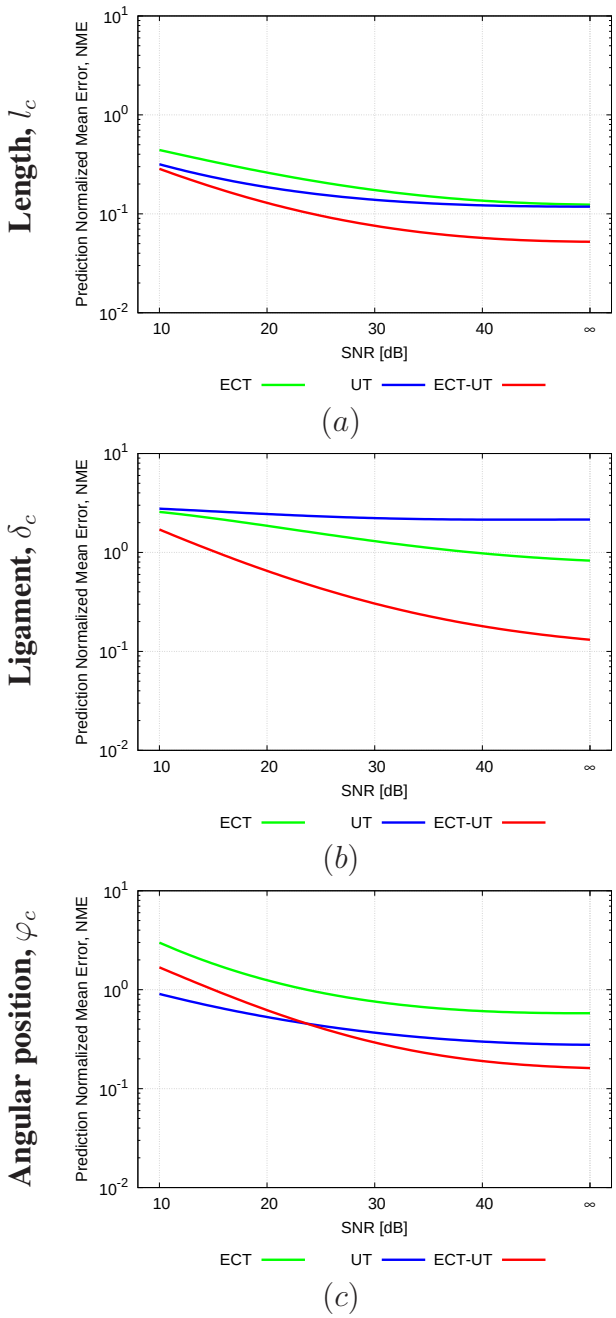
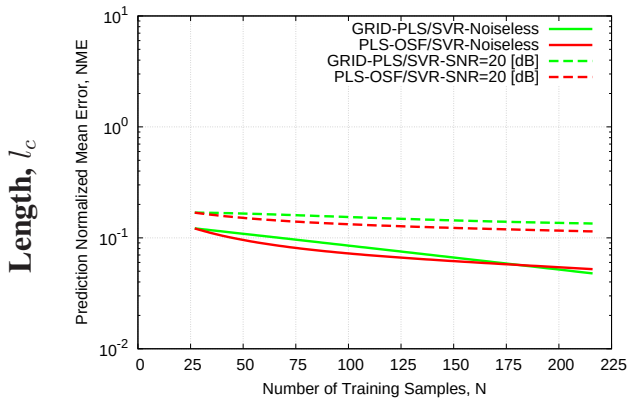
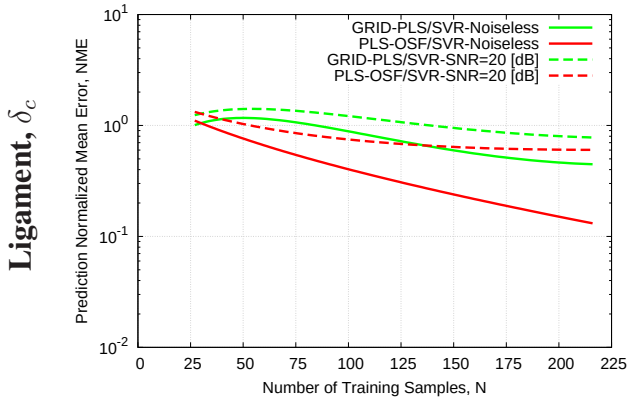


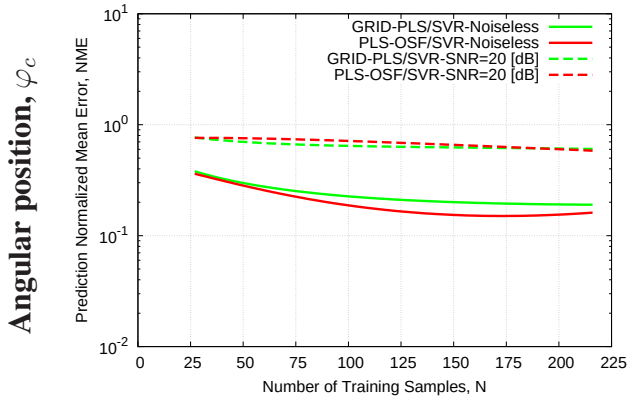
Figure 7.11: Numerical assessment: (crack length, $l_c = [3,\ 10]$ mm, ligament, $\delta_c = [0,\ 4]$ mm, and angular position, $\varphi_c = [0,\ 90]$ deg.) - PLS-OSF/SVR prediction results are shown in terms of prediction error (Normalized Mean Error, NME) vs. SNR for crack (a) length l_c, (b) ligament δ_c and (c) angular position φ_c estimation for $N = 216$, $J = 20$, for test set at noiseless and $SNR = 10,\ 20,\ 30,\ 40$ dB with $M = 1000$ through ECT, UT and ECT-UT.



(a)



(b)



(c)

Figure 7.12: Numerical assessment: (crack length, $l_c = [3,\ 10]$ mm, ligament, $\delta_c = [0,\ 4]$ mm, and angular position, $\varphi_c = [0,\ 90]$ deg.) - PLS-OSF/SVR vs. GRID-PLS/SVR prediction results are shown in terms of prediction error (Normalized Mean Error, NME) vs. number of training samples N for crack (a) length l_c, (b) ligament δ_c and (c) angular position φ_c estimation for $J = 20$, $SNR = 20$ dB and $Noiseless$ test set with $M = 1000$ through ECT-UT.

Until now we have seen the preliminary results of inversion performance by applying ECT-UT data fusion method through PLS-OSF/SVR approach. Multi-physics data fusion can also be treated by GRID-PLS sampling strategy. From Fig. 7.12, we can see that PLS-OSF/SVR has shown equal or higher prediction accuracy than GRID-PLS/SVR approach while we are considering ECT-UT data fusion method. By increasing N adaptively, PLS-OSF/SVR strategy improves crack ligament δ_c estimation compared to GRID-PLS/SVR strategy (Fig. 7.12 (b)) which also indicates better estimation accuracy is obtained by PLS-OSF/SVR. In this test case, the number of parameters to be estimated is $Q = 3$, therefore, even at $N = 216$, GRID-PLS/SVR can show good prediction accuracy. As we have discussed in Sec. 3.3.3, GRID-PLS sampling can be limited for dealing with higher parametric inversion. With increasing the number of parameters (i.e., Q) to be inverted, the number of total samples N increases exponentially. We need to consider also for dealing with $F = 9963$ features before applying PLS feature extraction. This requires higher simulation time in offline phase and limits prediction accuracy. As we have shown in Chapter 6, applying PLS-OSF can be a suitable solution for higher Q to have better prediction accuracy for lower N.

7.7 Conclusions

In this chapter, we have shown the preliminary results obtained with a multi-physics data fusion strategy for crack dimension and position estimation. PLS-OSF/SVR strategy has been applied for solving the complex NDE problem by utilizing ECT signals, UT signals and multi-physics data fusion (i.e., ECT-UT). ECT-UT shows better prediction accuracy than ECT and UT methods separately for crack dimension and position estimation both for noisy and noiseless test set. By combining two different NDT methods, we first retrieved the variation of actual ECT and UT signals for changing crack parameters. Applying PLS feature extraction retrieves most significant information from the actual combined ECT-UT data set. ECT signals are more corrupted for imposing noise than UT signals. By combining ECT and UT signals we can be benefited by both ECT and UT signals which improves the learning ability of the learning algorithm (Fig. 7.8 - Fig 7.11). With increasing number of training samples, N prediction accuracy increases for all NDT methods (i.e., ECT, UT and ECT-UT) cases. However, ECT-UT data fusion improves inversion performance and shows robustness in presence of noise. Finally, inversion performance by PLS-OSF/SVR strategy has been compared with non-adaptive GRID-PLS/SVR approach on ECT-UT data. PLS-OSF/SVR shows almost comparable inversion results than GRID-PLS/SVR for l_c and φ_c estimation, while much improved prediction accuracy for δ_c estimation (Fig. 7.12). One should keep in mind that, due to the curse of dimensionality, increasing the number of parameters to be estimated will greatly impact on the performance of GRID-PLS/SVR which will be worsen compared to PLS-OSF/SVR.

Chapter 8

Conclusions

The aim of this thesis was to develop a robust inversion schema in order to build real time diagnosis system in nondestructive eddy current testing (ECT). Learning by Examples (LBE) methodology has been adopted for crack(s)/defect(s) localization, characterization problems within the inspected medium. Though LBE has been mostly applied to deal with ECT signals, the developed strategies have been extended to other physics by applying it to ultrasound testing (UT) data too. In the latter part of the work, both ECT and UT signals were combined in order to enhance the inversion capabilities.

8.1 Concluding remarks

In Chapter 1, we have provided an overview on the different NDT methods that are used in industries for inspecting defect(s)/flaw(s) within the inspected medium. In the same chapter, a review on the state of art of the application of ECT NDT was emphasized. Within the context of ECT, forward and inverse problems have been outlined. Among different inverse problem solutions, non-iterative LBE strategy has been briefly introduced which is followed by the goal of the thesis.
In Chapter 2, LBE has been introduced for solving an inverse problem within the context of ECT NDT. In general, LBE is a two phases approach, where a fast and accurate training model is built in offline phase based on an optimal training set made of input-output pairs by regression/prediction techniques. The developed (trained) model is then used to predict the output associated to an unknown test sample during the online phase. Within the framework of LBE, the necessities of feature selection and extraction steps were discussed for reducing higher dimensional ECT signal space to a much lower dimensional extracted feature space. In the same chapter, a discussion on the exhaustive representation of extracted feature space for treating most appropriate and informative contents from actual ECT signal space has been done. Finally, the choice of suitable prediction technique was explained in order to obtain better inversion performance.
Application of different sampling strategies for obtaining optimal training set within

the scope of LBE are dilated in Chapter 3. Standard sampling approaches such as full factorial GRID, Latin Hyper-cube Sampling (LHS), Output Space Filling (OSF) sampling have been enhanced by applying Partial Least Squares (PLS) feature extraction. The obtained sampling strategies labelled as GRID-PLS, LHS-PLS, PLS-OSF have been analyzed through highlighting pros and cons of each of them. By applying feature extraction, we can not only reduce the actual NDT signal space, but also improve the learning ability of the prediction technique (e.g., support vector regressor (SVR)). In Chapter 4, after obtaining optimal training sets through the sampling strategies detailed in Chapter 3, the procedure to obtain accurate inverse model by the prediction algorithm (i.e., SVR) has been introduced. Afterward, the crucial choice to obtain optimal calibration parameters of PLS and SVR have been elaborated with error metric and signal to noise ratio formulation.
The proposed LBE strategies are applied for quasi real time inversion on different NDT problems from Chapter 5 to Chapter 7. In Chapter 5, we have assessed the performance of PLS-OSF/SVR strategy within the framework of LBE for real-time flaw characterization based on ECT signals. The prediction accuracy for crack dimension estimation by applying PLS-OSF/SVR has been compared with GRID/SVR through numerical simulations as well as experimental data. By applying PLS-OSF sampling strategy, an optimal training set can be obtained which increases the learning capability of the employed algorithm (i.e., SVR). PLS-OSF/SVR exhibits higher prediction accuracy than GRID/SVR on synthetic noisy test set as well as on experimental data.
PLS-OSF/SVR has been applied for multi-cracks characterization and localization problem by applying ECT signals in Chapter 6. The goal of this chapter was to analyze and show how LBE approach scales with the problem of high dimensionality. Standard high dimensional parametric problem implies the generation of huge training set which is also time demanding. By applying PLS-OSF/SVR, higher prediction accuracy has been shown for performing inversion on the studied problem having six parameters.
One of the significant advantages of PLS-OSF strategy is to project the high dimensional ECT signals to a different and much lower dimensional space through PLS feature extraction. Therefore, PLS-OSF/SVR approach has been applied to other NDT methods. In this chapter 7, PLS-OSF/SVR has been used for both ECT and UT signals for crack dimension and position estimation of a complex NDT problem. A data fusion of ECT and UT signals (i.e., ECT-UT) has been shown to be able to increase inversion performance for the treated problem. By combining two different NDT methods, we first retrieved the variation of actual ECT and UT signals for changing crack parameters. Also in the case of combined NDT signals, PLS feature extraction was able to retrieve the most significant information from the actual combined ECT-UT data set. By combining ECT and UT signals we can be benefited of both ECT and UT signals informative contents which improves the prediction capability of the learning algorithm. The preliminary studies on the proposed data fusion approach have shown promising results. In fact, ECT-UT shows better prediction accuracy than ECT and UT methods separately for crack dimension and position estimation, which

is also robust on synthetic noisy test set.

8.2 Perspectives and future works

In this thesis, we have built an inversion schema by using LBE methodology. PLS feature extraction technique is applied for extracting most suitable features from the actual NDT signal space. Other feature extraction techniques such as principle component analysis, linear discriminant analysis, canonical correlation analysis and their nonlinear extensions (i.e., kernel version) can also be applied within the framework of LBE. Depending on the formulation of each of these methods, inversion performance may be varied. This research work can be extended to other NDT methods (e.g., ultrasound, thermography etc.) in order to evaluate the performance of these approaches.
During this thesis we have experienced that different learning algorithms can be useful for estimating different crack parameters by means of optimal prediction accuracy. Therefore, adaptive boosting (known as AdaBoost) can be explored within the framework of LBE by combining different learning algorithms in order to improve the overall defect parameters estimation.
Dealing with real life NDT problems, we need to consider the impacts of various sources of experimental noise which can be dissimilar to AWGN noise. Concerning these unknown sources of noise, the learning capability of the inverse model can be enhanced by employing experimental data with the simulated data for training model development. Therefore, the robustness of the inversion schema would be improved to deal with unknown experimental noisy data.
For what concerns UT signals in Chapter 7, we have treated maximum amplitude of the UT signals collected from each single inspection point (i.e., C-scan). It is believed that higher information contents can be retrieved from the finer domain of UT signals. Unfortunately, such UT signals have much higher dimensional data set than a simple C-scan. Thus, the signals should be ad-hoc pre-processed before applying feature extraction techniques. More research work should be done to find out appropriate solutions dealing with higher dimensional UT data set. In [73], authors have utilized SVR and PCA for crack characterization using ultrasonic arrays. LBE strategy can be extended to deal with such type of UT solution.
We have shown multi-physics data fusion by combining ECT and UT signals for enhancing inversion performance based on simulation data. Extensive research would be emphasized through experimental validation and solving different NDT problems. Consequently, alternative methods would be investigated for combining the extracted features from different NDT data. Similarly, the developed LBE strategies can be adopted to perform data fusion with other NDT methods such as thermography.
One of the main goals of this thesis was to develop real time inversion strategy that can be applied on real industrial applications. After having an accurate training model obtained through LBE, it can be applied in portable devices (e.g., inspecting equipments) to provide real time prediction functionalities. That is, feature extraction and inverse model results can be fully described in a very compact form (e.g., through ma-

trix and vectors) of reasonable size. Thus storage limitation of portable devices would not be a big issue which translates up to a certain level into the possibility to obtain very fast predictions directly via embedded electronics.

Bibliography

[1] P. O. Moore and S. S. Udpa, *Nondestructive Testing Handbook: Electromagnetic testing*, vol. 5. American Society for Nondestructive Testing, 2004.

[2] J. G. Martin, J. G. Gil, and E. V. Sanchez, "Non-destructive techniques based on eddy current testing," *Sensors*, vol. 11, no. 3, pp. 2525–2565, 2011.

[3] Y. Deng and X. Liu, "Electromagnetic imaging methods for nondestructive evaluation applications," *Sensors*, vol. 11, no. 12, pp. 11774–11808, 2011.

[4] R. F. Anastasi and E. I. Madaras, "Terahertz NDE for Under Paint Corrosion Detection and Evaluation," in *Review of Progress in Quantitative Nondestructive Evaluation Volume 25* (D. O. Thompson and D. E. Chimenti, eds.), vol. 820 of *American Institute of Physics Conference Series*, pp. 515–522, Mar. 2006.

[5] K. J. Karwoski, *Circumferential cracking of steam generator tubes*. NUREG-1604, 1997.

[6] D. Mercier, J. Lesage, X. Decoopman, and D. Chicot, "Eddy currents and hardness testing for evaluation of steel decarburizing," *NDT & E International*, vol. 39, no. 8, pp. 652 – 660, 2006.

[7] L. B. Pedersen, K. Å. Magnusson, and Y. Zhengsheng, "Eddy current testing of thin layers using co-planar coils," *Research in Nondestructive Evaluation*, vol. 12, pp. 53–64, Mar 2000.

[8] R. Pohl, A. Erhard, H.-J. Montag, H.-M. Thomas, and H. WÃ$\frac{1}{4}$stenberg, "Ndt techniques for railroad wheel and gauge corner inspection," *NDT & E International*, vol. 37, no. 2, pp. 89 – 94, 2004.

[9] J. W. Wilson, G. Y. Tian, and S. Barrans, "Residual magnetic field sensing for stress measurement," *Sensors and Actuators A: Physical*, vol. 135, no. 2, pp. 381 – 387, 2007.

[10] H. C. Schonekess, W. Ricken, and W. J. Becker, "Improved multi-sensor for force measurement on pre-stressed steel cables by means of eddy current technique," in *Proceedings of IEEE Sensors, 2004.*, pp. 260–263 vol.1, Oct 2004.

[11] R. Grimberg, "Electromagnetic nondestructive evaluation: Present and future," *Journal of Mechanical Engineering*, vol. 57, no. 3, pp. 204–217, 2011.

[12] Comsol, "COMSOL: Multiphysics Modeling Software." https://www.comsol.com/. [Online; accessed Nov. 2017].

[13] G2ELab, "Flux: software for electromagnetic and thermal simulations." http://www.cedrat.com/software/flux/. [Online; accessed Nov. 2017].

[14] CST, "CST: EMC STUDIO." https://www.cst.com/products/cstemcs/solvers/. [Online; accessed Nov. 2017].

[15] CEA-LIST, "CIVA: Simulation and Analysis for NDT." http://www-civa.cea.fr/en/. [Online; accessed Nov. 2017].

[16] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. New Haven: Yale University Press, 1923.

[17] S. Caorsi, A. Massa, and M. Pastorino, "Electromagnetic imaging of penetrable configurations by means of a ga/cg method," *Measurement*, vol. 36, no. 3, pp. 271 – 278, 2004. Imaging Measurement Systems.

[18] P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, and A. Massa, "Evolutionary optimization as applied to inverse scattering problems," *Inverse Problems*, vol. 25, no. 12, p. 123003, 2009.

[19] M. Salucci, L. Poli, N. Anselmi, and A. Massa, "Multifrequency particle swarm optimization for enhanced multiresolution gpr microwave imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 1305–1317, March 2017.

[20] N. Anselmi, M. Salucci, G. Oliveri, and A. Massa, "Wavelet-based compressive imaging of sparse targets," *IEEE Transactions on Antennas and Propagation*, vol. 63, pp. 4889–4900, Nov 2015.

[21] L. Poli, G. Oliveri, P. Rocca, and A. Massa, "Model based inversion algorithms based on bayesian compressive sensing," in *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, pp. 492–495, July 2011.

[22] M. Salucci, G. Oliveri, and A. Massa, "Gpr prospecting through an inverse-scattering frequency-hopping multifocusing approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 6573–6592, Dec 2015.

[23] T. Henriksson, M. Lambert, and D.Lesselier, "Non-iterative music-type algorithm for eddy-currentusa nondestructive evaluation of metal plates," *Studies in Applied Electromagnetics and Mechanics: Elec- tromagnetic Nondestructive Evaluation (XIV)*, vol. 35, pp. 22–29, 2011.

[24] A. Tamburrino, A. Vento, S. Ventre, and A. Maffucci, "Monotonicity imaging method for flaw detection in aeronautical applications studies in applied," *Electromagnetics and Mechanics: Electromagnet ic Non- destructive Evaluation (XIX)*, vol. 41, pp. 284–292, 2016.

[25] S. Ahmed, R. Miorelli, M. Salucci, and A. Massa, "Real-time flaw characterization through learning- by-examples technique s: A comparative study applied to ect," *Studies in Applied Electromagnetics and Mechanics: Electromagnetic Nondestructive Evaluation (XX)*, vol. 42, pp. 228–235, 2017.

[26] S. Ahmed, R. Miorelli, P. Calmon, N. Anselmi, and M. Salucci, "Real time flaw detection and characterization in tube through partial least squares and svr: Application to eddy current testing," in *44th Annual Review of Progress in QNDE*, Jul. 2017. USA.

[27] S. Ahmed, M. Salucci, R. Miorelli, N. Anselmi, G. Oliveri, P. Calmon, C. Reboud, and A. Massa, "Real- time groove characterization combining partial least squares and svr strategies: Application to eddy current testing," *Journal of Physics: Conference Series*, vol. 904, 2017.

[28] S. Ahmed, R. Miorelli, C. Reboud, P. Calmon, N. Anselmi, and M. Salucci, "Fast characterization of multiple cracks in conductive media based on adaptive feature extraction and svr," in *22nd International Workshop on Electromagnetic Nondestructive Evaluation*, Sept. 2017. France.

[29] M. Salucci, S. Ahmed, and A. Massa, "An adaptive learning-by-examples strategy for efficient eddy current testing of conductive structures," in *10th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–4, Apr. 2016.

[30] M. Salucci, S. Ahmed, N. Anselmi, G. Oliveri, P. Calmon, R. Miorelli, C. Reboud, and A. Massa, "Real-time crack characterization in conductive tubes through an adaptive partial least squares approach," in *IEEE International Symposium on Antennas and Propagation USNC/URSI National Radio Science Meeting*, pp. 21–22, Jul. 2017.

[31] M. Salucci, N. Anselmi, G. Oliveri, P. Calmon, R. Miorelli, C. Reboud, and A. Massa, "Real-time ndt-nde through an innovative adaptive partial least squares svr inversion approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, pp. 6818–6832, Nov 2016.

[32] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[33] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer Series in Statistics, New York, NY, USA: Springer New York Inc., 2001.

[34] S. Theodoridis, *Machine Learning: A Bayesian and Optimization Perspective*. Academic Press, 1st ed., 2015.

[35] L. Udpa and S. Udpa, "Eddy current defect characterization using neural networks: Materials evaluation," *NDT International*, vol. 23, no. 6, p. 358, 1990.

[36] B. Rao, B. Raj, T. Jayakumar, and P. Kalyanasundaram, "An artificial neural network for eddy current testing of austenitic stainless steel welds," *NDT & E International*, vol. 35, no. 6, pp. 393 – 398, 2002.

[37] M. Cacciola, F. L. Foresta, F. C. Morabito, and M. Versaci, "Advanced use of soft computing and eddy current test to evaluate mechanical integrity of metallic plates," *NDT & E International*, vol. 40, no. 5, pp. 357 – 362, 2007.

[38] Y. L. Diraison, P.-Y. Joubert, and D. Placko, "Characterization of subsurface defects in aeronautical riveted lap-joints using multi-frequency eddy current imaging," *NDT & E International*, vol. 42, no. 2, pp. 133 – 140, 2009.

[39] J. Kim, G. Yang, L. Udpa, and S. Udpa, "Classification of pulsed eddy current gmr data on aircraft structures," *NDT & E International*, vol. 43, no. 2, pp. 141 – 144, 2010.

[40] 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 International*, vol. 67, no. Supplement C, pp. 46 – 54, 2014.

[41] D. Pasadas, A. Ribeiro, T. Rocha, and H. Ramos, "2d surface defect images applying tikhonov regularized inversion and ect," *NDT & E International*, vol. 80, no. Supplement C, pp. 48 – 57, 2016.

[42] H. A. Sabbagh, "Splines and their reciprocal-bases in volume-integral equations," *IEEE Transactions on Magnetics*, vol. 28, pp. 1138–1141, Mar 1992.

[43] D. Lesselier and A. Razek, "Eddy current scattering and inverse scattering, Green's integral and variational formulations.," in *Scattering. Scattering and Inverse Scattering in Pure and Applied Science: Part 1 – Scattering of Waves by Macroscopic Targets,* (E. Pike and P. C. Sabatier, eds.), pp. 486–507, Academic Press, 2002.

[44] J. Arenas-García, K. B. Petersen, G. Camps-Valls, and L. K. Hansen, "Kernel multivariate analysis framework for supervised subspace learning: A tutorial on linear and kernel multivariate methods," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 16–29, 2013.

[45] S. Bilicz, M. Lambert, and S. Gyimothy, "Kriging-based generation of optimal databases as forward and inverse surrogate models," *Inverse Problems*, vol. 26, no. 7, p. 074012, 2010.

[46] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, Sep 1995.

[47] C. Saunders, A. Gammerman, and V. Vovk, "Ridge regression learning algorithm in dual variables," in *Proceedings of the Fifteenth International Conference on Machine Learning*, ICML '98, (San Francisco, CA, USA), pp. 515–521, Morgan Kaufmann Publishers Inc., 1998.

[48] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sept. 2001.

[49] Sui, Yun-Kang, Li, Shan-Po, and Guo, Ying-Qiao, "An efficient global optimization algorithm based on augmented radial basis function," *Int. J. Simul. Multidisci. Des. Optim.*, vol. 2, no. 1, pp. 49–55, 2008.

[50] E. Bermani, A. Boni, S. Caorsi, and A. Massa, "An innovative real-time technique for buried object detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 927–931, April 2003.

[51] S. Caorsi, D. Anguita, E. Bermani, A. Boni, M. Donelli, and A. Massa, "A comparative study of nn and svm-based electromagnetic inverse scattering approaches to on-line detection of buried objects," *Applied Computational Electromagnetics Society Journal*, vol. 18, pp. 1–11, July 2003.

[52] D. C. Montgomery, *Design and Analysis of Experiments*. John Wiley & Sons, 2006.

[53] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 42, pp. 55–61, Feb. 2000.

[54] S. H. Paskov and J. F. Traub, "Faster valuation of financial derivatives," working papers, Santa Fe Institute, 1995.

[55] K. P. F.R.S., "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901.

[56] H. Wold, "Soft modeling: the basic design and some extensions," *Systems under indirect observation*, vol. 2, pp. 589–591, 1982.

[57] K. Fukunaga, *Introduction to Statistical Pattern Recognition (2nd Ed.)*. San Diego, CA, USA: Academic Press Professional, Inc., 1990.

[58] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[59] S. de Jong, "Simpls: An alternative approach to partial least squares regression," *Chemometrics and Intelligent Laboratory Systems*, vol. 18, no. 3, pp. 251 – 263, 1993.

[60] V. Vapnik and A. Lerner, "Pattern Recognition using Generalized Portrait Method," *Automation and Remote Control*, vol. 24, 1963.

[61] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, second ed., November 1999.

[62] E. Bermani, A. Boni, A. Kerhet, and A. Massa, "Kernels evaluation of SVM-based estimators for inverse scattering problems," *Progress In Electromagnetics Research*, vol. 53, pp. 167–188, 2005.

[63] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[64] S. Wold, M. Sjostrom, and L. Eriksson, "Pls-regression: a basic tool of chemometrics," *Chemometrics and Intelligent Laboratory Systems*, vol. 58, pp. 109–130, 2001.

[65] Q. Mistral, T. Charret, and J. Martens, "Array probe implementation (smx) on edf steam generator tubes," in *10th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurizeds*, pp. 329–333, Oct. 2013.

[66] R. Miorelli, C. Reboud, D. Lesselier, and T. Theodoulidis, "Eddy current modeling of narrow cracks in planar-layered metal structures," *IEEE Transactions on Magnetics*, vol. 48, pp. 2551–2559, Oct 2012.

[67] R. Miorelli, C. Reboud, T. Theodoulidis, N. Poulakis, and D. Lesselier, "Efficient modeling of ect signals for realistic cracks in layered half-space," *IEEE Transactions on Magnetics*, vol. 49, pp. 2886–2892, June 2013.

[68] Wfndec, "World Federation of NDE Centers, 2008 Eddy Current Benchmarks." https://www.wfndec.org/. [Online; accessed Nov. 2017].

[69] D. Hopkins, M. Datuin, J. Aldrin, M. Warchol, L. Warchol, and D. Forsyth, "Localization and characterization of fatigue cracks around fastener holes using spherically focused ultrasonic probes," *AIP Conference Proceedings*, vol. 1806, no. 1, p. 090007, 2017.

[70] M. Darmon, S. Chatillon, S. Mahaut, P. Calmon, L. J. Fradkin, and V. Zernov, "Recent advances in semi-analytical scattering models for ndt simulation," *Journal of Physics: Conference Series*, vol. 269, no. 1, p. 012013, 2011.

[71] R. Miorelli, C. Reboud, T. Theodoulidis, J. Martinos, N. Poulakis, and D. Lesselier, "Coupled approach vim-bem for efficient modeling of ect signal due to

narrow cracks and volumetric flaws in planar layered media," *NDT & E International*, vol. 62, no. Supplement C, pp. 178 – 183, 2014.

[72] A. Kamta Djakou, *Modeling of diffraction effects for specimen echoes simulations in ultrasonic Non-Destructive Testing (NDT)*. Theses, Université du Maine, Sept. 2016.

[73] L. Bai, A. Velichko, and B. W. Drinkwater, "Characterization of defects using ultrasonic arrays: a dynamic classifier approach," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, pp. 2146–2160, December 2015.

[74] M. Darmon, N. Leymarie, S. Chatillon, and S. Mahaut, *Modelling of scattering of ultrasounds by flaws for NDT*, pp. 61–71. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[75] J. Achenbach, A. Gautesen, and H. McMaken, *Ray Methods for Waves in Elastic Solids: with Applications to Scattering by Cracks*. Pitman Advanced Publishing Program, 1982.

[76] A. K. Djakou, M. Darmon, L. Fradkin, and C. Potel, "The uniform geometrical theory of diffraction for elastodynamics: Plane wave scattering from a half-plane," *The Journal of the Acoustical Society of America*, vol. 138, no. 5, pp. 3272–3281, 2015.

[77] V. Zernov, L. Fradkin, and M. Darmon, "A refinement of the kirchhoff approximation to the scattered elastic fields," *Ultrasonics*, vol. 52, no. 7, pp. 830 – 835, 2012.

[78] M. Darmon, V. Dorval, A. K. Djakou, L. Fradkin, and S. Chatillon, "A system model for ultrasonic ndt based on the physical theory of diffraction (ptd)," *Ultrasonics*, vol. 64, no. Supplement C, pp. 115 – 127, 2016.

Appendix A

Ultrasound Testing Formulation

A.1 Problem definition

An aluminum plate alloy has been investigated by Ultrasonic Testing (UT) approach. The plate includes a fastener hole and a crack of length (l_c), ligament (δ_c), and angular distance (φ_c).

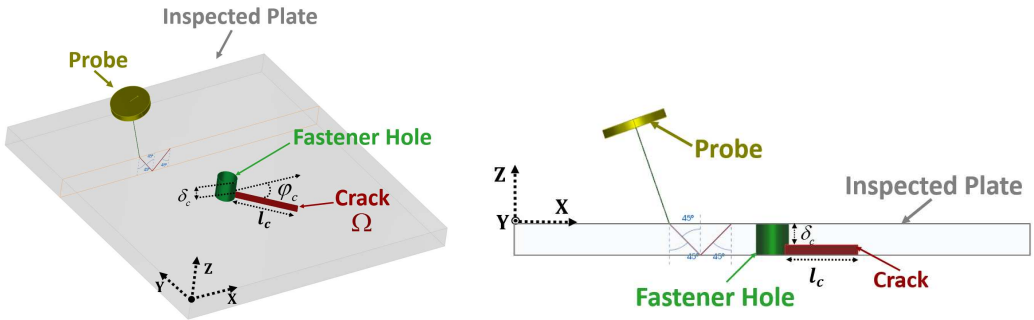


Figure A.1: Geometry of the problem of plate structure (a) Top view and (b) Front view.

Figure A.1 shows the geometrical view of the problem that is used for UT measurement. The plate is inspected by a probe (a single transmitting/receiving probe) that generates incident ray at position $\mathbf{r}_k = (x_k, y_k, z_k)$ (considering the associate Cartesian coordinate system (x, y, z) while it is moving during measurement phase). The probe is positioned in way that the incident angle of the transmitting UT ray generates 45 deg shear waves (i.e., secondary waves/ S-waves) in the aluminum plate. The next section describes the mathematical formulation of UT signal.

A.2 Theoretical background of Ultrasonic Testing (UT):

The ultrasonic simulation tools in CIVA allow to fully predict a real ultrasonic inspection in a various range of applications which requires the computation of the beam propagated, as well as its interaction with flaws. The beam propagation model is based upon a semi-analytical method which calculates the impulse response of the probe inside the component, assuming individual source points distributed over the radiating surface of the probe. Each elementary source point contribution of the probe toward the computation point is therefore evaluated using a so-called pencil method applied to elastodynamics. This model allows to compute the ultrasonic field in the component for wedge coupled or immersed probes of arbitrary shapes, for monolithic or phased-array transducers. Firstly developed for homogeneous and isotropic materials, it has been extended to deal with anisotropic and heterogeneous cases. The CIVA tools are also dedicated to the prediction of flaw responses. The flaw scattering model is based upon various approximations depending on the inspection technique and the nature of flaw (void, solid inclusion). These approximations will be detailed in the following part. The signal received by the probe is obtained by summing up all the scattered contributions [74]. The amplitude of one contribution is proportional to the probe sensitivity at the location of the source. The model assumes the transmission-reception reciprocity for the transducers so that this sensibility is directly deduced from the transmitted field computation. To deal with the wave scattering, the different classical approximations are applied depending of the scatterer. In the case of specimen boundaries, calibration reflectors (side drilled hole, flat bottom hole, etc...) or large voids, the high frequency Kirchhoff Approximation (KA) is used. In the case of cracks, the KA and/or Geometrical Theory of Diffraction (GTD) can be used depending on the considered echo mechanism (tip diffraction; specular reflection). A generic model based on Physical Theory of Diffraction (PTD) which mixes KA and GTD is the more effective model to simulate the UT response signals.

A.2.1 Geometrical Elastodynamics

In the CIVA platform software, the computation of the field radiated by a transducer is based on Geometrical Elastodynamics (GE) [75] and more precisely the pencil method. GE is the simplest way to model propagation and scattering of elastic waves. It is a ray method which just considers incident rays, refracted and reflected rays by interfaces scatterers. Along a ray, the characteristics of the GE, field (amplitude, phase, propagation and polarisation directions) are determined using Snell-Descartes law of reflection and the energy conservation. Where, a pencil is a tube of rays which emanate from a source point. The axial ray of this tube respects Snell-Descartes law at interfaces. The pencil is then parameterized by this axial ray and by the cone aperture angle which is the solid angle covered by the pencil at the point source. The pencil then respects refraction and reflection laws at interfaces and is characterized by its time of flight, its amplitude, its mean polarisation and the wave vector directions.

When the pencil propagates, the energy through its cross section remains constant. Therefore, the divergence factor of the pencil corresponds to the one of a spherical wave emitted by a point source. The pencil method for field computation is also used for calculating the specimen echoes in CIVA. The corresponding model of specimen echoes calculation is called "specular model" since it models only the specular reflection on the specimen boundaries. The "specular model" in CIVA for the specimen echoes simulation is suitable to model specular reflection on the specimen boundaries but it presents discontinuities at the incident and reflected shadow boundaries and is not valid at caustics. These limitations of the GE are overcome by the Kirchhoff approximation.

A.2.2 Kirchhoff Approximation (KA)

Kirchhoff approximation is a high frequency approximation as the GE. This approximation is used to calculate the unknown Crack Opening Displacement COD (i.e. the jump in the total field displacement across the crack). The COD is a factor which is contained in the integrand of the exact integral representation of the scattered field at an observation point (Eq. (A.1)). For a stress-free crack, in elastodynamics, this exact integral representation of the scattered field also called the Rayleigh-Sommerfeld integral which can be expressed from [72]

$$u_{\wp}^{Scat(Kir)}(\mathbf{r}) = \int_S [u_i(\mathbf{r}')]\Sigma_{ij}^{(\wp)}(\mathbf{r},\, \mathbf{r}')\eta_j(\mathbf{r}')d^2\mathbf{r}'. \tag{A.1}$$

where $u_{\wp}^{Scat(Kir)}(\mathbf{r})$ is the $\wp$-th coordinate of the displacement scattered field at the observation vector $\mathbf{r}$, of the surface S. $\Sigma_{ij}^{(\wp)}(\mathbf{r},\, \mathbf{r}')$ is the (ij) component of the Green's stress tensor $\mathbf{\Sigma}^{(\wp)}(\mathbf{r},\, \mathbf{r}')$ (the stress produced at $\mathbf{r}$ by an unit traction acting along the $\wp$-axis at a point $\mathbf{r}'$ on the surface S). η is the outward normal to S and $[u(\mathbf{r}')]$ is the COD. In far field, the integral of Eq. (A.1) is reduced to

$$u_{\wp}^{Scat(Kir)}(\mathbf{r}) = \int_S u_\alpha(\mathbf{r}')D_\beta^\alpha(\upsilon^\alpha, \upsilon_\beta)\frac{e^{i\lambda_\beta S_\beta}}{S_\beta}d^2\mathbf{r}' \tag{A.2}$$

where, $\alpha = L,\, TV,\, or\, TH$ (Longitudinal, Transverse Vertical or Transverse Horizontal, respectively) incident type wave vector and $\beta = L,\, TV,\, or\, TH$ is the scatter type wave vector. Therefore, $u_\alpha(\mathbf{r}')$ is the incident field amplitude at a point on the scatter surface. D_β^α is the Kirchhoff diffraction coefficient which is dependent on the incident propagation vector υ^α and on the observation propagation vector υ_β. S_β is the distance from the integration point on the scatter surface to the observation point and λ_β is the wave number. KA describes particularly well specular reflection since it is based on GE. It also models edge diffracted wave contrary to GE, and hence gives rise to a continuous field in the whole space, even at caustics.

For NDT applications echo calculation, observation points are points on the receiver transducer crystal. KA implementation in CIVA is thus based on the plane

wave approximation for the incident field and on the field reciprocity principle for calculating the echo generated by the scatterer surface to the receiver transducer. Indeed, the field radiated by the emitter transducer at a point on the scatterer surface $u^{\alpha}(\mathbf{r}')$ is calculated using the pencil method. In far-field, this field can be approximated at each point on the scatterer surface by a local plane wave by extracting information from the field computation as its amplitude, its time of flight, its mean polarization and the wave vector directions at a given observation point. However, the amplitude of the diffracted waves from a crack edge predicted by KA is not accurate, more specifically for transversal waves. GTD is well known for describing more precisely the edge diffraction phenomena.

A.2.3 Geometrical Theory of Diffraction (GTD)

The geometrical theory of diffraction extends the GE by adding to it diffraction phenomena. It postulates the existence of rays diffracted from the structure irregularities such as edges or tips, in addition to the incident and reflected rays of GE. It also gives a recipe for calculating the amplitudes carried by these diffracted rays. For an elastic incident plane wave of type $\alpha = L, TV, or\, TH$ on a half-plane or wedge, the total displacement field is expressed as [72, 76]

$$\mathbf{u}^{tot}(\mathbf{r}) = \mathbf{u}^{GE}(\mathbf{r}) + \sum_{\beta} \mathbf{u}_{\beta}^{diff}(\mathbf{r})$$

$\mathbf{u}^{GE}$ is the GE displacement field and $\mathbf{u}_{\beta}^{diff}$ is the diffracted displacement field of the scattered wave of type β. The GTD permits to approximate the diffracted displacement field $\mathbf{u}_{\beta}^{diff}$ by

$$\mathbf{u}_{\beta}^{diff(GTD)}(\mathbf{r}) = \sum_{\beta} u_{\alpha}(\mathbf{r}_{\beta}^{\alpha}) D_{\beta}^{\alpha(GTD)} \frac{e^{i\lambda_{\beta} S_{\beta}}}{\sqrt{\lambda_{\beta} L_{\beta}}} \mathbf{e}_{\beta}(\mathbf{r})$$

where

$$u_{\alpha}(\mathbf{r}_{\beta}^{\alpha}) = \mathbf{u}^{\alpha}(\mathbf{r}_{\beta}^{\alpha}).\mathbf{d}^{\alpha}$$

$\mathbf{u}^{\alpha}(\mathbf{r}_{\beta}^{\alpha})$ is the incident field at the diffraction point $\mathbf{r}_{\beta}^{\alpha}$ and $\mathbf{d}^{\alpha}$ is its polarisation vector. $\beta = L, TV, or\, TH$ is the scattered wave mode, λ_{β} its wave number and $\mathbf{e}_{\beta}$ is its polarisation. S_{β} is the distance between the diffraction point $\mathbf{r}_{\beta}^{\alpha}$ and the observation point $\mathbf{r}$. L_{β} is a parameter distance.

We can find diffraction contribution in the scattered field by assuming incident field $u_{\alpha}(\mathbf{r}_{\beta}^{\alpha}) \approx 1$ from [77],

$$\mathbf{u}_{\beta}^{diff(GTD)}(\mathbf{r}) = \sum_{\beta} D_{\beta}^{\alpha(GTD)} \frac{e^{i\lambda_{\beta} S_{\beta}}}{\sqrt{\lambda_{\beta} L_{\beta}}} \mathbf{e}_{\beta}(\mathbf{r})$$

Thus the resultant scattered field can be expressed as

$$\mathbf{u}^{Scat(GTD)}(\mathbf{r}) = \mathbf{u}^{GE}(\mathbf{r}) + \mathbf{u}^{Rayleigh}(\mathbf{r}) + \mathbf{u}_{\beta}^{diff(GTD)}(\mathbf{r})$$

where the geometrical elastrodynamics field $\mathbf{u}^{GE}(\mathbf{r})$ and Rayleigh field $\mathbf{u}^{Rayleigh}(\mathbf{r})$ [75] comprise reflected waves and surface waves, respectively.

GTD gives a good analytical description of diffracted waves. However, as GE, it fails at caustic regions and also in the transition zones (also called penumbrae) surrounding shadow boundaries, where edge diffracted waves interfere with incident or reflected waves. Indeed the GTD solution is infinite at such boundaries. For this last reason, GTD is said to be non-uniform. To overcome this shortcoming of GTD, an uniform extension of GTD named Physical Theory of Diffraction (PTD) has been developed recently in elastodynamics.

A.2.4 The Physical Theory of Diffraction (PTD)

The physical theory of diffraction combines GTD and KA to overcome the limitations of both models. The PTD consists in correcting the KA contribution for edge diffracted waves thanks to the GTD. This correction is done by adding to the KA field a corrective term which is the difference between the wave amplitudes diffracted by the edge given by GTD and KA. For an incident plane wave of type $\alpha = L,\, TV,\, or\, TH$, on an obstacle, the total PTD displacement field is expressed as [72, 78]

$$\mathbf{u}^{tot(PTD)}(\mathbf{r}) = \mathbf{u}^{\alpha}(\mathbf{r}) + \mathbf{u}^{diff(PTD)}(\mathbf{r})$$

where,

$$\mathbf{u}^{diff(PTD)}(\mathbf{r}) = \sum_{\beta} \left[\mathbf{u}_{\beta}^{\alpha(KA)}(\mathbf{r}) + u_{\alpha}(\mathbf{r}_{\beta}^{\alpha}) \left(D_{\beta}^{\alpha(GTD)}(\mathbf{r}) - D_{\beta}^{\alpha(KA)}(\mathbf{r}) \right) \frac{e^{i\lambda_{\beta} S_{\beta}}}{\sqrt{\lambda_{\beta} L_{\beta}}} \mathbf{e}_{\beta}(\mathbf{r}) \right] \tag{A.3}$$

In A.3, β is the type of the scattered wave and the scattered field $\mathbf{u}_{\beta}^{\alpha(KA)}(\mathbf{r})$ is obtained by Eq. (A.1) with the COD being calculated with GE. $D_{\beta}^{\alpha(GTD)}$ is the GTD diffraction coefficient and $D_{\beta}^{\alpha(KA)}$ is the Kirchhoff edge diffraction coefficient and is obtained by an asymptotic evaluation (for $\lambda_{\beta} S_{\beta} \gg 1$) of the Kirchhoff integral. It corresponds to the contribution of the scatterer edge to the Kirchhoff integral. Near incident and specular directions, the GTD diffraction coefficient $D_{\beta}^{\alpha(GTD)}$ diverges in the same manner as the KA diffraction coefficient $D_{\beta}^{\alpha(KA)}$ leading to a cancellation of the diffracted field. Similar to the previous case, assuming unit incident field ($u_{\alpha}(\mathbf{r}_{\beta}^{\alpha}) \approx 1$), the diffracted field due to PTD is obtained,

$$\mathbf{u}^{diff(PTD)}(\mathbf{r}) = \sum_{\beta} \left[\mathbf{u}_{\beta}^{\alpha(KA)}(\mathbf{r}) + \left(D_{\beta}^{\alpha(GTD)}(\mathbf{r}) - D_{\beta}^{\alpha(KA)}(\mathbf{r}) \right) \frac{e^{i\lambda_{\beta} S_{\beta}}}{\sqrt{\lambda_{\beta} L_{\beta}}} \mathbf{e}_{\beta}(\mathbf{r}) \right]$$

$$= \mathbf{u}^{Kir}(\mathbf{r}) + \sum_{\beta} \left[\left(D_{\beta}^{\alpha(GTD)}(\mathbf{r}) - D_{\beta}^{\alpha(KA)}(\mathbf{r}) \right) \frac{e^{i\lambda_\beta S_\beta}}{\sqrt{\lambda_\beta L_\beta}} \mathbf{e}_\beta(\mathbf{r}) \right].$$

Therefore, the resultant scattered field, for the PTD model can be expressed by following the notion described in [77]

$$\mathbf{u}^{Scat(PTD)}(\mathbf{r}) = \mathbf{u}^{Rayleigh}(\mathbf{r}) + \mathbf{u}^{Kir}(\mathbf{r}) + \sum_{\beta} \left[\left(D_{\beta}^{\alpha(GTD)}(\mathbf{r}) - D_{\beta}^{\alpha(KA)}(\mathbf{r}) \right) \frac{e^{i\lambda_\beta S_\beta}}{\sqrt{\lambda_\beta L_\beta}} \mathbf{e}_\beta(\mathbf{r}) \right].$$

Résumé

Le domaine du contrôle non-destructif (CND) regroupe l'ensemble des méthodes utilisées pour tester, sans les affecter, les matériaux et détecter/caractériser de potentielles anomalies telles que la présence de défauts, de vides, de la corrosion. Une autre famille d'application de ces techniques concerne l'évaluation des propriétés physiques (électromagnétiques, acoustiques, et/ou mécaniques par exemple) d'un matériau, toujours sans entraîner des dégradations pour ses utilisations futures. L'évaluation non-destructive (END) est utilisée pour identifier, localier et caractériser d'éventuels défauts dans le milieu sous contrôle. Habituellement, les procédures END impliquent l'utilisation d'une source telle que, par exemple, la lumière visible, des ondes ultrasonores, des rayons gamma, des ondes électromagnétiques, etc. Ces méthodes sont utilisées pour garantir les normes de contrôle de qualité et effectuer des tests d'intégrité pendant toute la durée de vie de la pièce contrôlée. Les méthodes d'END peuvent être répartis en deux grands familles : les méthodes qualitatives et quantitatives. Une END qualitative est utilisée pour établir la présence d'un défaut à l'intérieur de la pièce inspectée. On parle d'END quantitative lorsque la localisation et le dimensionnement des défauts ou des variations de propriétés du matériau sont envisagés.

Parmi les différentes méthodes de CND, on peut citer les contrôles par courants de Foucault (CF) et par ultrasons (US) sont les méthodes d'inspection parmi les plus répandues. Dans le cas d'une inspection CF, les signaux électromagnétiques résultent d'un phénomène d'induction dans l'objet inspecté. L'énergie induite interagit avec le matériau de l'objet et la réponse de ce dernier (sous la forme d'une variation d'impédance pour une bobine simple ou d'une tension pour un circuit de détection plus complexe) peut être analysée pour connaître l'état du matériau. Le contrôle par CF est largement utilisé en raison de la possibilité de concevoir des sondes très compactes et sans contact. Le CND par CF est particulièrement adapté à la détection de petites fissures débouchant du côté du capteur ou légèrement enfouis. Dans le cas des inspections par US, des faisceaux d'ondes ultrasonores sont injectés dans le volume des matériaux inspectés. Ces ondes se propagent à travers les matériaux, sont atténuées partiellement et réfléchies aux interfaces. Les différents échos générés par les réflexions aux interfaces de la pièce ou de zones inhomogènes sont analysés pour définir la présence et l'emplacement des discontinuités.

Les outils de simulation sont largement utilisés dans le domaine du CND pour évaluer la performance de la procédure d'inspection, pour la conception de sondes d'inspection, ou simplement pour faciliter l'interprétation des signaux mesurés. Ces outils de simulation numériques peuvent être utilisés pour résoudre deux grands types de problèmes. Le développement des modèles physiques à partir des paramètres physiques et géométriques du système pour prédire avec précision les signaux d'inspection (e.g., champs électromagnétiques, variation d'impédance, etc.) est connu sous le nom de problème direct. Lorsque les signaux d'inspection par CND sont utilisés pour remonter aux paramètres d'intérêt on parle de problème inverse.

Le but de cette thèse est de développer un approche d'inversion robuste afin d'établir procédure de diagnostic en quasi temps réel dans le cadre du CND par courants Foucault. La méthodologie adoptée est basée sur des techniques d'apprentissage supervisé afin de la localiser et caractériser des fissures dans des géométries canoniques. Bien que la méthodologie proposée ait été principalement appliquée pour des cas d'inspection par CF, les stratégies développées peuvent être étendues à d'autres physiques comme le CND par US. Cette partie a été brièvement étudiée dans cette thèse, lorsque les signaux CF et US ont été combinés afin d'améliorer les performances

d'inversion.

1 Apprentissage par les exemples appliqués au CND par CF

En général, l'apprentissage par exemple (en anglais Learning by Examples, LBE) est une approche qui peut être découpée en deux phases. Pendant la phase préliminaire (dite phase hors ligne), un modèle rapide et précis est construit sur la base d'un ensemble d'apprentissage, constitué de paires entrée-sortie, par des techniques de régression/prédiction. Le modèle obtenu est également connu sous le nom de modèle inverse/entraîné. Le modèle développé dans la phase hors ligne est ensuite utilisé pour prédire des sorties, associée à des échantillons de test inconnus, pendant la seconde phase dit « en ligne ». Ces deux phases peuvent être divisées en trois étapes suivantes :

- Sélection et extraction de caractéristiques
- Représentation exhaustive de l'espace des caractéristiques extraites
- Choix de la technique de prédiction (i.e., le modèle inverse).

1.1 Définition du problème

Par souci de simplicité et pour introduire la notation mathématique nécessaire dans ce manuscrit, nous considérons ci-après un problème de CND par CF simple utilisé pour expliquer au mieux l'approche LBE. Considérons une configuration axisymétrique 2*D* composée d'un tube conducteur avec une conductivité σ, et une perméabilité relative $\mu_r = 1$ (Fig. 1). Le tube est inspecté par deux bobines axiales excitées par un courant harmonique. Les bobines sont centrées à $x = 0$ et $z = 0$ et le fonctionnent est en mode différentiel. Le déplacement des bobines s'effectue le long de l'axe y du tube. Le tube est affecté par une gorge axisymétrique (i.e., défaut circonférentiel) qui occupe un volume Ω (Fig. 1) dans le tube (i.e., $\sigma(\mathbf{r}) = \sigma \forall\, \mathbf{r} \in \Omega$).

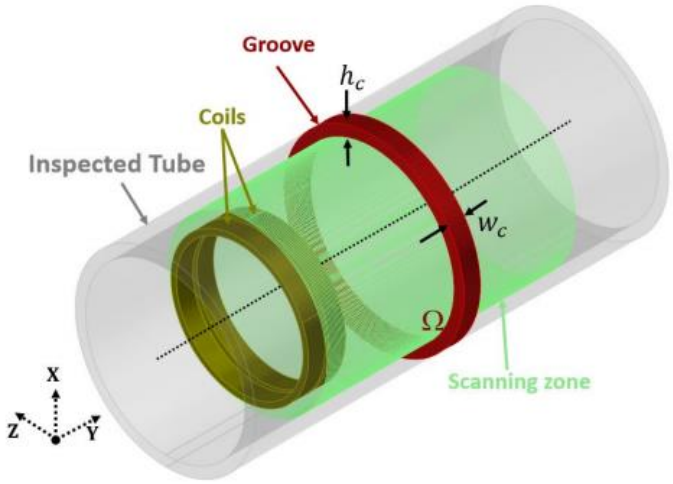


Figure 1: Exemple de géométrie de tube étudiée.

1.2 Sélection et extraction de caractéristiques

1.2.1 Sélection de caractéristiques

La sélection des caractéristiques est un processus à travers lequel les signaux CF sont collectés en termes de variation d'impédance des bobines d'inspection à partir de l'inspection du milieu à contrôler. Ces caractéristiques représentent l'ensemble des mesures significatives dans l'espace du signal CF. Depuis le début de cette thèse, nous avons utilisé le simulateur CIVA [14] comme opérateur direct $\Phi\{.\}$ pour générer des signaux CF synthétiques. Le signal CF est décrit en termes de variation d'impédance à la *k*-ième ($k = 1,...,K$) position de la sonde comme

$$\Psi_k = -\frac{1}{I^2}\int_{\Omega} \mathbf{E}^{inc}(\mathbf{r}\,|\,\mathbf{r}_k)\cdot\boldsymbol{\rho}(\mathbf{r}\,|\,\mathbf{r}_k)d\mathbf{r} \qquad (1)$$

I est le courant circulant à l'intérieur des bobines, $\mathbf{E}^{inc}(\mathbf{r}\,|\,\mathbf{r}_k)$ est le champ d'incident généré à la position **r** dans le tube sans défaut ($\mathbf{r}_k = y_k$ représente la *k*-ième position de la bobine le long du tube). $\boldsymbol{\rho}(\mathbf{r}\,|\,\mathbf{r}_k)$ est la densité du courant du dipôle dans le volume de la gorge et représente l'inconnue, associée à la présence de la gorge. Elle est lié au champ total, $\mathbf{E}^{tot}(\mathbf{r}\,|\,\mathbf{r}_k)$ qui peut être exprimé par :

$$\boldsymbol{\rho}(\mathbf{r}\,|\,\mathbf{r}_k) = [\sigma(\mathbf{r}) - \sigma]\mathbf{E}^{tot}(\mathbf{r}\,|\,\mathbf{r}_k). \qquad (2)$$

Plus de détails sur l'approche théorique peuvent être trouvés dans les références [41, 42]. Le formalisme évoqué ici est un de ceux développés dans la plateforme pour la simulation d'inspection CND par CF. Nous définissons Ψ comme vecteur des signaux CF associé aux caractéristiques sélectionnées comme $\underline{\Psi} = (\Psi_k; k = 1,..., K)$. Le vecteur décrivant les paramètres associés au défaut est noté $\underline{p} = (p_q; q = 1,..., Q)$, c'est à dire que $\Psi_k = \Phi\{\mathbf{r}_k, \underline{p}\}$ vérifie l'Eq. (1). La gorge considérée dans notre problème occupe la région Ω qui peut être définie à l'aide de $Q = 2$ paramètres, comme $\underline{p} = (h_c, w_c)$ ou h_c et w_c représentent la hauteur et la largeur de la gorge, respectivement (Fig. 1). En raison de la nature complexe des signaux CF, le vecteur de signal associé peut être écrit comme $\underline{\Psi} = (\Re\{\Psi_k\}, \Im\{\Psi_k\}; k = 1,..., K)$ avec $\underline{\Psi} \in \mathbb{R}^{1\times F}$. Par conséquent, l'espace des signaux CF noté $\aleph_\Psi$ est de dimension $F = 2K$.

En traitant un ensemble plus grand de signaux CF (c'est-à-dire, par exemple, la variation d'impédance collectée à partir d'un nombre plus élevé de points de mesure *K*), le nombre de caractéristiques ECT traitées *F* augmente. Plus le nombre de fonctions est élevé, plus la phase d'entraînement est compliquée et le nombre d'échantillons d'entraînement nécessaires pour développer un modèle optimal devient important. Ce problème est également connu sous le nom de malédiction de la dimensionnalité pour les algorithmes d'apprentissage. D'un autre côté, plus nombre d'échantillons augmente, plus le temps d'entraînement ainsi que la complexité de la tâche d'apprentissage augmentent. Pour faciliter ce problème, l'extraction de caractéristiques au moyen de techniques de réduction de dimensionnalité est utilisée.

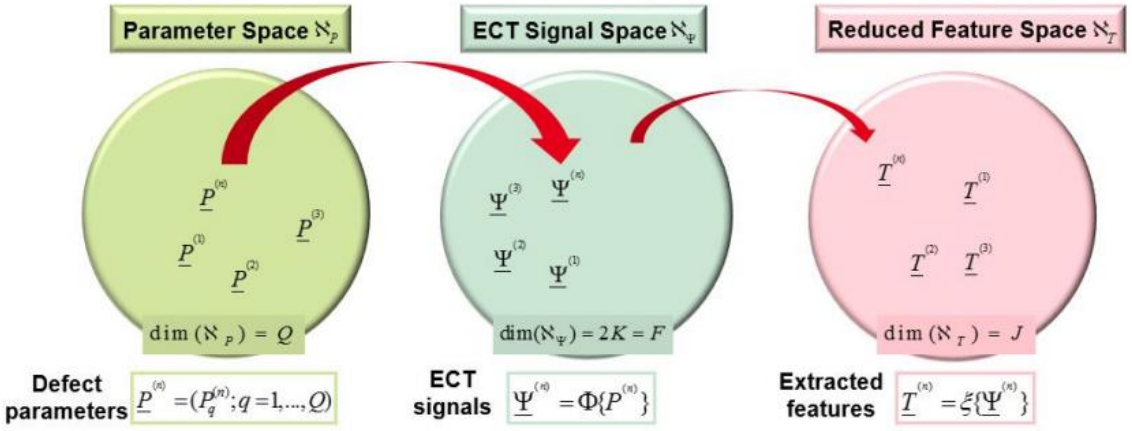


Figure 2: Mappage entre l'espace des paramètres, l'espace des signaux ECT et l'espace réduit des caractéristiques.

1.2.2 Extraction de caractéristiques

Le but de l'extraction des caractéristiques est de réduire l'espace des caractéristiques CF de $F = 2K$ à un espace des fonctions extraites ayant une dimension $\aleph_T$ beaucoup plus faible. C'est-à-dire que $\aleph_T$ représente l'espace caractéristique réduit qui contient l'ensemble des caractéristiques extraites T de dimension $\dim(\aleph_T) = J$ où J<<F. La Fig. 2 montre le schéma associé au processus d'extraction de caractéristiques. Les signaux CF sont sont projetés dans un espace de dimension inférieure $\aleph_T$ par un modèle d'extraction linéaire ou non linéaire selon la méthode de réduction de la dimensionnalité utilisée. Par exemple, dans la Fig. 3, nous avons montré la représentation de l'espace des paramètres et l'espace des signaux CF réduit projeté dans un espace bidimensionnel à l'aide de la technique d'extraction de caractéristique PLS (Partial Least Square en anglais).

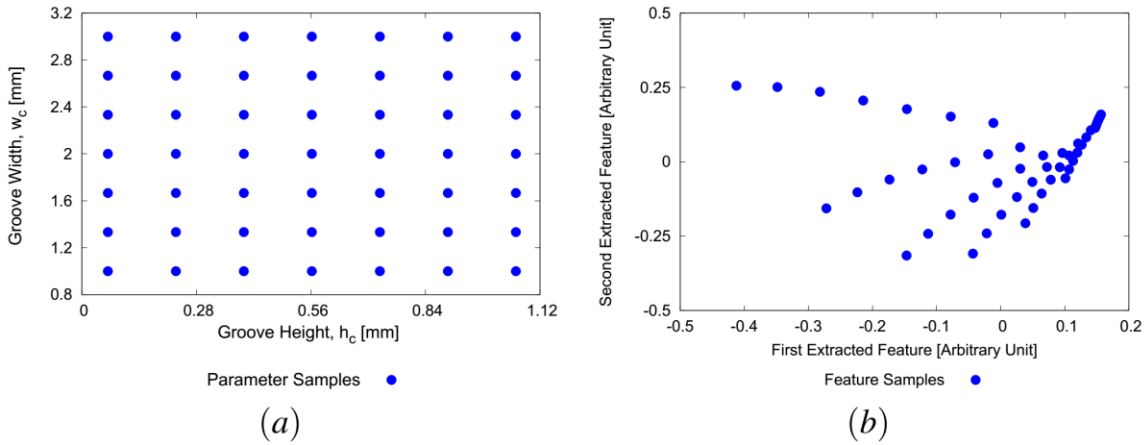


Figure 3: Exemples de signaux calculés sur (*a*) l'espace des paramètres $\aleph_p$ et (*b*) sur l'espace des caractéristiques extraites résultant $\aleph_T$.

La Fig. 3(a) représente la vue 2D de l'espace des paramètres $\aleph_P$ contenant les paramètres de la gorge obtenue par la stratégie d'échantillonnage factoriel (GRID). L'espace des caractéristiques extraites résultantes $\aleph_T$ caractéristiques extraites est montré en Fig. 3 (b). Pour l'illustration graphique, l'espace de caractéristiques extrait est affiché pour J = 2 caractéristiques extraites. A partir de la Fig. 3 (b), nous pouvons observer que les entités extraites dans $\aleph_T$ présentent des profils avec des zones sur-échantillonnées et une zone sous-échantillonnée. Ce phénomène dépend de l'espace original des signaux CF et de la façon dont les paramètres sont distribués dans l'espace des paramètres. La zone sur-échantillonnée contient des informations redondantes pour les algorithmes d'apprentissage, ce qui augmente également le nombre d'échantillons d'apprentissage à traiter. Par conséquent, en évitant la zone sur-échantillonnée dans $\aleph_T$ tout en échantillonnant uniformément l'espace des caractéristiques extraites, nous pouvons augmenter la capacité d'apprentissage des techniques de prédiction et obtenir un meilleur résultat d'inversion. Dans la section suivante, une brève description de cette stratégie est illustrée. D'autres types de techniques d'extraction de caractéristiques telles que l'analyse en composantes principales (PCA en anglais), ainsi que la version basée sur les noyaux du PCA et PLS, etc. peuvent également être appliquées [32, 33].

1.3 Représentation exhaustive de l'espace des caractéristiques extraits

Le signal ECT varie avec la variation des différents paramètres de défaut. Par exemple, la variation de la hauteur de la gorge a une plus grande influence sur la modification des signaux CF que la variation de sa largeur. Par conséquent, si un certain budget de simulations est fourni, il est essentiel de sélectionner un ensemble optimal d'échantillons permettant de concevoir un modèle le plus précis possible. En regardant la Fig. 3, on peut remarquer qu'une répartition égale des échantillons dans l'espace des paramètres contribue à la distribution irrégulière de l'espace caractéristique extrait sur lequel la tâche d'apprentissage est déjà effectuée. Pour avoir un modèle adapté et plus précis, une distribution uniforme de l'espace des caractéristiques extraites est nécessaire.

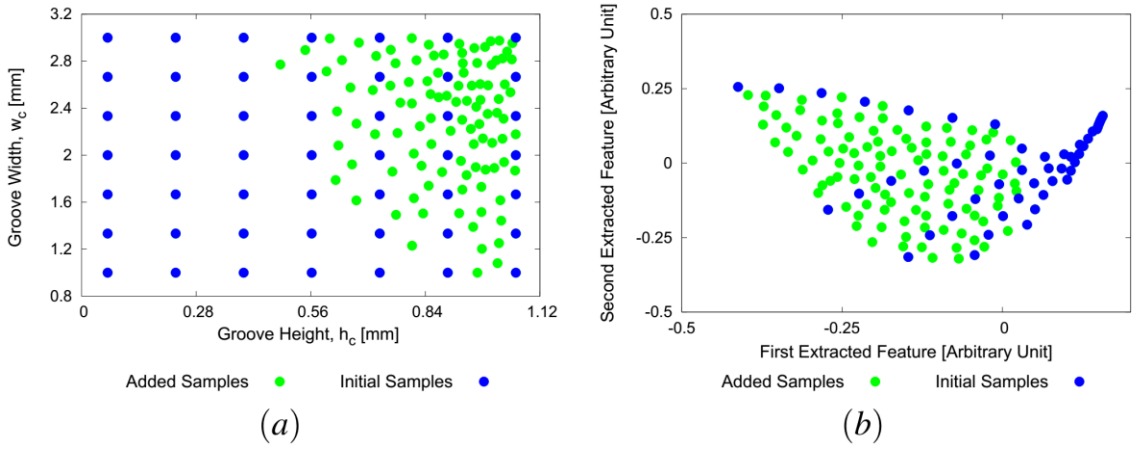


Figure 4: Entraînement des emplacements d'échantillons mappés sur (*a*) paramètre et (*b*) espace des caractéristiques extraites résultantes.

De plus, en répartissant uniformément l'espace de caractéristiques extrait, des échantillons redondants et inutiles (c'est-à-dire des points de caractéristiques trop proches) pour différentes

configurations de paramètres de défauts peuvent être évités. Un modèle plus précis peut donc être obtenu pour un nombre donné d'échantillons d'apprentissage. Dans cette thèse, nous avons appliqué la technique d'extraction des caractéristiques PLS avec une version modifiée de l'algorithme dit de remplissage de l'espace de sortie (en anglais OSF) [44] que nous appellerons PLS-OSF. Les résultats cette procédure d'échantillonnage adaptative de l'espace des caractéristiques est montre dans la Fig. 4 (b) ou on peut noter le remplissage quasi uniforme de l'espace des caractéristiques extraites. Dans l'espace de paramètres, on peut noter que cet échantillonnage optimal n'est pas trivial (Fig. 4 (a).

1.4 Choix de la technique de prédiction

Le choix de la technique de prédiction (appelée algorithme d'apprentissage) est l'une des étapes cruciales des méthodes de LBE. Traitant un problème paramétrique, le modèle de prédiction est développé en cartographiant les signaux NDT avec les paramètres de fissure/défaut correspondants. Dans le cadre de LBE, des algorithmes d'apprentissage machine tels que Support Vector Regressor (SVR) [45], Kernel Ridge Regressor (KRR) [46], Relevance Vector Machine (RVM) [47] et l'interpolateur de la fonction de base radiale augmentée (A-RBF) [48] ont été implémentés et comparés.

D'une manière générale, l'interpolateur A-RBF est obtenu en ajoutant un terme polynomial à l'interpolateur RBF standard. Le terme polynomial d'A-RBF facilite l'ajustement d'un modèle d'apprentissage en utilisant les données d'apprentissage. Plus le modèle d'apprentissage utilise d'échantillons, plus le modèle est précis. Cependant, cela conduit à un problème de sur-ajustement (dit overfitting en anglais) et le modèle d'apprentissage devient plus complexe en présence de valeurs aberrantes ou de bruit. Par conséquent, la précision de prédiction d'A-RBF est plus élevée lors du traitement avec des données sans bruit et, inversement, diminue lorsque l'ensemble de test inconnu est corrompu.

Disposant de paramètres de régularisation, les régresseurs KRR, RVM et SVR sont bien adaptés pour traiter les données bruitées qui ont été traitées dans [24] dans le cadre de LBE. Étant donné un ensemble de données $D_{N_q} = \left\{ \left(\underline{\Psi}^{(n)}; p_q{}^{(n)} \right); n = 1,..., N; q = 1,..., Q \right\}$, le but de la tâche de régression est de trouver une fonction $\Theta_q\{.\}$ estimant la relation entre $\underline{\Psi}^{(n)}$ et $p_q{}^{(n)}$ telle que $\Theta_q\left\{\underline{\Psi}^{(n)}\right\}$ fournit un sortie prédite/estimée qui est proche de $p_q{}^{(n)}$ pour tout *n*. Cette tâche est généralement effectuée en utilisant une fonction de perte *L* qui mesure l'écart entre $p_q{}^{(n)}$ (le paramètre réel) et la valeur prédite $\Theta_q\left\{\underline{\Psi}^{(n)}\right\}$ et minimise le « risque empirique ». Par exemple, la méthode KRR se fonde sur l'utilisation de la fonction Least Squares Error (LSE), $L = \sum_{n=1}^{N} \left(p_q{}^{(n)} - \Theta_q\left\{\underline{\Psi}^{(n)}\right\} \right)^2$. Cette fonction de perte au carré est également appelée fonction de perte L_2, qui est sensible à la présence de valeurs aberrantes ou de bruit. Par conséquent, tout $\underline{\Psi}^{(n)}$ pour le paramètre correspondant $p_q{}^{(n)}$ différant beaucoup du reste du signal ECT dans $\underline{\underline{\Psi}}$ (où $\underline{\underline{\Psi}} = \left(\underline{\Psi}^{(n)}; n = 1,..., N \right)$) aura un effet disproportionné sur les paramètres KRR. Ainsi, la capacité de

prédiction de KRR se dégrade significativement en présence de bruit [33] en raison d'un sur-ajustement.

Comme il utilise une fonction de perte L_1, cet effet n'est pas dominant pour l'application de SVR. Bien que SVR a un grand succès dans le schéma d'apprentissage supervisé, il présente les inconvénients suivants :

- Le nombre de vecteurs de support augmente linéairement avec la taille de l'ensemble d'apprentissage. Une certaine forme de post-traitement est également nécessaire pour réduire la complexité de calcul.

- SVR produit une estimation ponctuelle pour la régression et une décision binaire difficile pour la classification. L'approche probabiliste n'est pas utilisée.

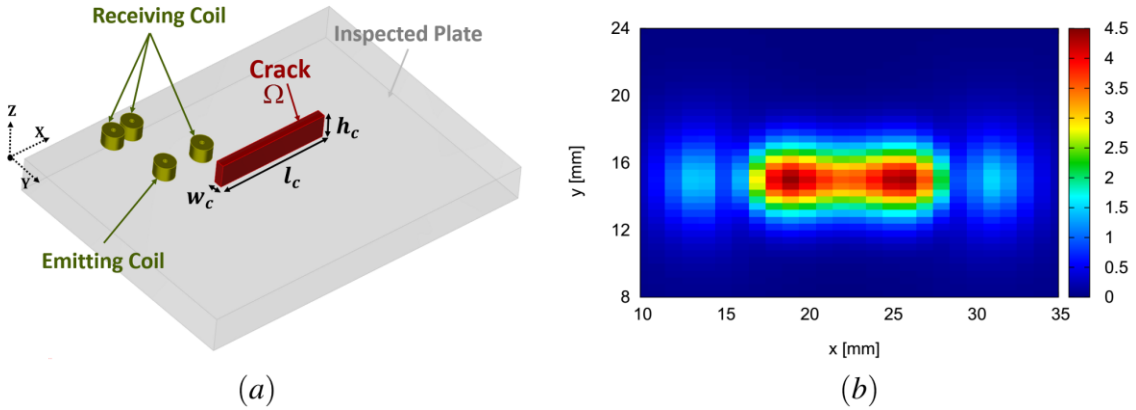
RVM a été introduit en utilisant un traitement bayésien qui ne souffre pas de toutes ces limitations. Il utilise un a priori sur les poids du modèle, régis par un ensemble d'hyper paramètres, chacun étant associé à un poids particulier. Les valeurs les plus probables sont estimées itérativement à partir des données. Les distributions postérieures de plusieurs des poids sont maximales autour de zéro, ce qui entraîne un certaine sparsité. Le vecteur d'apprentissage associés à ces poids non-zéros restants sont appelés machine à vecteurs de relevance (RVM en anglais). La caractéristique importante de RVM est qu'il utilise généralement moins de vecteurs supports que d'autres méthodes comme SVR par exemple.

Cette dernière a été appliquée avec succès dans les applications électromagnétiques [49, 50] pour la résolution de problèmes inverses et sa robustesse sur des données bruitées est bien connue. De plus, dans [24], les auteurs ont montré que la SVR surpasse la RVM pour traiter de vraies données expérimentales dans les problèmes de NDE. Par conséquent, nous considérons SVR comme la technique de prédiction la plus appropriée et nous avons utilisé SVR tout au long de ce travail de recherche. De plus, il convient de noter que la précision de prédiction et la robustesse des différents algorithmes d'apprentissage en présence de données bruitées dépendent de la génération d'un ensemble d'apprentissage approprié, de différents paramètres de défaut (par exemple, relation entre les signaux ECT pour un paramètre particulier) et des capacités de généralisation desdits algorithmes. Pour des problèmes dimensionnels plus élevés, il est assez difficile d'utiliser d'autres algorithmes d'apprentissage automatique populaires tels que Réseaux des Neurones Artificiels (RNA) comme technique de prédiction. Cependant, en traitant de manière appropriée les signaux ECT en appliquant une extraction de caractéristiques et/ou en appliquant un échantillonnage adaptatif, RNA peut également être utilisé en tant que technique de prédiction. Dans les sections suivantes, LBE a été appliqué pour résoudre les problèmes de localisation et de caractérisation des fissures dans différents cas de test.

2 Cas de test 1: Caractérisation de fissure (s) en plaque (s) conductrice (s)

Cette section décrit le problème de caractérisation de fissure(s) en temps réel en exploitant la stratégie LBE dans le contexte du CND par CF. Un ensemble d'apprentissage optimal a été généré en phase hors ligne en adoptant l'extraction des caractéristiques PLS, combinée à une version personnalisée du remplissage de l'espace de sortie (OSF). L'algorithme SVR est utilisé pour

développer un modèle précis basé sur l'ensemble d'apprentissage et ensuite l'inversion en temps réel (phase en ligne) a été effectué sur un ensemble de test inconnu. La robustesse de l'approche référencée ci-après comme approche PLS-OSF/SVR est évaluée numériquement en présence d'un ensemble de test bruité synthétique et comparée à une approche GRID/SVR plus standard [24].



(a) (b)

Figure 5: Exemples de (*a*) géométrie de plaque étudiée et (*b*) carte de signaux CF en termes de valeur absolue de variation de tension de bobine.

Considérons une plaque conductrice d'épaisseur 1.27 mm, ayant une conductivité électrique de 1.02 MS/m et une perméabilité magnétique relative de 1. La plaque est affectée par une seule fissure rectangulaire débouchant coté oppose à la sonde ayant une région géométrique Ω dans la structure à tester (SAT) (Fig. 5). La fissure est caractérisée par un total de $Q = 3$ descripteurs $\underline{p} = (l_c, h_c, w_c)$, où l_c, h_c, w_c représentent la longueur, la hauteur et la largeur de la fissure, respectivement. La plaque est inspectée par un motif de 4 bobines à une fréquence de 200 kHz : dans cette configuration seule la bobine émettrice et la bobine réceptrice la plus proche sont actives. Plus de détails sur les paramètres des bobines peuvent être trouvés dans [64]. La bobine réceptrice enregistre le signal CF avec un pas de 1 mm et de 0.4 mm suivant les axes *X* et *Y*, respectivement. Le signal CF (c'est-à-dire, ici la variation de tension ΔV aux bornes de la bobine réceptrice) sont collectés à partir de $K = 28 \times 23 = 644$ de points inspectés, qui sont représentés par $F = 2K = 1288$ valeurs (une partie réelle et une partie imaginaire par position).

2.1 Résultats et discussion

Dans cette section, la performance (en termes d'estimation paramétrique) prédictive de la méthode SVR est comparée en appliquant les deux approches d'échantillonnage GRID et PLS-OSF. On utilise ici $N_0 = 27$ et $N = 512$ échantillons minimum et maximum, respectivement, pour la génération d'ensembles d'apprentissage en appliquant des stratégies d'échantillonnage PLS-OSF et GRID. Le nombre d'échantillons de l'ensemble de test M a été choisi égal à 1000 avec une stratégie d'échantillonnage de type hyper cube latin (LHS en anglais). Dans le cas d'un échantillonnage PLS-OSF, $J = 5$ caractéristiques ont été extraites, parmi les 1288 caractéristiques associées aux mesures CF, en utilisant la méthode PLS pour les ensembles d'apprentissage et de test. En traitant un problème dimensionnel supérieur (par exemple, F=1288 caractéristiques ECT), l'échantillonnage GRID nécessite d'avoir un plus grand nombre d'échantillons d'apprentissage (N) pour obtenir suffisamment d'informations et construire un modèle précis [24]. Cela rend également

difficile pour le modèle SVR de traiter les données bruitées pour un nombre inférieur d'échantillons d'entraînement *N*.

PLS-OSF / SVR montre une précision de prédiction améliorée pour des valeurs N plus faibles, ce qui est également robuste dans le cas d'un ensemble de test bruité. En appliquant l'extraction de caractéristiques PLS, les ensembles d'apprentissage et de test peuvent être représentés en dimension inférieure (J<<F) avec des grandeurs caractéristiques plus significatives. Ceci permet à la technique de prédiction d'être plus précise dans la création du modèle d'apprentissage. En explorant uniformément l'espace caractéristique et l'espace de paramètre de fissure associé, la stratégie d'échantillonnage PLS-OSF peut avoir une représentation optimale (c'est-à-dire, complète et non redondante) de l'espace caractéristique extrait. Il en résulte une erreur de prédiction plus faible sur l'estimation de la hauteur de la fissure pour un *N* plus faible sur les données de test silencieuses ainsi que sur les données bruitées.

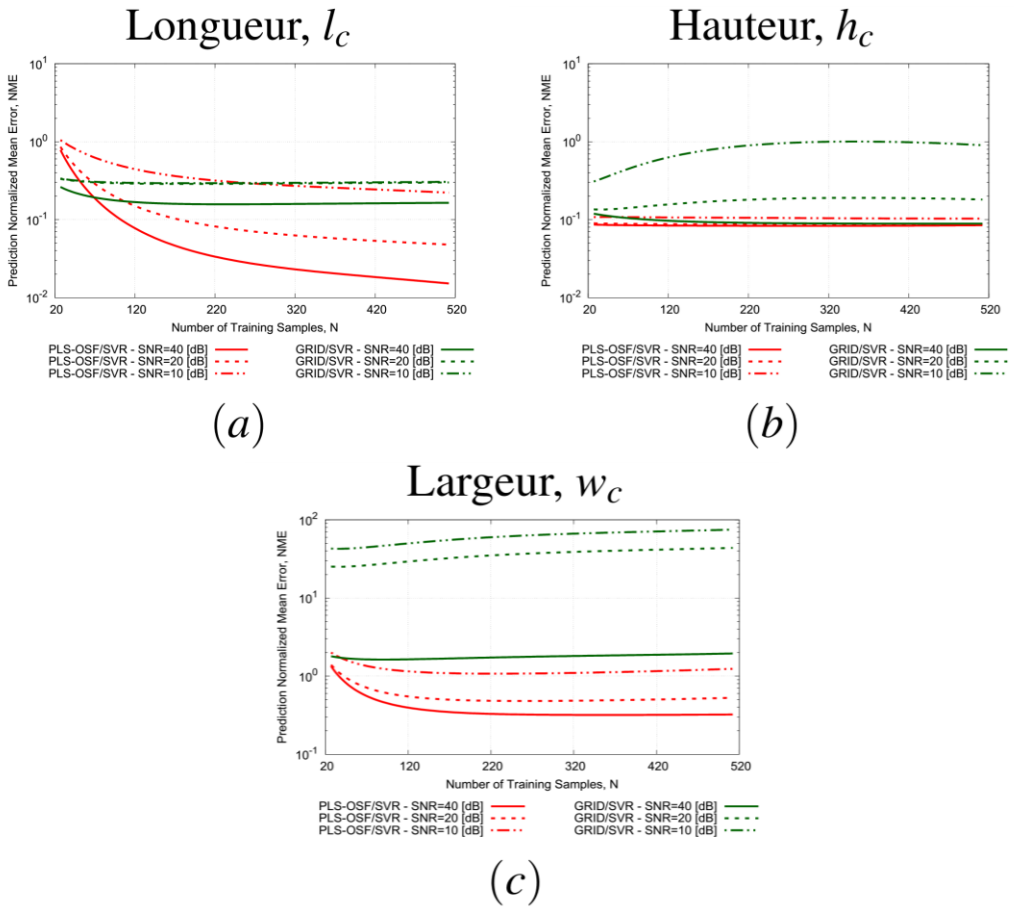


Figure 6: Comparaisons de prédiction PLS-OSF/SVR et GRID/SVR en fonction de *N* pour la (*a*) longueur l_c,(*b*) hauteur h_c et (*c*) largeur w_c de la fissure. Différents rapports de signal à bruit *SNR* = 10, 20, 40 dB ont été considérés, avec les paramètres *F* = 1288, *J* = 5, *M* = 1000.

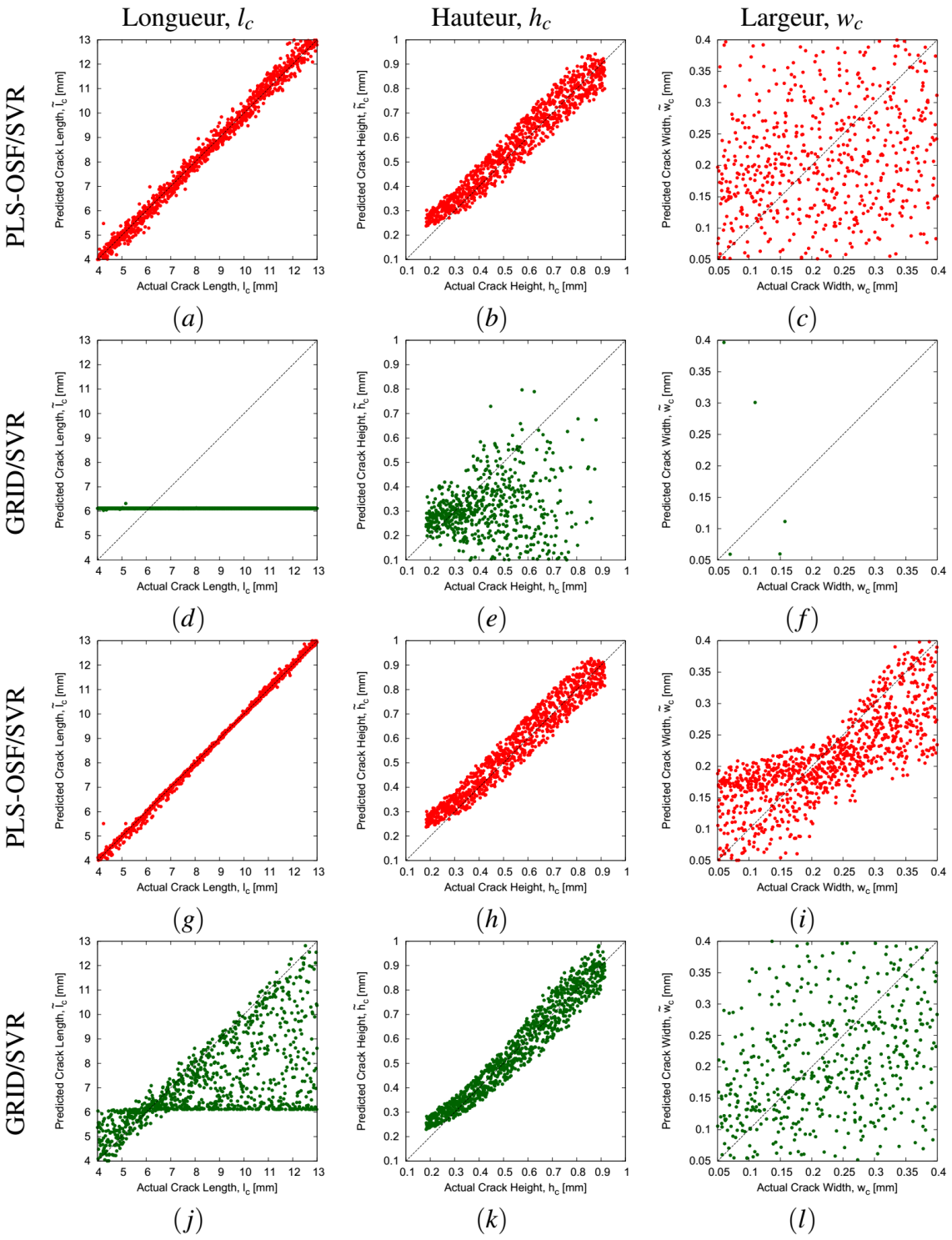


Figure 7: Comparaisons des prédictions PLS-OSF/SVR et GRID/SVR représentées en termes de courbes de régression vraies et prédites (longueur de fissure l_c, hauteur h_c, largeur w_c) sur l'ensemble de test en (*a*) −(*f*) *SNR*= 10 dB et en (*g*)−(*l*) *SNR*= 40 dB, avec N = 512, F = 1288, J =5, M = 1000.

Les figures Fig. 6 et Fig. 8 résument la robustesse de la précision de l'inversion en présence d'un ensemble de test bruité en utilisant l'approche PLS-OSF/SVR et en le comparant avec la performance de l'approche GRID/SVR. Nous pouvons observer qu'avec une augmentation de *N*, l'erreur de prédiction, *NME* diminue légèrement pour un *SNR* plus élevé en utilisant GRID/SVR pour l'estimation des paramètres l_c, h_c et w_c. Cela indique que pour traiter un problème de dimension $F = 1288$, le nombre d'échantillons d'entraînement appliqués $N = 512$ n'est pas suffisant pour un développement précis du modèle d'entraînement par GRID/SVR (i.e., pas d'extraction des caractéristiques). Par conséquent, avec un *SNR* décroissant, la méthode GRID/SVR ne peut prédire aucun paramètre de fissure. Alors que l'approche PLS-OSF/SVR ne traite que de J=5 caractéristiques extraites nécessitant un nombre inférieur d'échantillons d'apprentissage pour avoir un modèle d'apprentissage précis. En conséquence, la précision de la prédiction de l'approche PLS-OSF/SVR augmente fortement tandis que *N* augmente grâce à l'échantillonnage adaptatif. Du point de vue du calcul, le temps moyen de prédiction n'était que de 0.03 secondes pour prédire des échantillons de test M=1000 pendant la phase en ligne par PLS-OSF/SVR sur un ordinateur portable standard, ce qui permet de qualifier l'estimation de quasi temps réelle. En appliquant la stratégie d'échantillonnage PLS-OSF, un ensemble d'apprentissage optimal peut être obtenu, ce qui implique une plus grande précision de prédiction en présence de données bruitées. De plus, la stratégie d'échantillonnage PLS-OSF augmente la capacité d'apprentissage de l'algorithme employé.

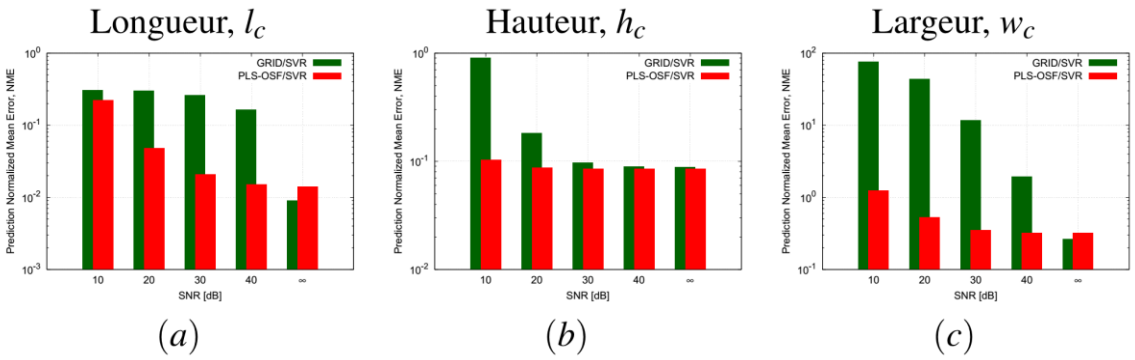


Figure 8: Comparaisons des prédictions PLS-OSF/SVR vs. GRID/SVR en termes de *SNR* vs. *NME* dans les mêmes conditions que précédemment.

3 Cas test 2: Localisation et caractérisation de fissures basées sur la fusion de données NDT multi physiques

Cette section décrit une méthodologie d'inversion exploitant la technique d'apprentissage par les exemples en utilisant une approche de simulation multi physique et de fusion de données. Dans le cadre de LBE, nous avons appliqué les méthodes CND aux essais par courants de Foucault et ultrasons pour résoudre un problème complexe de localisation et de caractérisation des fissures en développant des modèles d'entraînement précis pour chaque méthode. Suite à cela, une approche innovante de fusion de données multi-physiques (CF-US) a été adoptée pour développer un schéma d'inversion robuste. L'extraction des caractéristiques des moindres carrés partiels (PLS) combinée à une version personnalisée du remplissage de l'espace de sortie (OSF) (échantillonnage PLS-OSF) a été appliquée avec l'algorithme SVR pour développer un modèle d'entraînement précis basé sur

l'ensemble d'apprentissage et ensuite l'inversion en temps réel (phase en ligne) a été effectuée sur un ensemble de test inconnu. La robustesse de l'approche CF-US proposée est évaluée numériquement en présence d'un bruit de fond synthétique par la stratégie PLS-OSF/SVR et comparée à chaque méthode de CND séparément (c'est-à-dire, CF et US).

Une plaque homogène faite d'un alliage d'aluminium 2024 d'une épaisseur de 6 mm a été étudiée à la fois par les méthodes d'essai par courants de Foucault et par ultrasons séparément. La plaque est constituée d'une fixation (alésage) de rayon 3.75 mm et 6.00 mm de hauteur. La plaque est affectée par une seule entaille (par exemple, une fissure étroite) de volume Ω ayant une largeur fixe de 0.01 mm et une hauteur de 2 mm (Fig. 9). La fissure étroite est caractérisée par un total $Q = 3$ descripteurs de longueur (*lc*), de ligament (δ*c*) et de distance angulaire (φ_c) (i.e., $\underline{p} = (l_c, \delta_c, \varphi_c)$.

Les signaux CF et US sont échantillonnés sur 81 positions dans la direction *X* avec un pas de 0.5 mm et sur 41 positions dans la direction *Y* avec un pas de 1 mm. L'inspection de cette pièce par des méthodes CF et US va générer des signaux aux propriétés très différentes. L'intérêt de développer un schéma d'inversion exploitant ces deux mesures nous permet de maximiser la capacité de caractérisation et de localisation des fissures. Nous avons appliqué une stratégie de fusion de données multi-physiques pour améliorer les performances d'inversion. Dans cette approche de fusion de données, les signaux CF et les signaux US sont générés séparément par leur propre solveur direct (CIVA) et les ensembles d'apprentissage de ces deux ensembles de données sont fusionnés en concaténant les données CF et US représentées avec CF-US.

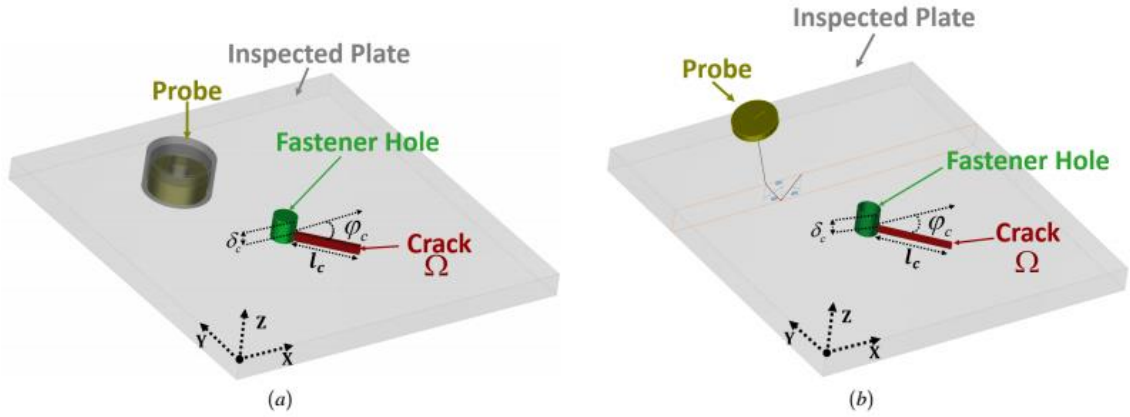


Figure 9: Des exemples de (*a*) géométrie de plaque étudiée pour CF et (*b*) géométrie de plaque étudiée pour US.

3.1 Résultats et discussion

Dans le cadre de l'évaluation, on utilise $N_0 = 27$ et $N = 216$ comme nombre minimum et maximum d'échantillons pour la formation des ensembles d'apprentissage en appliquant la stratégies échantillonnage PLS-OSF pour la fusion de données CF, US et CF-US. Considérant que, $M = 1000$ échantillons générés par les stratégies d'échantillonnage LHS pour construire l'ensemble de test. $J = 20$ entités sont extraites de $F_{CF} = 2K = 6642$ (Caractéristiques ECT), $F_{US} = K = 3321$ (caractéristiques US) et $F_{CF-US} = 3K = 9963$ (caractéristiques ECT-UT) séparément par extraction de caractéristiques PLS pour les ensembles d'apprentissage et de test.

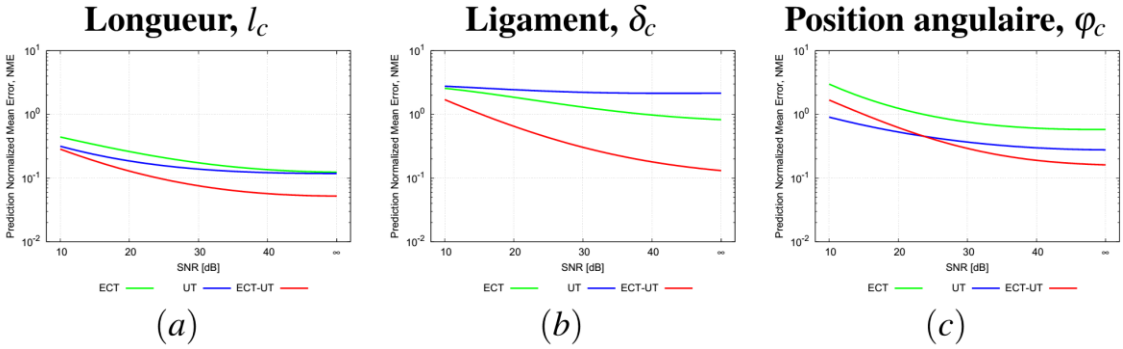


Figure 10: Evaluation numérique: (longueur de la fissure, *lc* = [3, 10] mm, ligament, δ*c* = [0, 4] mm, et position angulaire, φ_c = [0, 90] deg.)- Les résultats de prédiction PLS-OSF/SVR sont montrés en termes d'erreur de prédiction (erreur moyenne normalisée, *NME*) vs. *SNR* pour la fissure (a) longueur *lc*, (b) ligament δ*c* et (c) estimation de la position angulaire φ_c pour *N* = 216, *J*= 20, avec *M* = 1000 à travers CF, US et CF-Us.

En combinant les signaux CF et US, la fusion de données CF-US contient à la fois des informations provenant des signaux CF et US. Cela améliore la capacité d'apprentissage de la méthode SVR pour le développement du modèle de formation. Par conséquent, CF-US peut améliorer la précision de la prédiction en utilisant l'approche PLS-OSF/SVR. En appliquant l'extraction de caractéristiques PLS, nous pouvons extraire la plupart des informations significatives à partir des signaux combinés (c'est-à-dire, CF-US). Comme nous l'avons constaté, les signaux CF sont en grande partie corrompus pour imposer un bruit, en combinant les signaux CF et US, nous pouvons améliorer les performances globales de l'inversion. Dans la Fig. 10, la précision des performances d'inversion de la fusion de données ECT-UT en termes de NME pour différents *SNR* à *N* = 216 sont comparés aux méthodes ECT et UT. Ici, il apparaît que la fusion de données ECT-UT améliore la précision de prédiction et est robuste sur les données bruitées.

Il est également visible qu'en combinant les deux signaux, nous pouvons bénéficier à la fois des méthodes CND et l'augmentation adaptive des échantillons grâce à PLS-OSF, ce qui améliore la précision de la prédiction pour tous les paramètres de fissure. La fusion de données CF-US montre une meilleure précision de prédiction sur la longueur de fissure *lc*, la distance ligamentaire δ*c* et l'estimation de la position angulaire φ_c par stratégie PLS-OSF/SVR méthodes des CF et US. La Fig. 11 montre les diagrammes de dispersion des vrais paramètres de fissures vs. paramètres prédits obtenu par stratégie PLS-OSF/SVR pour *N* = 216. D'un point de vue qualitatif, la fusion de données CF-US améliore la capacité d'apprentissage qui fournit une meilleure estimation de *lc*, δ*c* et φ_c en la comparant à celle des signaux CF et US. En utilisant des signaux CF, l'estimation de *lc* devient plus difficile qu'à partir des signaux US en raison de la présence d'un dispositif de fixation. Cependant, CF-US fournit une plus grande précision de prédiction pour l'estimation de *lc* sur un ensemble de test silencieux (voir le chapitre 7).

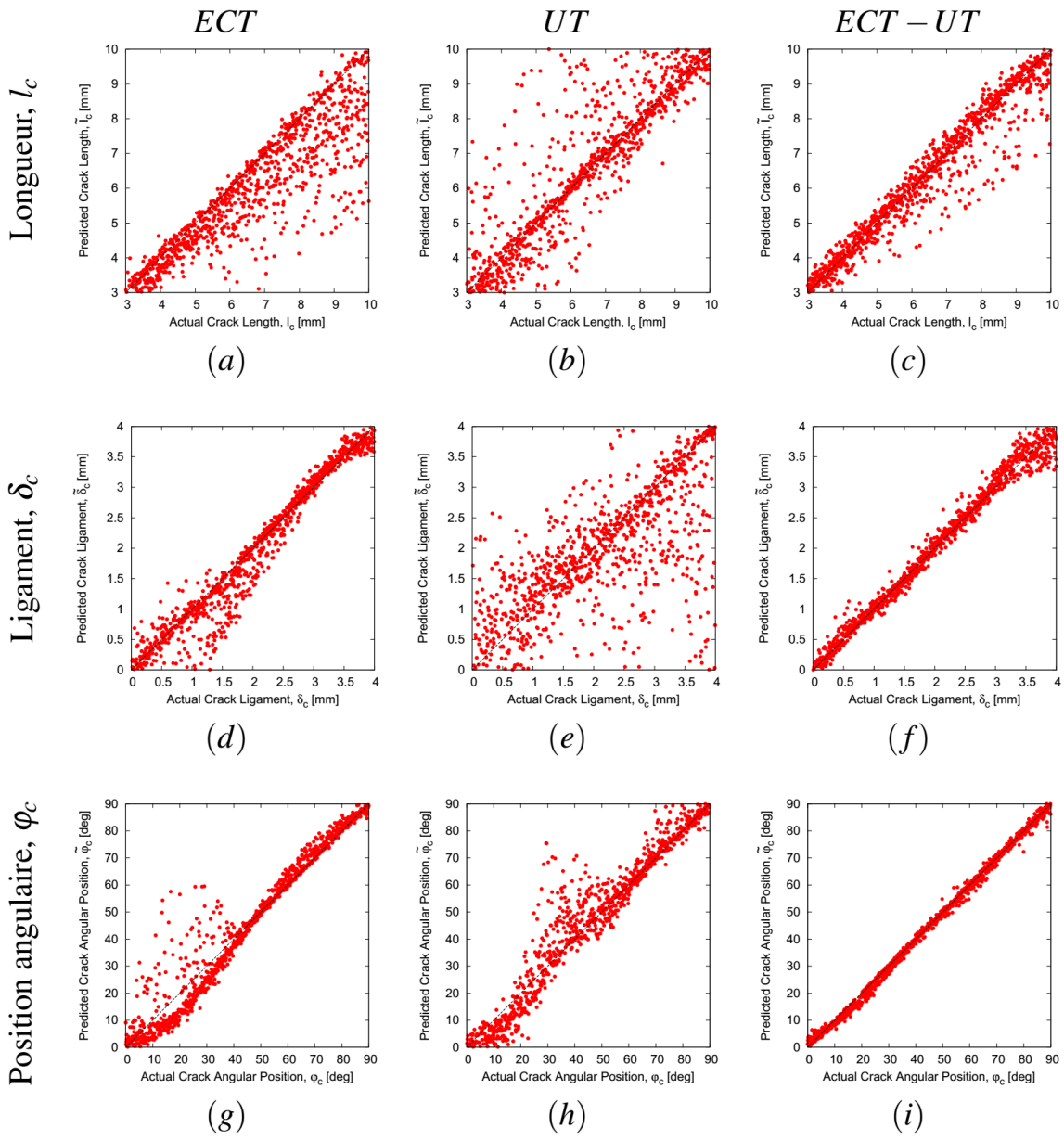


Figure 11: Evaluation numérique: (longueur de la fissure, *lc* = [3, 10] mm, ligament, δ*c* = [0, 4] mm, et position angulaire, φ_c = [0, 90] deg.) - Les résultats de prédiction PLS-OSF / SVR sont représentés en termes de courbes de régression vraies et prédites (longueur de la fissure *lc*, ligament δ*c* et position angulaire φ_c) Test de bruit pour *N* = 216, *J* = 20, *M* = 1000 à CF, US et CF-US.

4 Conclusions et perspectives

Le but de cette thèse était de développer un schéma d'inversion robuste afin de construire un système de diagnostic en temps réel pour le contrôle non destructif par courants de Foucault. La méthodologie LBE a été adoptée pour la localisation des fissures et des défauts, ainsi que pour des

problèmes de caractérisation dans le milieu inspecté. Bien que le LBE ait été principalement appliqué pour traiter les signaux CF, les stratégies développées ont été étendues à d'autres physiques en l'appliquant aux données de test par ultrasons. Dans la dernière partie du travail, les signaux CF et US ont été combinés afin d'améliorer les capacités d'inversion.

La technique d'extraction de caractéristiques PLS a été appliquée pour extraire les caractéristiques les plus appropriées de l'espace réel du signal CND. D'autres techniques d'extraction de caractéristiques telles que l'analyse de composantes principales, l'analyse discriminante linéaire, l'analyse de corrélation canonique et leurs extensions non linéaires (c'est-à-dire la version de noyau) peuvent également être appliquées dans le cadre de LBE.

Selon la formulation de chacune de ces méthodes, les performances d'inversion peuvent varier. Ce travail de recherche peut être étendu à d'autres méthodes CND (par exemple, ultrasons, thermographie, etc.) afin d'évaluer la performance de ces approches. Pour ce qui concerne les signaux US au chapitre 7, nous avons traité l'amplitude maximale des signaux US collectés à partir de chaque point d'inspection unique (C-scan). Les signaux US constituent des données beaucoup plus riches, et lourds en termes de taille, qu'un simple C-scan. Ainsi, afin de tirer parti de toute l'information disponible, les signaux doivent être prétraités avant d'appliquer des techniques d'extraction de caractéristiques. Nous avons montré la fusion de données multi-physiques en combinant les signaux CF et US pour améliorer les performances d'inversion. De même, les stratégies LBE développées peuvent être adoptées pour effectuer la fusion de données avec d'autres méthodes CND telles que la thermographie.

Un des principaux objectifs de cette thèse était de développer une stratégie d'inversion en temps réel applicable aux cas industrielles réalistes. Après avoir obtenu un modèle d'apprentissage précis grâce à la méthode LBE, il pourrait être introduit dans des dispositifs portables (par exemple, des équipements d'inspection) pour fournir des fonctionnalités de prédiction en temps réel.

Titre: Développement d'une méthodologie robuste d'inversion dédiée au CND par courants de Foucault.

Mots clés: Inversion, Contrôle non destructif, L'apprentissage par l'exemple, Apprentissage automatique.

Résumé: Ce travail de thèse porte sur l'étude et le développement de stratégies innovantes pour la résolution, basée sur l'utilisation de la simulation et de la théorie de l'apprentissage statistique, de problèmes inverses dans le domaine contrôle non destructif (CND) par méthodes électromagnétiques. L'approche générale adoptée consiste à estimer un ensemble des paramètres inconnus, constituant un sous-ensemble des paramètres décrivant le scénario de contrôle étudié. Dans les cas de CND, les trois applications classiquement visées sont la détection, la localisation et la caractérisation de défauts localisés dans le matériau inspecté. Ce travail concerne d'une part la localisation et la caractérisation des fissures et d'autre part l'estimation de certains paramètres de sonde difficiles à maîtriser ou inconnus. Dans la littérature, de nombreuses méthodes permettant de remonter aux paramètres inconnus ont été étudiées. Les approches d'optimisation standard sont basées sur la minimisation d'une fonction de coût, décrivant l'écart entre les mesures et les données simulées avec un solveur numérique. Les algorithmes les plus répandus se fondent sur des approches itératives déterministes ou stochastiques. Cette thèse considère le problème de l'estimation de paramètres inconnus dans une perspective d'apprentissage statistique/automatique. L'approche supervisée adoptée est connue sous le nom de d'apprentissage par l'exemple (LBE en anglais). Elle se compose d'une première phase, dite hors ligne, pendant laquelle un « modèle inverse » est construit sur la base de la connaissance d'un ensemble de couples entrée/sortie connu, appelé ensemble d'entraînement. Une fois la phase d'apprentissage terminée et le modèle généré, le modèle est utilisé dans une phase dite en ligne pour prédire des sorties inconnues (les paramètres d'intérêt) en fonction de nouvelles entrées (signaux CND mesurés appartenant à un second ensemble dit de test) en temps quasi-réel. Lorsqu'on considère des situations pratiques d'inspection, en raison du grand nombre de variables impliquées, la création d'un modèle précis et robuste n'est pas une tâche triviale (problème connu comme la malédiction de la dimensionnalité). Grace à une étude approfondie et systématique, l'approche développée dans cette thèse a conduit à la mise en place de différentes solutions capables d'atteindre une bonne précision dans l'estimation des paramètres inversés tout en conservant de très bonnes performances en temps de calcul. Le schéma LBE proposé dans cette thèse a été testé avec succès sur un ensemble des cas réels, en utilisant à la fois des données synthétiques bruitées et des mesures expérimentales.

Title: Development of a robust inversion methodology in nondestructive eddy current testing.

Keywords: Inversion, Nondestructive testing and evaluation, Learning by example, Machine learning.

Abstract: The research activity of the PhD thesis focuses on the study and development of innovative strategies for the solution of inverse problems arising in the field of Non-Destructive Testing and Evaluation (NDT-NDE), based on the use of statistical learning theory. Generally speaking, the objective of the optimization stage is the retrieval of the unknown parameters within the studied electromagnetic scenario. In the case of NDT-NDE, the optimization problem, in terms of parameters to estimate, is divided into three stages, namely detection, localization and characterization. This work mainly addresses localization and characterization of crack(s) and/or estimation of probe(s) parameters. Unknown parameters, constituting a subset of the parameters set describing the electromagnetic scenario, are robustly estimated using several approaches. Standard optimization approaches are based on the minimization, by means of iterative approaches like stochastic and/or deterministic algorithms, of a cost function describing the discrepancy between measurements and prediction. This thesis considers the estimation problem in a machine learning perspective, adopting well known Learning-By-Example (LBE) paradigm. In a so-called offline phase, a surrogate inverse model is first fitted on a set of known input/output couples, generated through numerical simulations. Then, in a so-called online phase, the model predicts unknown outputs (the parameters of interest) based on new inputs (measured NDT signals) in quasi-real time. When considering practical inspection situations, due to the large number of variables involved (known as curse of dimensionality), obtaining an accurate and robust model is not a trivial task. This thesis carries out a deep and systematic study of different strategies and solutions to achieve simultaneously good accuracy and computational time efficiency in the parameters estimation. Moreover, a particular emphasis is put on the different approaches adopted for mitigating the curse of dimensionality issue. The proposed LBE schema has been tested with success on a wide set of practical problems, using both synthetic noisy data and experimental measurements.