Università degli Studi di Napoli "Federico II"

SCUOLA POLITECNICA E DELLE SCIENZE DI BASE
DIPARTIMENTO DI INGEGNERIA INDUSTRIALE

TESI DI LAUREA IN INGEGNERIA AEROSPAZIALE

CLASSE DELLE LAUREE MAGISTRALI IN INGEGNERIA AEROSPAZIALE E ASTRONAUTICA
(LM 20)

Characterization of Lamb waves behavior in CFRP pseudo-damaged configurations for SHM applications

Relatore:
Ernesto Monaco

Correlatori:
Vittorio Memmolo
Tim Behrens
Maria Boix-Bonet
Peter Wierach

Candidate:
Michele Mormone

Matricola:
XXXXXX

Anno accademico 2025-2026

Abstract

The structural framework of modern aircraft is subject to some of the most demanding engineering constraints, balancing the imperative of uncompromising safety with the need for high operational readiness. In pursuit of lighter, more efficient airframes, aerospace designers have turned extensively to carbon fiber-reinforced polymers (CFRPs), whose superior mechanical properties per unit weight have revolutionized structural design. However, this advancement comes at a cost: CFRPs do not fail like conventional metals. Their layered, anisotropic nature gives rise to subtle and often hidden damage mechanisms. Chief among these is barely visible impact damage (BVID)—a phenomenon where an apparently intact surface conceals internal disruption that may critically undermine the structural integrity over time.

To meet the growing demand for continuous, reliable assessment of such complex materials, Structural Health Monitoring (SHM) has emerged as a frontier field, with guided ultrasonic waves—particularly Lamb waves—offering a non-intrusive route to damage interrogation. These waves, by virtue of their ability to travel long distances and remain sensitive to internal inhomogeneities, are well-suited for inspecting CFRP components. Yet a persistent obstacle hinders progress: the high cost and destructive nature of creating true damage scenarios, which are essential for validating SHM methodologies. As a response to this constraint, the idea of pseudo-damages has taken shape—engineered disturbances designed to mimic real structural defects while preserving the functionality of the material.

This thesis investigates the interaction of Lamb waves throught such controlled, reversible damage proxies in order to assess the possibility to reliably emulate the ultrasonic signature of BVID, thereby enabling more scalable and cost-effective SHM development by validating pseudo-damages as faithful surrogates for real impact-induced damage. This is achieved within a controlled experimental framework in which Lamb wave acquisitions are systematically performed over a designed set of pseudo-damage configurations—varying material type, defect size, and adhesive tecnique—while maintaining consistent excitation and sensing conditions. The resulting dataset is then processed using two complementary approaches: classical inference methods, which rely on the explicit extraction of physics-based features from the measured signals and convolutional neural networks (CNNs), which take the minimally pre-processed acquisition data as input, learning discriminative representations of the pseudo-damage response directly from the measurements without manual feature engineering.

Contents

Chapter 1

The Idea of Pseudo-damages

1.1 Introduction

In the context of aeronautics, structures play a role of significative importance: high demands are imposed to these structures with respect to both safety and operational availability, while, on the other hand, low weight prerequisites for designs lead to the usage of lightweight, tailored and complex new composite materials. From those driver specifications and connecting to the new, modern and often embedded systems, a possibility to define a design in which reliability of structures connects with the weights requirements are represented by modern Structural Health Monitoring (SHM) approaches.

In the industry, modern composite materials are applied for the design of damage tolerant structures with a high specific load-bearing capacity, which is connected to the increasing implementation of carbon fiber-reinforced plastics (CFRPs)[1]. CFRP is deployed for structures such as thin-walled plates and shells, eventually reinforced by stringers and frames, and it is often complex in its layup because it is designed over the characteristics that the structural element has to accomplish. Not only complex because of the design but also complex by its nature, that lead to complex behavior in terms of damage: in contrast to structures made of classical construction materials such as steel and aluminum a low-energy impact strike on a CFRP component may result in barely visible impact damage (BVID) which is externally not visible but internally relevant. So a strike impact may implicate delaminations or cracks in the laminates that might spread out during operation and may cause, in the worst case, failure if not detected.
In this scenario, SHM systems able to investigate damages in CFRP aeronautical structures correspond to an incredibly promising field in term of reliability and cost reduction by presenting an alternative to the common strategy to prescribe maintenance as a function of number of hours of operation and not as a function of the usable residual life of the structural element. The mean by which SHM is able to derive information about the residual strength of aeronautical structures are ultrasonically guided waves, such as Lamb waves. Lamb waves are particularly noteworthy in the SHM context because they are low attenuating and they are highly sensible to the even slightly different mean internal characteristics variation. Using Lamb waves, SHM systems need to be able to verify the presence of the damage, localizing it and characterize it in

[1]The implementation of CFRP structural elements is alone accompanied by a huge mass reduction of airplanes in the range of 30% compared to the purely isotropic designs.

a often complex construction and in order to do so, the system must refer to plenty of different damaged configuration for the same construction. This clearly highlights one of the problems that is needed to face in order to obtain a valid SHM system: a cost-effective way to reproduce damaging over different and often complex structures.

In this context, the concept of pseudo-damages (PD) emerges as a promising solution: artificially crafted damage surrogates that are cost-effective, quasi reversible and that may be able to simulate real damage without compromising the structural integrity of the tested component. However, developing pseudo-damages as a viable alternative to actual structural damage presents significant challenges, primarily because their interactions with ultrasonic waves may differ from those of real defects. Since it is difficult to fully characterize wave behavior within pseudo-damages, a large experimental dataset is essential to evaluate their effectiveness in replicating the features of real damage.

This work investigates the feasibility of utilizing pseudo-damages to simulate BVID in CFRP plates by analyzing physics-based features derived from experimentally acquired data. In Chapter 2, the analytical background of Lamb waves is addressed, providing the reader with all the relevant information for the presented work. In Chapter 3, the experimental environment is presented to assess the interaction of Lamb waves with pseudo-damages, using this interaction as a benchmark to compare their wave behavior against that of real BVIDs. The experimental campaign, structured through a Design of Experiment (DoE) approach, the feature definition, and all constitute the core of the chapter and aim to meet the two essential criteria for the tested pseudo-damages: high reversibility and a strong correlation with the wave signatures observed in real impact damage. In Chapter 4, the feature derived from experimental data is presented and analyzed in relation to the crafting parameters of pseudo-damages. To elucidate the motivation behind defining these features ad hoc, a brief overview of the types of experimental data is provided. With the objective of assessing the similarity between Lamb modes interaction with BVID and with PD, Impact Testing is conducted, which serves as the core of Chapter 5. Subsequently, Chapters 6 and 7 present a mathematical framework within which the similarity assessment is conducted, utilizing the feature defined in Chapter 3 and generalizing it through a transfer learning approach applied to a pre-trained Convolutional Neural Network. Both approaches are employed to comprehend the DoE dataset and compare it to the BVID dataset and are closely related to the data analysis strategy used. The final Chapter serves as a concluding chapter, dedicated to future improvements and potential avenues for further research.

The result of this thesis open up exciting prospects for both research and industrial applications: for instance, gaining data over the defined DoE represent a source for producing more information about the behavior of Lamb waves through inhomogeneities. Furthermore, if proven reliable, pseudo-damages could serve as powerful tools for generating extensive and diverse training dataset for machine learning algorithms without the need to induce irreversible damage in costly aerospace components. This would address a significant bottleneck in the development of data-driven SHM systems, where the scarcity of real damage data often limits the generalization capabilities of trained models.
Moreover, the integration of pseudo-damages into SHM testbeds could enable controlled, repeatable, and customizable scenarios for the development and validation of SHM diagnostic

algorithms. This approach hold particular promise in accelerating the deployment of intelligent, real-time monitoring systems for composite airframes, where safety, data volume and test flexibility are critical. In the long term, the use of pseudo-damages may contribute to more robust and trustworthy AI models capable of distinguishing between various damage types and severities, enhancing predictive maintenance strategies and reducing the operational costs of modern aircraft fleets.

Chapter 2

Analytical Foundation of Lamb waves

2.1 Wave propagations in elastic solids

In case of an infinite elastic medium it is known that exactly two types of waves are capable to propagate: the compression waves and the shear waves, more commonly known as P-waves and S-waves. In the case of a semi-infinite or, as it is in reality, a finite media the existence of a boundary affects the wave propagation and a phenomenon of mode conversion may occurs when waves encounter a free boundary. This means that in the case of a S-wave and a P-wave both may be reflected. The solution of the respective mathematical boundary value problem is interpreted as a surface wave which is a third type of wave: this new kind of waves has been named Rayleigh waves, after Lord Rayleigh (1842-1919) who discovered them in 1885. Rayleigh waves are characterized by a high attenuation ratio through the dept of the materials and by the fact that they are not dispersive, i.e. the wave speed depends only on the medium characteristics.
Lamb waves have been first described in 1917 by the English mathematician Sir Horace Lamb (1849-1934)[1]. They belong to a type of waves which is commonly found in traction-free thin plate and shell-like structures. From the analytical solution of the governing equation one obtains that an infinite number of wave modes propagate in plates and that at least two modes exist for a fixed frequency, one symmetrical and another one asymmetrical. In SHM the fundamental symmetric (S0) and antisymmetric (A0) modes are typically generated without excitation of higher modes: those waves are generally dispersive and, if the waveguide is anisotropic, they are direction-dependent. Nevertheless they are able to propagate over long distances with high speed and low attenuation, making them really attractive for health monitoring techniques.

In this section the mathematical foundation of Lamb waves is addressed [2], in order to assess the characteristics of Lamb waves and their complex, thus exciting, behavior in different media. In particular a derivation and description of the dispersion relations for single and multiple-layered isotropic carbon fiber-reinforced plastics (CFRP) is presented. In particular the mathematical formulation, based on the Lamè-Navier equations, of wave propagation in thin-walled isotropic solids and their dispersive behavior is addressed, therefore leading to the definition of the dispersion relations. Finally, a procedure for the computation of the disper-

[1] As it is common for any kind of particular wave, they were named after him.

[2] The results, encapsulated in this chapter, are in their completeness written in [4]; here we review part of the mathematical background, useful for the purpose.

sion relations in multiple-layered anisotropic solids as well as their assets and drawbacks is presented.
The aim of this framework is to first present the behavior of Lamb waves and their propagation modalities inside different means and how to predict, using analytical and numerical tools, their behavior. On the other hand, this framework is useful to assess the characteristics of the given set of CFRP plates, used as the mean for acquiring the experimental data.

The general formulation of the wave equation can be derived recalling the linearized strain-displacement equation (2.1), the generalized Hooke's law (2.2) and the balance of momentum equation (2.3):

$$\underline{\underline{\varepsilon}} = \frac{1}{2}\left(\underline{\underline{\nabla u}} + \underline{\underline{\nabla u}}^T\right) \tag{2.1}$$

$$\underline{\underline{\sigma}} = \underline{\underline{\underline{\underline{\mathbb{K}}}}} : \underline{\underline{\varepsilon}} \tag{2.2}$$

$$\underline{\nabla} \cdot \underline{\underline{\sigma}} + \rho \underline{\mathbf{f}} = \rho \underline{\ddot{u}} \tag{2.3}$$

Where (2.2) is reported in its general form, in which the stiffness tensor $\underline{\underline{\underline{\underline{\mathbb{K}}}}}$ is defined as a tensor of fourth order of $3^4 = 81$ different components, depending on the material properties. Emanating from symmetries of strain and stress tensors, the number of independent elastic parameters decreases from 81 to 36, enabling us to use the common Voigt notation:

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \tau_{12} \\ \tau_{23} \\ \tau_{13} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \gamma_{12} \\ \gamma_{23} \\ \gamma_{13} \end{pmatrix} \tag{2.4}$$

By using (2.2) and (2.1) inside the balance of momentum and thus neglecting the influence of the force per unit volume f we get the wave equation for a 3D material domain:

$$\underline{\nabla} \cdot (\underline{\underline{\underline{\underline{C}}}} : \underline{\underline{\nabla u}}) = \rho \cdot \underline{\ddot{u}} \tag{2.5}$$

2.2 Isotropic material

Assuming that the material properties corresponds to a linear elastic one, we can further more reduce the independent component of the elastic matrix from 36 to 2 and obtain the expression for C_{ij} as a function of Lamè costants λ and G:

$$C_{abcd} = \lambda\delta_{ab}\delta_{cd} + G\left(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}\right) \tag{2.6}$$

Where the two Lamè costants are related to the Poisson ratio ν and the Young modulus E:

$$\lambda = \frac{\nu \cdot E}{(1+\nu)(1-2\nu)}$$

$$G = \frac{E}{2(1+\nu)}$$

And thus the elastic matrix becomes:

$$\underline{\underline{C}} = \begin{pmatrix} \frac{E}{(1-\nu^2)} & \frac{\nu E}{(1-\nu^2)} & \frac{\nu E}{(1-\nu^2)} & 0 & 0 & 0 \\ \frac{\nu E}{(1-\nu^2)} & \frac{E}{(1-\nu^2)} & \frac{\nu E}{(1-\nu^2)} & 0 & 0 & 0 \\ \frac{\nu E}{(1-\nu^2)} & \frac{\nu E}{(1-\nu^2)} & \frac{E}{(1-\nu^2)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{E}{2(1+\nu)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{E}{2(1+\nu)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{E}{2(1+\nu)} \end{pmatrix}$$

Recalling (2.5) and using the new form granted by (2.6) we get the Lamè-Navier equations:

$$(\lambda + G)\underline{\nabla}(\underline{\nabla} \cdot \underline{u}) + G \cdot \underline{\nabla} \cdot (\underline{\underline{\nabla u}}) = \rho \cdot \underline{\ddot{u}} \tag{2.7}$$

That in its scalar form appears as:

$$\begin{cases} (\lambda+G)\dfrac{\partial}{\partial x}\left(\dfrac{\partial u}{\partial x}+\dfrac{\partial v}{\partial y}+\dfrac{\partial w}{\partial z}\right)+G\left(\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}+\dfrac{\partial^2 u}{\partial z^2}\right)=\rho\ddot{u}, \\ (\lambda+G)\dfrac{\partial}{\partial y}\left(\dfrac{\partial u}{\partial x}+\dfrac{\partial v}{\partial y}+\dfrac{\partial w}{\partial z}\right)+G\left(\dfrac{\partial^2 v}{\partial x^2}+\dfrac{\partial^2 v}{\partial y^2}+\dfrac{\partial^2 v}{\partial z^2}\right)=\rho\ddot{v}, \\ (\lambda+G)\dfrac{\partial}{\partial z}\left(\dfrac{\partial u}{\partial x}+\dfrac{\partial v}{\partial y}+\dfrac{\partial w}{\partial z}\right)+G\left(\dfrac{\partial^2 w}{\partial x^2}+\dfrac{\partial^2 w}{\partial y^2}+\dfrac{\partial^2 w}{\partial z^2}\right)=\rho\ddot{w}. \end{cases} \tag{2.8}$$

The Lamè-Navier equations are a system of coupled, second-order linear differential equations, whose closed-form solution with taking account boundary conditions is only feasible in particular cases, needing numerical methods generally.

2.2.1 Waves in infinite solids

A classical analytical solution of the Lamè-Navier equation is obtained in the case of a three-dimensional infinite domain: in those hypothesis the Helmholtz decomposition theorem is usable. According to Helmholtz theorem a vector field, as the displacement field $\underline{u}$, can be decomposed into a curl-free part and a solenoidal part, which in terms means that two potential functions ϕ and ψ can be employed such as:

$$\underline{u} = \underline{u}_{curl-free} + \underline{u}_{div-free} = -\underline{\nabla}\phi + \underline{\nabla} \times \underline{\psi}$$

Following the theorem of Schwarz that grants that $\underline{\nabla}(\underline{\nabla} \cdot (\underline{\nabla}\phi)) = \underline{\nabla} \cdot (\underline{\nabla}(\underline{\nabla}\phi)$ and considering that the $\underline{\nabla} \cdot (\underline{\nabla} \times \underline{\psi})) = 0$ Eq. (2.7) becomes:

$$(\lambda + 2\mu)\underline{\nabla}(\underline{\nabla} \cdot (\underline{\nabla}\phi)) - \rho \cdot (\underline{\nabla}\ddot{\phi}) + G \cdot \underline{\nabla} \cdot (\underline{\nabla}(\underline{\nabla} \times \underline{\psi})) - \rho(\underline{\nabla} \times \ddot{\underline{\psi}}) = 0 \tag{2.9}$$

This equation yields for every point in the infinite domain we are considering, thus we get two sub-equations that must be satisfied for the potentials functions ϕ and $\underline{\psi}$:

$$\begin{aligned} \nabla \cdot (\underline{\nabla}\phi) &= \frac{1}{c_L^2}\ddot{\phi} \\ \underline{\nabla} \cdot (\underline{\nabla}\psi) &= \frac{1}{c_t^2}\ddot{\underline{\psi}} \end{aligned} \tag{2.10}$$

where the propagation velocities c_L and c_T are defined as:

$$c_L = \sqrt{\frac{\lambda + 2G}{\rho}} \quad \text{and} \quad c_T = \sqrt{\frac{G}{\rho}}$$

We found the equation from which compression and shear waves arises: from a mathematical stand point, it is interesting to note that the P-waves are well described by a single scalar function, which is consistent with the fact that compression waves generates displacement in the direction of propagation of the wave as compression and rarefaction of the medium; S-waves instead need a vector field because they introduce rotational motion in the media, physically described by the fact that S-waves are not parallel to the propagation direction.
This fundamental difference explains why P-waves typically travel faster than the S-waves and why only S-waves can propagate inside a fluid media, which is because a fluid cannot resist shear forces without deforming indefinitely.

2.2.2 Waves in thin plates

In the case of propagation of waves inside plate-like solids, the solution of the Lamè-Navier equations are obtained by including the boundary conditions imposed by the free surfaces. The interaction of the compression and shear waves with the boundaries causes reflections and thus a generation of an additional wave type called Lamb wave.

Whereas there are just two types of waves into an infinite solid, into a finite dimension thin plate an infinite number of Lamb waves arises.

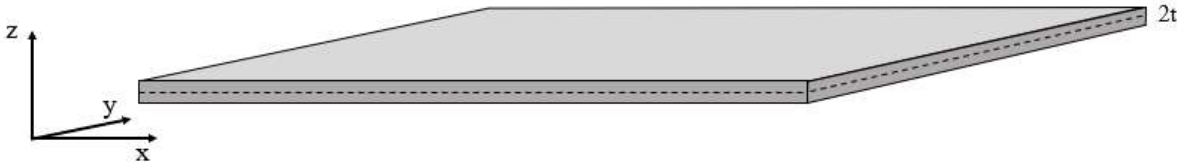


Figure 2.1: thin plate-like 3D dominium

If we use the classical assumption of thin plates, we can image them as infinitely extended plate (into the x-y plane) having a thickness of $2t$ into the z direction. As said for infinite solid, only trasversal waves and parallel waves can arise, with a movement that depends only onto x or y and z coordinates. The following mathematical derivation is done by studying thew cross sectional view of the plate into the x-z plane. The infinite dimension in the x-y plane and the consideration that waves into solids are related only to 2 of the three dimensions, a plain strain state is assumed.
As done into the case of infinite solids, the starting point is to express the displacement field according to Helmholtz decomposition theorem:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\partial \phi}{\partial x} \\ \frac{\partial \phi}{\partial y} \\ \frac{\partial \phi}{\partial z} \end{pmatrix} + \begin{pmatrix} \frac{\partial \psi_3}{\partial y} - \frac{\partial \psi_2}{\partial z} \\ \frac{\partial \psi_1}{\partial z} - \frac{\partial \psi_3}{\partial x} \\ \frac{\partial \psi_2}{\partial x} - \frac{\partial \psi_1}{\partial y} \end{pmatrix} \tag{2.11}$$

From the assumption that the displacement is y-indipendent, all the $\frac{\partial}{\partial y}$ derivatives vanish. Thus (2.11) becomes:

$$u(x,z,t) = \frac{\partial \phi}{\partial x} - \frac{\partial \psi_2}{\partial z} \tag{2.12}$$

$$v(x,z,t) = \frac{\partial \psi_1}{\partial z} - \frac{\partial \psi_3}{\partial x} \tag{2.13}$$

$$w(x,z,t) = \frac{\partial \phi}{\partial z} + \frac{\partial \psi_2}{\partial x} \tag{2.14}$$

The displacements components u,w are dependent to ϕ and ψ_2, where instead the displacement v is non dependent for them, so the displacement field components are decoupled in terms of x-z plane and y-direction and they will be considered separately. While (2.13) indicate the horizontal shear motion of the plate, the expression for the u-w displacements in (2.12-2.14) contain the term related to the pressure wave (P-waves from ϕ) and the vertical shear waves (SV-waves from ψ_2). The displacement field obtained by the summation of S and P waves leads to Lamb waves. Hence, the further derivation only includes the displacement in the x-z plane and for reason of clarity it is assumed that $\psi_2 = \psi$. So the displacements become:

$$u(x,z,t) = \frac{\partial \phi}{\partial x} - \frac{\partial \psi}{\partial z} \tag{2.15}$$

$$w(x,z,t) = \frac{\partial \phi}{\partial z} + \frac{\partial \psi}{\partial x} \tag{2.16}$$

An accademic approach to resolving D'Alembert differential problem [3] is to assume the separation of variables. In this case the details of the technique will not be presented, but the resulting expression for ϕ and ψ are the following:

$$\phi = \Phi(z) \cdot e^{i(kx-wt)} \tag{2.17}$$

$$\psi = \Psi(z) \cdot e^{i(kx-wt)} \tag{2.18}$$

The factor Φ and Ψ refer to stationary waves into perpendicular direction, wether the imaginary exponent refers to the wave propagation into the x direction. With the separation technique the (2.7) solution read as follows:

$$\Phi = A_1 sin(pz) + A_2 cos(pz) \tag{2.19}$$

$$\Psi = B_1 sin(qz) + B_2 cos(qz) \tag{2.20}$$

with:

$$p^2 = \frac{w^2}{c_L^2} - k^2, \qquad q^2 = \frac{w^2}{c_T^2} - k^2 \tag{2.21}$$

And therefore the in-plane u and out-of-plane displacements w become:

$$\begin{aligned} u &= \left(ik\left[A_1 \sin(pz) + A_2 \cos(pz)\right] - \left[qB_1 \cos(qz) - qB_2 \sin(qz)\right]\right) e^{i(kx-\omega t)} \\ w &= \left(\left[pA_1 \cos(pz) - pA_2 \sin(pz)\right] + ik\left[B_1 \sin(qz) + B_2 \cos(qz)\right]\right) e^{i(kx-\omega t)} \end{aligned} \tag{2.22}$$

The solution in terms of the displacement can be decomposed taking into account that sine and cosine are respectively odd and even functions of z: that leads to a net distiction of two wave modes, respectively representing the symmetrical and antisymmetrical component of Lamb waves. Therefore for the symmetric mode, one obtains:

$$\begin{aligned} u_{sym} &= [i \cdot kA_2 \cos(pz) - qB_1 \cos(qz)] e^{i(kx-\omega t)} \\ w_{sym} &= [-p \cdot A_2 \sin(pz) + ikB_1 \sin(qz)] e^{i(kx-\omega t)} \end{aligned} \tag{2.23}$$

And in a similar fashion the antisymmetrical mode can be obtained as well:

$$\begin{aligned} u_{anti} &= [i \cdot kA_1 \sin(pz) + qB_2 \sin(qz)] e^{i(kx-\omega t)} \\ w_{anti} &= [+p \cdot A_1 \cos(pz) + ikB_2 \cos(qz)] e^{i(kx-\omega t)} \end{aligned} \tag{2.24}$$

[3] A typical approach in mathematics to relate to wave propagation problems is the use of D'Alembert equation: this is a second-order linear differential equation that models many physical phenomena in nature, including the vibration of a a string in one dimension, a thin membrane in two dimensions and pressure fluctuation in a three-dimensional acoustic volume.

At this point, we obtained the general solution of the thin plate wave problem in terms of a set of functions u and w, all possible varying A_i and B_i: in order to identify which function is the real solution of the problem, we need to take into account boundary conditions. In case of a plane strain state,we can recall (2.1), (2.2) and (2.3) to express the tension in the boundaries parallel to the x-z plane:

$$\begin{aligned} \sigma_{31} &= G(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}) \\ \sigma_{33} &= G(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}) + (\lambda + 2G)\frac{\partial w}{\partial z} = G\frac{\partial u}{\partial x} + (\lambda + 2G)\frac{\partial w}{\partial z} \end{aligned} \tag{2.25}$$

Substituting into those equations (2.24-2.27) leads to the stress equations of the symmetric waves:

$$\begin{aligned} \sigma_{31}^{\text{sym}} &= G\left[-2ikpA_2\sin(pz) + \left(q^2 - k^2\right)B_1\sin(qz)\right]e^{i(kx-\omega t)} \\ \sigma_{33}^{\text{sym}} &= G\left[-\left(q^2 - k^2\right)A_2\cos(pz) + 2ikqB_1\cos(qz)\right]e^{i(kx-\omega t)} \end{aligned} \tag{2.26}$$

and the antisymmetric ones:

$$\begin{aligned} \sigma_{31}^{\text{anti}} &= G\left[2ikpA_1\cos(pz) + \left(q^2 - k^2\right)B_2\cos(qz)\right]e^{i(kx-\omega t)} \\ \sigma_{33}^{\text{sym}} &= G\left[-\left(q^2 - k^2\right)A_1\sin(px_3) - 2ikqB_2\sin(qz)\right]e^{i(kx-\omega t)} \end{aligned} \tag{2.27}$$

On the free surfaces, we have:

$$\sigma_{31}(x, z = \pm t) = \sigma_{33}(x, z = \pm t) = 0$$

imposing this condition leads to the linear, homoegeneous equation system for symmetrical and antisymmetrical stresses:

$$\begin{bmatrix} -2ikp\sin(pd) & (q^2 - k^2)\sin(qd) \\ -(q^2 - k^2)\cos(pd) & 2ikq\cos(qd) \end{bmatrix}\begin{bmatrix} A_2 \\ B_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.28}$$

$$\begin{bmatrix} 2ikp\cos(pd) & (q^2 - k^2)\cos(qd) \\ -(q^2 - k^2)\sin(pd) & -2ikq\sin(qd) \end{bmatrix}\begin{bmatrix} A_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.29}$$

Resolving the system with non trivial solution requires that the matrices in (2.28) and (2.29) have a determinant which is equal to zero. This condition provides the formulae for the determination of the symmetric and antisymmetric angular wavenumber k_{sym} and k_{anti}, commonly know as Rayleigh-Lamb wave equations:

$$\frac{\tan(qd)}{\tan(pd)} = \left[-\frac{4k^2pq}{(q^2 - k^2)^2}\right]^{\pm 1} \tag{2.30}$$

Where the exponent can be $+1$ or -1, depending on the evaluation of k_{sym} and k_{anti}. These equations have solutions for every angular frequency w, hence the phase velocity c_p may be calculated for a given pair of angular excitation frequency w and the related wavenumber k by the following relation:

$$c_p = \frac{w}{k}$$

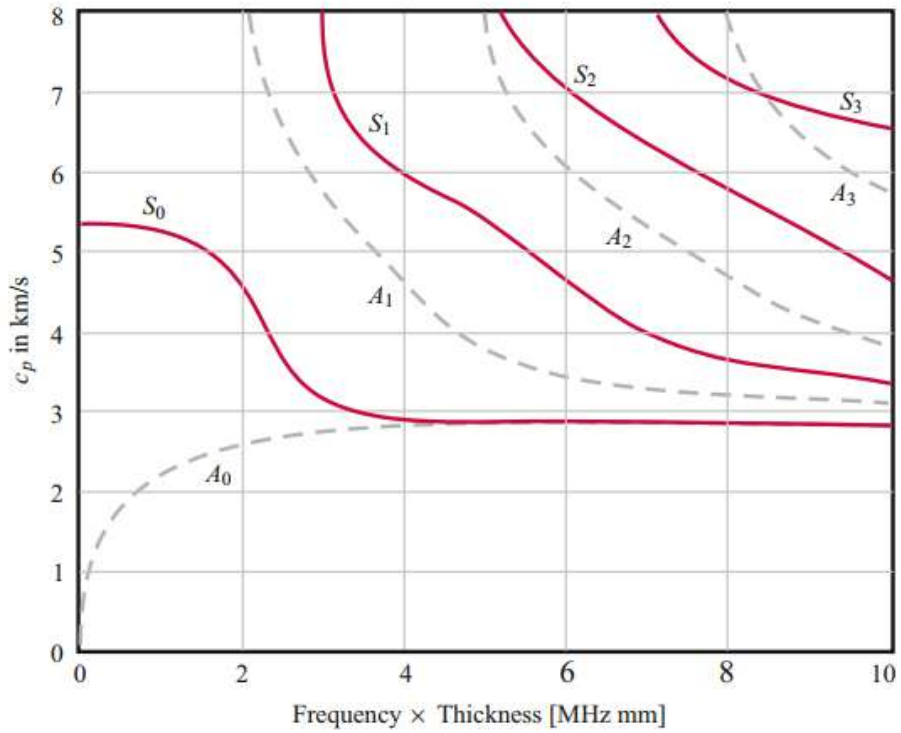


Figure 2.2: Dispersion curves for an aluminum plate (E = 70 GPa, $\nu = 0.33$, $\rho = 2700 kg/dm^3$) - plot from [4]

Thus, for every frequency and wavenumber we get a velocity c_p: the behavior of c_p for every k,w is denoted as dispersion curves.
Calculating dispersion curves is mandatory in the context of laboratory experiments involving wave propagation. Dispersion curves show how wave velocity depends on frequency and wavelength, providing essential information about how different wave modes travel through the defined mean. By understanding the dispersion curves, it is possible to identify frequency ranges where specific wave modes propagate with higher velocities than other modes. This leads to clearer signals and thus accurate interpretation of the wave behavior. On the other hand, Eq. (2.30) cannot be solved analitically, even in the case of isotropical mean.

The calculation of the angular wave number enable a qualitative representation of the displacement field over the thickness of the plate: The unknowns A_2 and B_1 (as well as A_1 and B_2) are pairwise linearly dependent and serve as a scaling factor for the trigonometric parts of the functions. For instance one can assume that:

$$B_1 = \frac{2ikp\sin(pd)}{(q^2 - k^2)\sin(qd)} A_2 \tag{2.31}$$

$$B_2 = -\frac{2ikp\cos(pd)}{(q^2 - k^2)\cos(qd)} A_1 \tag{2.32}$$

with the coefficients A_i freely selectable.

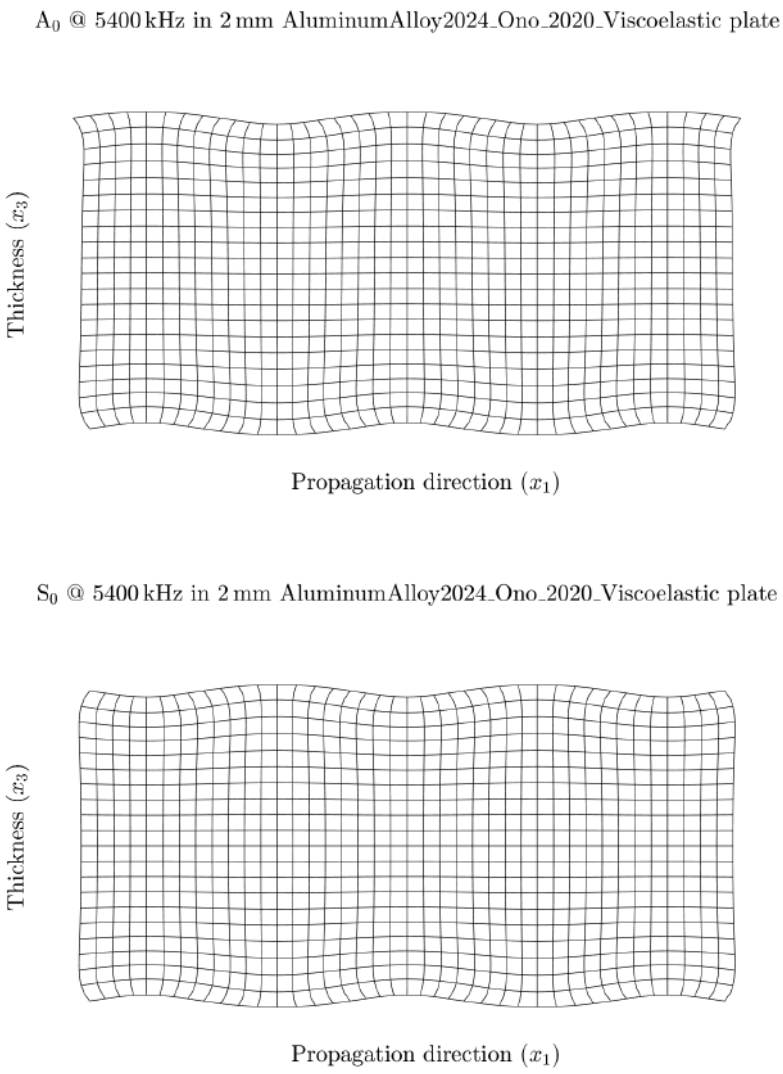


Figure 2.3: fundamental Lamb modes A0 and S0 for an aluminum plate (E = 70 GPa, $\nu = 0.33$, $\rho = 2700 kg/dm^3$) for a frequency of 5400 KHz - Evaluated via Dispersion Calculator (DC) - DLR

At this point, important insight of the complexity of Lamb modes can be appreciated: during the mathematical foundation, it has been demonstrated that there are a numerable infinite[4] of Lamb modes propagating inside a thin plate, each of them can be symmetrical or antisymmetrical. Furthermore, we demonstrated their dispersive characters: for each frequency there is a particular phase velocity, for each mode, and the locus of the phase velocities for each modes are called dispersion curves. Until here there are no anisotropies that can lead to their direction dependency because we assumed the isotropy of the mean: clearly this will not be the case for the CFRP plates, beside the fact that they are built of a quasi-isotropical layup.

[4]The solution for an eigenvalue problem such as the one represented by the D'Alembert equation we resolved are numerable infinites, in particular they are of cardinality $\aleph^0$.

2.3 Anisotropic material

The study of guided waves in plate-like structures traditionally distinguishes between Lamb waves and horizontal shear (SH) waves. While these wave types are naturally decoupled in standard isotropic plates, the introduction of structural anisotropy, such as in multi-layered composite materials like CFRP elements, fundamentally complicates this behavior.

Depending on the propagation medium and the wave's orientation relative to the material's principal axes, Lamb and SH waves can become strongly coupled[5]. The fundamental approach to modeling this complex wave behavior begins with the governing the wave equations and the definition of a plane wave solution, which ultimately leads to the formulation of the Christoffel equation.

For a given anisotropic medium, substituting the harmonic plane wave displacement solution into the equations of motion yields the Christoffel eigenvalue problem:

$$(\Gamma_{ik} - \rho v^2 \delta_{ik}) U_k = 0 \tag{2.33}$$

where ρ is the material density, v is the phase velocity, δ_{ik} is the Kronecker delta, U_k are the polarization vectors, and Γ_{ik} is the Christoffel stiffness tensor defined as:

$$\Gamma_{ik} = C_{ijkl} n_j n_l \tag{2.34}$$

Here, C_{ijkl} represents the elastic stiffness tensor of the material and n_j, n_l are the direction cosines of the wave normal. The off-diagonal terms of Γ_{ik} are responsible for the coupling between longitudinal and shear displacements, dictating whether Lamb and SH waves can be treated separately or must be analyzed as a fully coupled system.

Regarding the solvability of the Christoffel equation itself, it fundamentally represents a 3×3 eigenvalue problem. Because the resulting characteristic equation is a cubic polynomial in terms of v^2, exact analytical solutions for the bulk wave velocities theoretically exist via standard algebraic methods. Furthermore, when wave propagation occurs along the principal axes of symmetry in orthotropic materials, specific off-diagonal terms of Γ_{ik} vanish. This decouples the wave modes, yielding very simple, closed-form exact solutions for pure longitudinal and pure shear bulk waves.

However, for arbitrary propagation angles in generally anisotropic media, these analytical expressions become algebraically cumbersome. Numerical methods are therefore predominantly used in practice to efficiently compute these eigenvalues and eigenvectors for any given propagation angle. More importantly, it must be emphasized that while the bulk wave characteristics of a single infinite layer can be solved analytically, the subsequent boundary-value problem for guided waves in a finite, layered structure cannot. The global system matrices yield complex transcendental equations. Finding the roots of these dispersion equations to predict guided wave modes requires robust numerical root-finding algorithms to enforce stress and displacement continuity across all layer interfaces, as well as the traction-free boundary conditions at the upper and lower surfaces of the plate.

Three primary matrix-based methods are traditionally utilized to assemble these conditions:

[5]In modern literature, the term "Lamb waves" is frequently generalized to encompass these complex, coupled guided waves in both single and multi-layered anisotropic plates.

- Transfer Matrix Method (TMM): Originally developed by Thomson and Haskell, the TMM constructs a state vector V comprising displacements and interlaminar stresses and relates the state at the top interface of a layer to the bottom interface using a layer transfer matrix T_i:

$$V_i^{top} = T_i V_i^{bottom} \tag{2.35}$$

  By sequentially multiplying the transfer matrices of all individual layers, one can directly map the bottom surface of a laminate to its top surface. While mathematically elegant and computationally compact [6], TMM suffers from the notorious "large fd (frequency-thickness) problem": At high frequencies or in thick plates, the positive exponential terms e^{+kd} within the matrix grow disproportionately, causing catastrophic loss of precision and numerical instability during the root-finding process.

- Global Matrix Method (GMM): To resolve the instabilities of TMM, Knopoff introduced the GMM. Instead of propagating a state vector through the thickness, the GMM assembles the equations of motion for all layers, along with the boundary and continuity conditions, into a single, comprehensive global system:

$$GA = 0 \tag{2.36}$$

  where G is the global matrix and A contains the unknown partial wave amplitudes for all layers. This formulation eliminates the numerical instabilities associated with growing exponentials, ensuring absolute stability across all frequency ranges. However, because the dimension of G scales directly with the number of layers, the GMM is highly computationally expensive and memory-intensive for structures with many plies.

- Stiffness Matrix Method (SMM): Operating as an optimal middle ground, the SMM ,formalized by Rokhlin and Wang, reformulates the boundary value problem by relating the stresses at the layer interfaces directly to the displacements, analogous to stiffness methods used in finite element analysis. A recursive algorithm is employed to assemble the global stiffness matrix from the individual layer matrices. The SMM effectively circumvents the numerical instabilities inherent in the TMM while strictly avoiding the massive matrix dimensions and computational overhead of the GMM. Consequently, it has become a highly preferred, stable, and efficient choice for modeling complex anisotropic laminates.

The ultimate goal of employing methods like TMM, GMM, or SMM is to calculate the dispersion relations of the structure. By setting the determinant of the assembled system matrix to zero, one can find the roots corresponding to the allowable wave numbers for a given frequency.

Because anisotropic materials like CFRP are highly dispersive, the phase and group velocities of the guided waves change non-linearly with frequency and propagation direction. Predicting these characteristics is paramount,as will be shown in short, as it allows engineers to select optimal frequencies and modes that are sensitive to specific defects while minimizing signal attenuation and multi-modal overlapping.

To bridge the gap between complex theoretical root-finding and practical application, software tools such as Dispersion Calculator (DC) are heavily utilized. Dispersion Calculator

[6] the final system matrix size remains independent of the number of layers

provides a robust computational environment specifically designed to solve the dispersion equations for isotropic and multi-layered anisotropic plates. By leveraging stable matrix formalisms (often SMM or stable GMM variants) internally, DC bypasses the need for the user to code complex root-tracing algorithms from scratch. It directly generates the phase velocity, group velocity, and attenuation curves for a given CFRP lay-up. This facilitates a direct and visual understanding of the multi-modal wave propagation characteristics, drastically accelerating the theoretical pre-analysis required for experimental CFRP inspection.

2.4 CFRP plates

To facilitate a comprehensive study of guided wave interactions with various pseudo-damage configurations, a series of 10 CFRP plates was utilized. Each plate was fabricated with the same dimensions, namely 66x50 cm, and they share the same symmetric and quasi-isotropic layer sequence of $[0, \pm 45, 90]_s$, for a total thickness of 2.18 mm. This layup is commonly employed in aerospace structures due to its balanced mechanical properties, providing uniform stiffness and strength in multiple in-plane directions. Such a configuration is particularly advantageous for investigating the interaction of lamb wave with the pseudo-damage, as it supports the transmission of guided waves across diverse fiber orientations, thereby offering a representation that does not need to attack the direction-dependencies typically exhibited by Lamb waves propagating in non-isotropic means.

Figure 2.4: set of 10 CFRP plates 66x50 cm, layup $[0, \pm 45, 90]_s$, thickness of 2.18 mm.

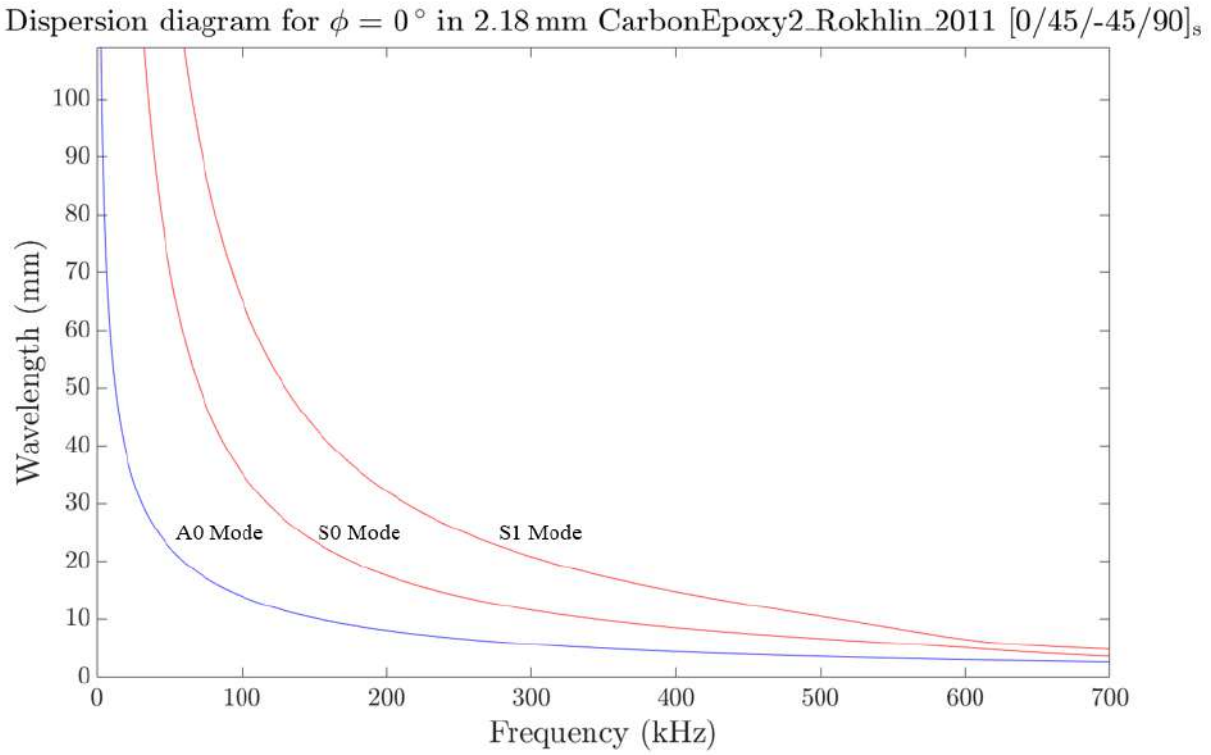
2.4.1 Dispersion Calculator ®

Prior to the experimental phase, an in-depth analysis of wave propagation behavior in the selected laminate was conducted by computing its dispersion characteristics using the Dispersion Calculator ® (DC) software. DC is an interactive and fully validated stand-alone software [2] for the computation of dispersion curves (phase velocity, energy velocity, attenuation, etc. vs frequency) and mode shapes (displacement, stress, strain, etc.) of guided waves in isotropic plates, rods and pipes (axial and circumferential waves) as well as in multilayered anisotropic plates. The material(s) can be perfectly elastic or viscoelastic, and the waveguide can be surrounded by and, in case of a pipe, be filled with inviscid fluids. DC can handle laminates consisting of several hundreds of layers, and it is able to distinguish the different mode families, like symmetric, anti-symmetric, and non-symmetric (leaky) Lamb, shear horizontal, and Scholte waves, depending on the symmetry and coupling properties of a given layup. DC features a highly efficient and robust dispersion curve tracing algorithm.DC is an open-source, MATLAB-based tool developed by Armin Huber and the German Aerospace Center (DLR), initially released in November 2018.
As an overview of what DC is capable of, a brief description of the eight different tabs is listed below:

- Isotropic: Calculation of dispersion diagrams and mode shapes for single layer isotropic plates, rods, and pipes (axial and circumferential waves)(fast computation by using the Rayleigh-Lamb equations, respectively, those specific to isotropic cylinders).
- Anisotropic: Calculation of dispersion diagrams and mode shapes for multilayered, strongly anisotropic plates using Stiffness Matrix Method and Trasfer Matrix Method [7].
- Signal simulator: Simulation of the signal of single and multiple guided wave modes (for plates only).
- Polar diagrams: Calculation of polar dispersion diagrams for multilayered anisotropic plates.
- Calculation of the bulk waves' phase and group velocities, slownesses, and polarization, propagating in anisotropic media as well as their scattering on solid-fluid interfaces.

The dispersion analysis provided by deploying DC gives critical insights into the behavior of fundamental Lamb wave modes inside the defined CFRP plates, expecially for the fundamental modes A0 and S0. For instance, reported below are both the dispersion diagrams for wavelenghts and phase velocities and the polar diagrams of the different Lamb modes propagating throught the CFRP plates.

[7]The Dispersion Calculator implements two numerical schemes for evaluating Lamb wave dispersion in anisotropic media: the Stiffness Matrix Method (SMM), which formulates the problem via the global stiffness matrix of the layered waveguide, and the Transfer Matrix Method (TMM), which propagates stress–displacement state vectors through each layer to enforce boundary conditions. For a full description of their implementation, see [2]



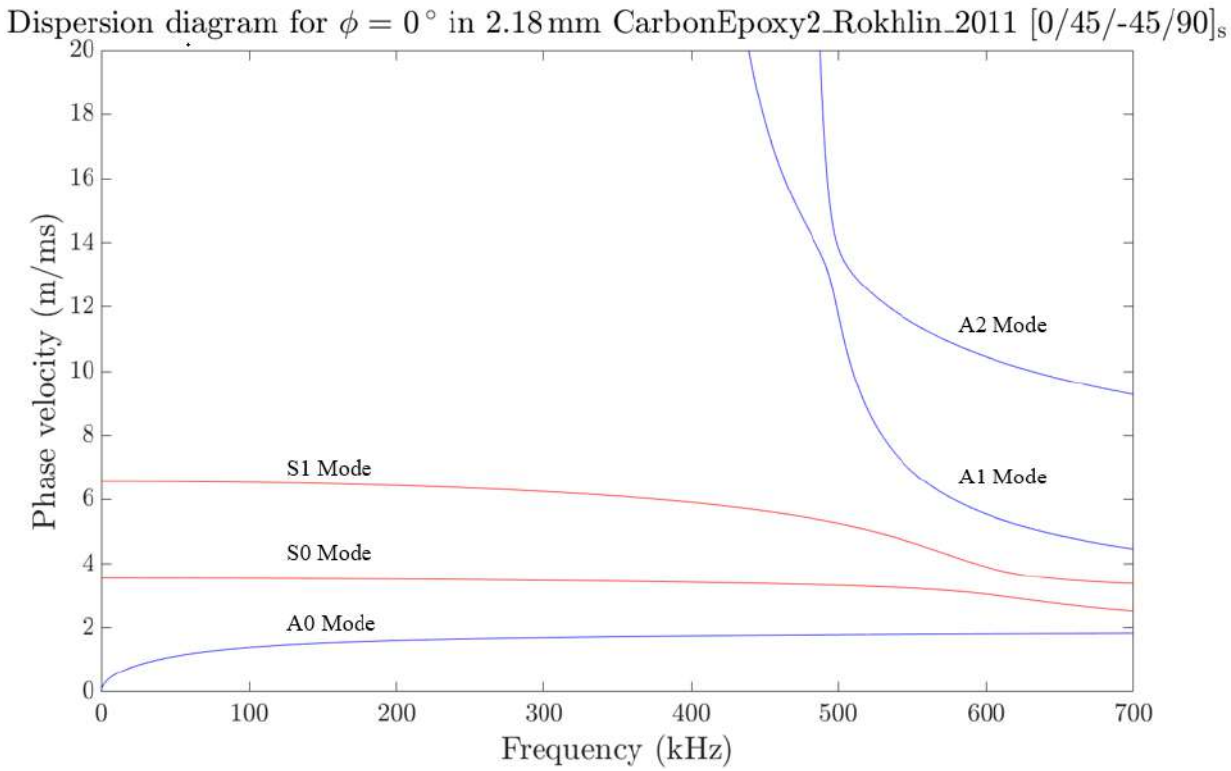


Figure 2.5: Dispersion relations for CFRP plates

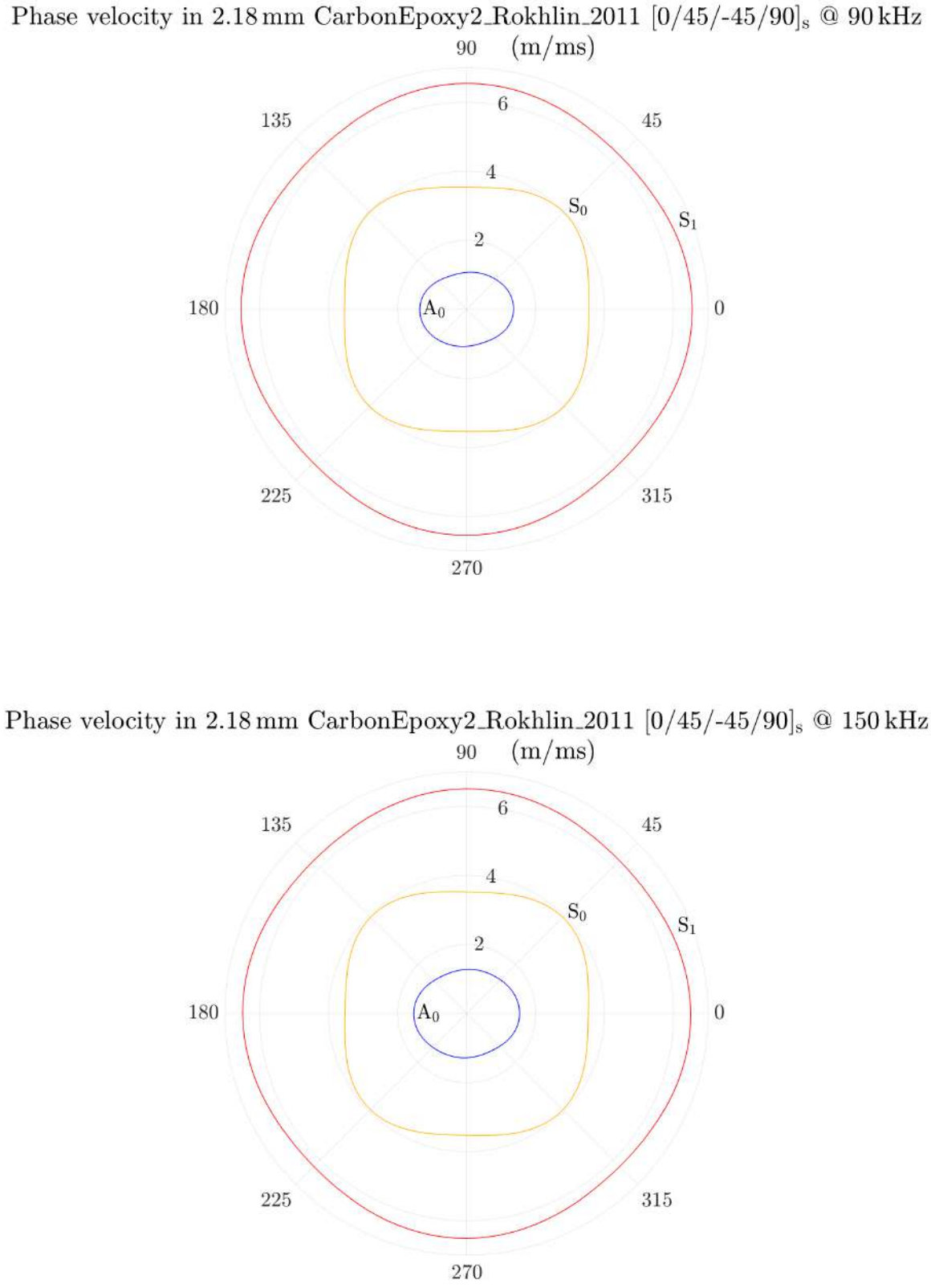


Figure 2.6: Polar Dispersion relations for CFRP plates - 90 KHz and 150 KHz.

Chapter 3

The experimental campaign

3.1 The experimental setup

As noted in the last chapter, depending on the excited frequency range, the amplitudes of the individual modes vary and one is dominant over the others. In order to enable a good characterization of the wave interactions, a single measurement for each fundamental mode is carried out. The investigation of Lamb wave interactions with different pseudo-damaged configurations requires an experimental arrangement capable of generating controlled wave excitations and capturing low noised responses, and ensuring precise synchronization between actuation and measurement. To meet these requirements, the entirety of the experiments was conducted at the Structural Health Monitoring (SHM) laboratory of the German Aerospace Center (DLR) in Braunschweig, Germany, operating within the Smart Structures (Systemleichtbau) department. The Experimental setup used in this framework is based on a technique called Ultrasonic Scanning Technique[1].
The ultrasonic scanning technique is an established Non-Destructive Testing (NDT) method for a wide range of materials and composites. An ultrasonic actuator is used for the excitation of dilatational waves field in specimen. The wave propagates through the specimen and interacts with its structure: any inhomogeneity in the structure influence the wave propagation. These can be cracks, impact damages, debondings, or even pores. The incident wave is partially reflected, part of the impacting energy is transmitted and mode conversion may occur. These interaction is coupled with a fluid between the specimen and the sensor. As coupling media in principle water, air, or other suitable fluids can be used. The interaction of the coupling fluid and the excited specimen is so acquired and processed.

In this context, the setup has to guarantee that: the measurements acquired from the sensor allow an accurate representation of the wave propagation without influencing the measurements; Furthermore, the different pseudo-damaged configurations must be analyzed in a way that is fairly simple, thus enabling the analysis of a large number of different cases; the excitation frequencies provided at the exciter need to highlight a single mode in order to characterize the interactions of the A0 and S0 modes without interference and, at last, the experimental data must be processable in a way that enables a proper understanding of the Lamb wave mode interactions.
Taking into account those requirements, the experimental setup is made up by:

[1] [4], cap. 5.2, page 94.

- An embedded piezo (DuraAct P-879K025) provides the excitation by means of a connection with a signal generator and a signal amplifier (WMA 300 Falco Systems);
- The signal generator (Tektronikx AFG 3022) generates a 3-sine burst of fixed frequency [2], opportunely amplified and transferred to the piezo.
- An ultrasonic probe (DeltaTron field 1/4" Microphone Type 4954A) is the actual sensor, acquiring the off-plane displacement [3] of the plane along a linear path.
- A CNC system, on which the rowing ultrasonic probe is acquiring the displacement.
- An acquisition system, which is the computer by which the experimental data is made available.

The excitation frequencies provided to the exciter needs to highlight just one mode, in order to characterize A0 and S0 interactions without interference: for the A0 mode the aim is to acquire the highest excitation frequency in order to have the smallest wavelenght[4] while keeping S0 mode negligible, for S0 the idea is to rise the excitation frequency in order to make it dominant. For this reason for the A0 mode the excitation frequency is defined to 90 KHz, for the S0 mode is defined to 150KHz.

The experimental setup is briefly schematized briefly down below in Fig. 3.1:

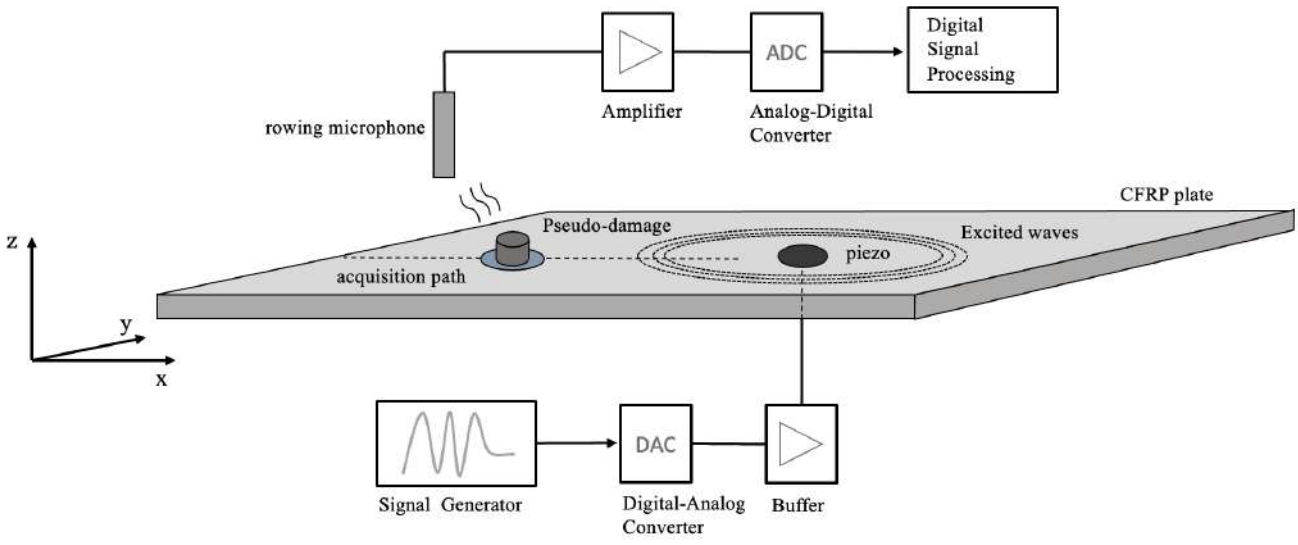


Figure 3.1: Logical scheme of the experimental setup

[2]Depending on the excitation of A0 and S0.

[3]Actually, Lamb waves propagate through the specimen, interact with its structure (and with the pseudo-damage) and radiate dilatational waves in the surrounding air: this is a limit for the technique taking into account that pressure waves into air are likely to attenuate, so the sensor need to be in the vicinity of the specimen.

[4]"The wavelength must be lower or at least comparable with the size of the defect", [5] pag 795.

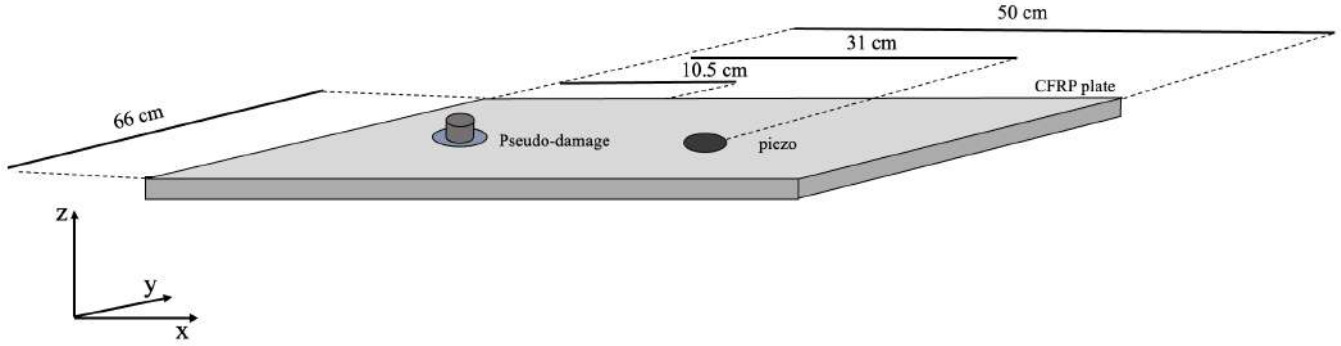


Figure 3.2: Geometry and measures of the experimental setup

Along the shown linear path in Fig. 3.2 an acquisition point is defined from the edge of the specimen for the acquisition with a resolution of 1 mm, and in an observation time windows of 0.4 microseconds at a sampling frequency of 1000 KHz the 311 different [5] amplitudes histories[6] are collected.

3.1.1 Imaging Method: B-Scan

The data in this environment can be displayed in different fashions: for instance, for each point along the linear path in the prescribed time windows an Amplitude vs Time, for this reason also called Amplitude-Scan (A-Scan) can be collected for both the experiments:

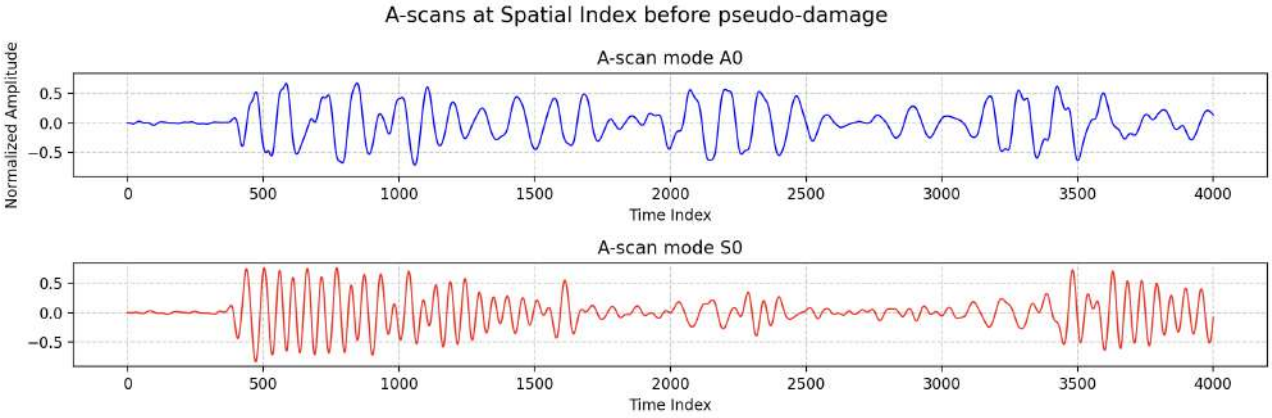


Figure 3.3

[5] 310 mm divided by our resolution + 1.
[6] Amplitudes for 4001 different time instants.

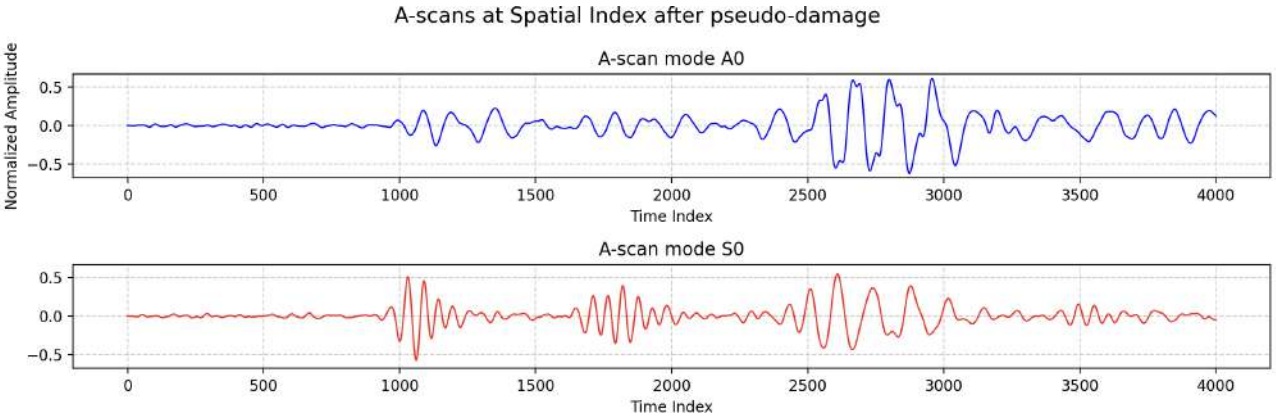


Figure 3.4

Figure 3.5: Comparison between A-scans of S0 and A0 modes before and after PD.

A-Scans of a scanning route can be arranged in an array by stacking along rows each A-Scan. The resulting imaging is called Brightness-Scan (B-scan), an array of dimensions (n.point along route) x (sampling instants), in this case 311 x 4001. In this framework, B-scan are used as a way to collect valuable information on the interaction of the 2 fundamental Lamb modes with the different pseudo-damages so, for each pseudo-damaged configurations, 2 B-Scans are collected.

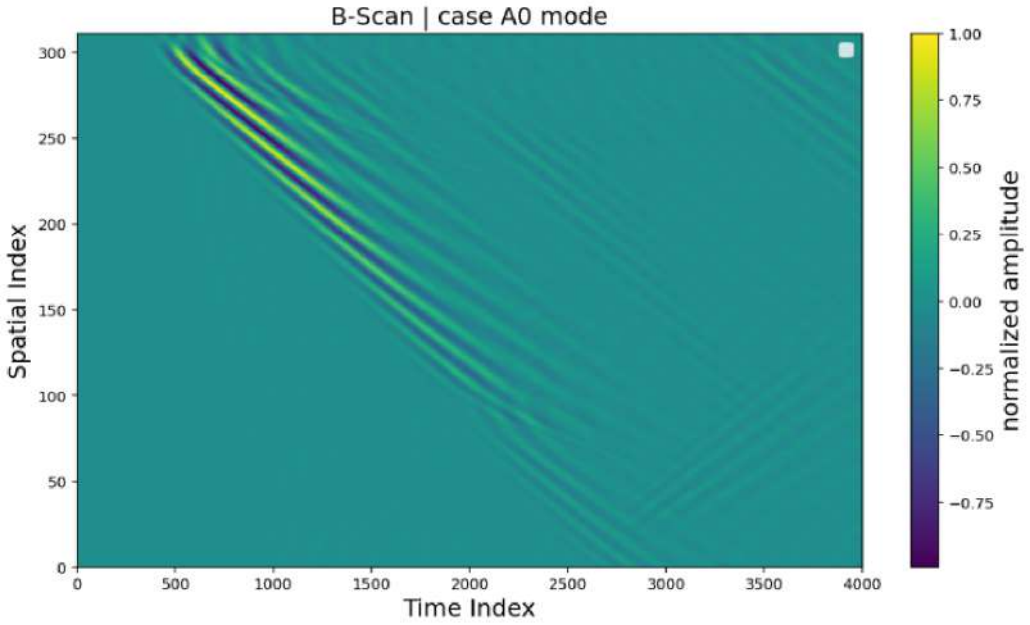


Figure 3.6: B-Scan for S0 mode.

B-scan imaging enables the viewer to comprehend the behavior of wave packets. As time progresses along the x-axis, it becomes evident that the two Lamb waves traverse the linear path from the exciter, interacting with the pseudo-damage and the aft boundary of the specimen. A closer examination of the interaction mechanisms with the pseudo-damage can be made by

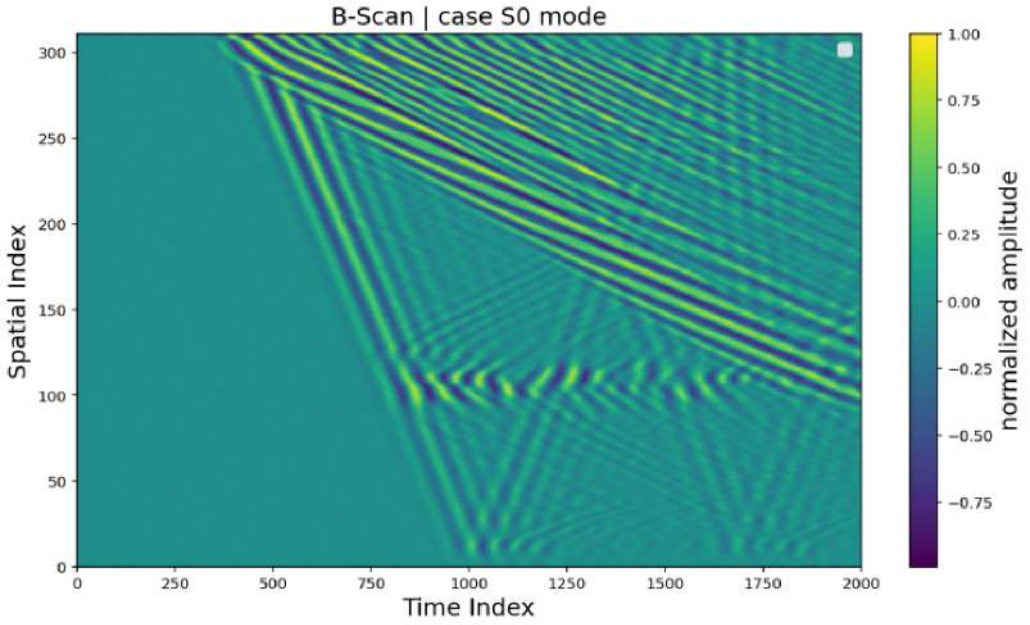


Figure 3.7

Figure 3.8: B-Scan for S0 mode.

considering, for example, the propagation of the S0 mode[7].

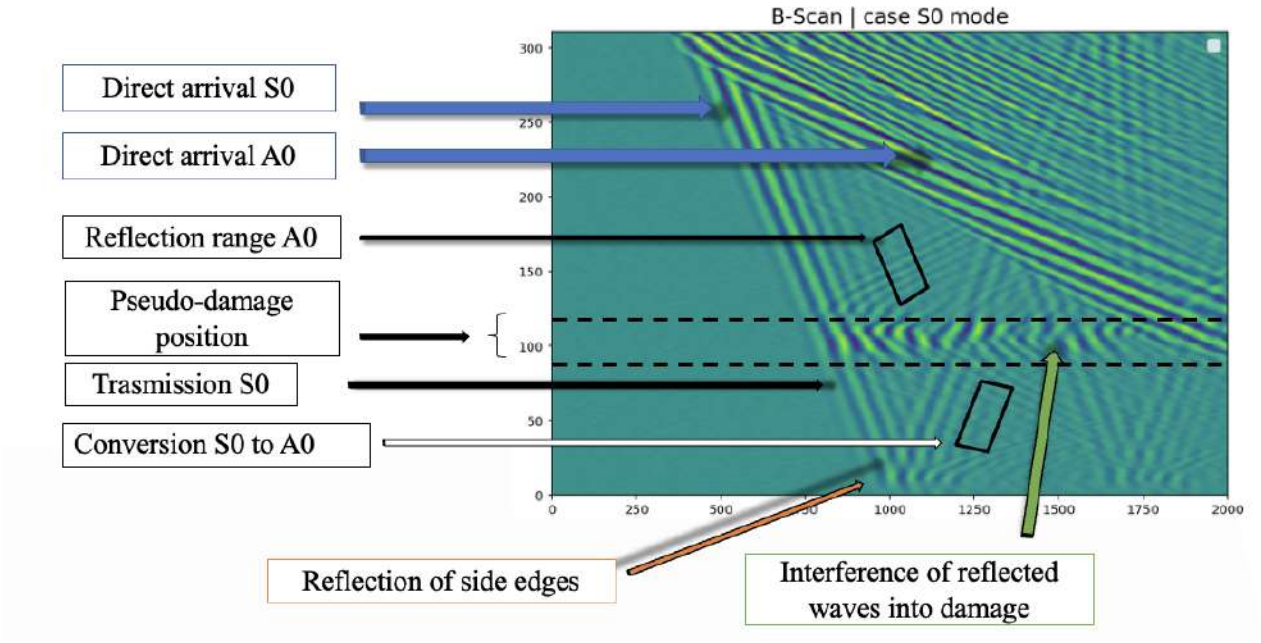


Figure 3.9: B-Scan of S0 mode analyzed.

[7]The considerations presented for the S0 mode can be readily extended to the other fundamental Lamb wave. This discussion focuses solely on the S0 mode for clarity, acknowledging that the A0 mode is also excited.

Fig. 3.9 presents an annotated B-scan: the image captures the key physical phenomena arising from the interaction of the guided waves with a pseudo-damage:

- The early high-velocity wave packet, labeled *Direct arrival S0*, corresponds to the direct propagation of the S0 mode, while the *Direct arrival A0* identifies the arrival of the slower antisymmetric component, which may be generated due to imperfect mode isolation during actuation.
- The region labeled *Reflection range A0* marks the temporal and spatial window where S0 mode is converted [8] and then reflected. The reflections from the damage site can provide information for backscatter-based diagnostics for the characteristics of the known defect.
- The central dashed lines indicate the spatial extent of the pseudo-damage. After this region, the transmitted portion of the S0 wave (*Transmission S0*) is detected downstream of the damage, often exhibiting amplitude attenuation and phase distortion due to scattering effects. Another mode conversion phenomenon is also evident (*Conversion S0 to A0*), where part of the incident S0 wave is transformed into an A0 mode upon interaction with the defect.
- Secondary features, such as the *Reflection of side edges*, are attributed to waveguide boundary reflections and are typically considered parasitic signals in the context of damage detection. Lastly, the region labeled *Interference of reflected waves into damage* denotes a zone where reflected and incident waves interfere constructively or destructively, often forming complex patterns that are sensitive to both the position and nature of the damage.

This figure serves as a reference framework for the interpretation of the wavefield dynamics and, as it will be discussed, supports the feature extraction methodology adopted in this study, including the analysis of reflected amplitudes, mode conversion energy within the regions of interest we are defining.

[8] Another reason to consider the symmetric Lamb mode is that it exhibit an exclusive behavior with the respect to his anti-symmetrical counterpart: when an defect is encountered mode conversion appears. This is particularly noteworthy: although this study found some information about the reasons behind it (cfr. [4], cap. 11 page 276), the full picture of the phenomena is still missing. What is interesting to note is that continuous mode conversion is a characteristic signature of local asymmetries or thickness variations.

3.1.2 Feature definition

So far the experimental data acquired during the experimental campaign and the environment in which the experiments will be conducted were discussed. Moreover, concerning the experimental data acquired, the B-scans, an explanation of the information about the A0 and S0 behavior has been given. At this point a way to characterize numerically the different experiences, referring different pseudo-damaged configuration of the specimens, is needed. In order to do so, a feature definition that extract phisics-based information from the interaction of the wave has been deployed.

The aim is to calculate characteristic values that are comparable between different types of pseudo-damages and, in order to achieve this, the following features [3] will refer to the amplitude of waves packets. More specifically, due to the interference of decaying wave packets and stochastic variation, common areas are defined in the B-scans for the two modes in which the mean value of the maximum amplitude is formed for equalization. The maximum amplitude is determined in the B-scans in the section of the wave packets corresponding to the edge of the respective zone [9] and from that the mean value of the maximum amplitudes is then calculated for each area.
For both A0 and S0, the zones are positioned relative to the leading and trailing edges of the pseudo-damage and, in order to determine the amplitudes of the wave packets for reflections and conversion, areas with low interference oscillations are selected.

For the A0 mode, a set of 4 different features is defined:

$$\underline{x}_{A_0} = \begin{pmatrix} x_{A_0 transition1} \\ x_{A_0 reflection1} \\ x_{A_0 reflection2} \\ x_{A_0 reflection3} \end{pmatrix}^T \tag{3.1}$$

- A0 before Damage: in this area we evaluate the impacting energy of A0 mode. This one will be useful as a normalizing factor for the following features, assuring that the energies exctracted in the different experiences will be comparable.
- A0 transition 1: this is strongly related to how the impacting wave energies differs in terms of transmission through different pseudo-damaged configurations.
- A0 reflection 1, 2 and 3: in these areas the information of how much reflecting energies is emanated from the pseudo-damages. The idea to use 3 different areas is related to the fact that it can be seen that the wave in the area of the damages is trapped, i.e. reflected at the front and rear edges of the defect: this means that the reflected A0 mode is distributed over time and those zones are defined in order to depict this effect.

[9] See Figures 3.10-3.11 for visualization.

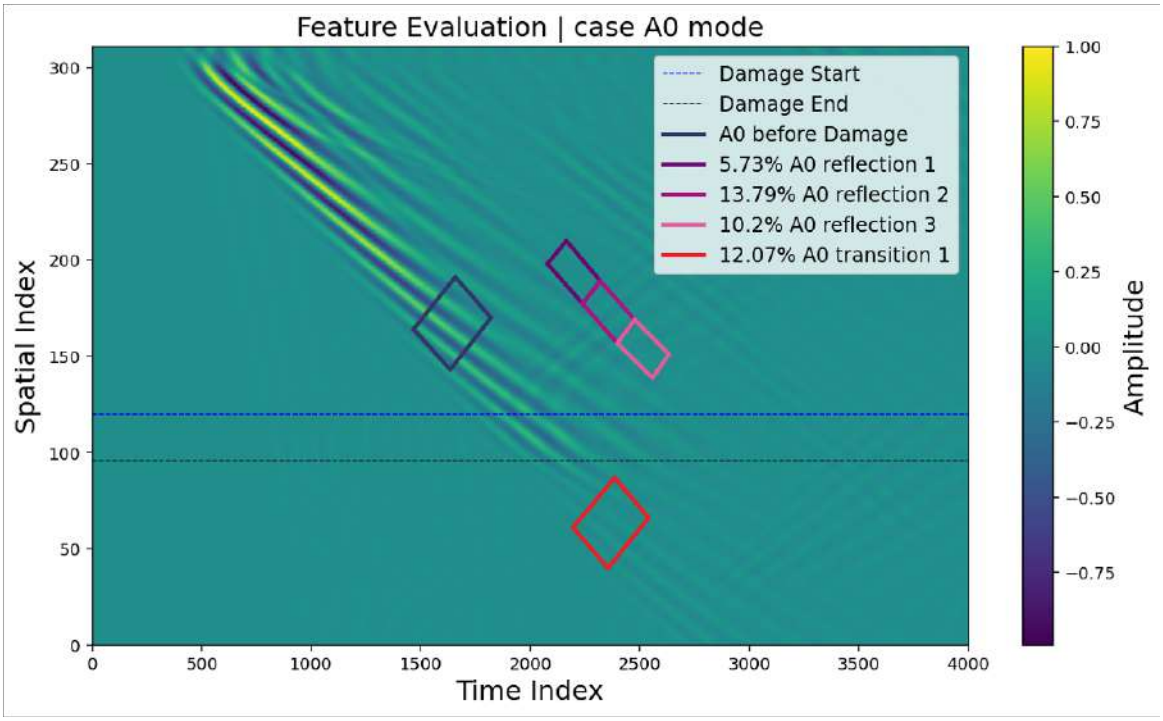


Figure 3.10: common areas defined for A0 feature extraction.

For the S0 mode, a set of 6 different features is prescrived:

$$\underline{x}_{S_0} = \begin{pmatrix} x_{S_0 transition1} \\ x_{A_0 reflection1} \\ x_{A_0 reflection2} \\ x_{A_0 reflection3} \\ x_{A_0 transition1} \\ x_{A_0 transition2} \end{pmatrix}^T \tag{3.2}$$

- S0 before Damage: in this area we evaluate the impacting energy of S0 mode; equivalent of the one for A0, here the energy is evaluated as a scaling factor for comparability.
- S0 transition 1: as made for A0, this area represent the transmitted energy of the impacting Lamb waves and it enable to evaluate how much it differs in terms of transmission through different pseudo-damaged configurations.
- A0 reflection 1, 2 and 3: in these areas the information of how much S0 energy is converted into A0 mode energies and emanated from the leading edge of th epseudo-damages to the aft edge of the specimen. Using again three areas for the same reason as for the reflection in the A0 experience, but now taking into account that is energy obtained from continuous mode conversion.
- A0 transition 1 and 2: in these areas the information of how much S0 energy is converted into A0 mode energies and emanated from the trailing edge to the pseudo-damages to the specimens rear edge.

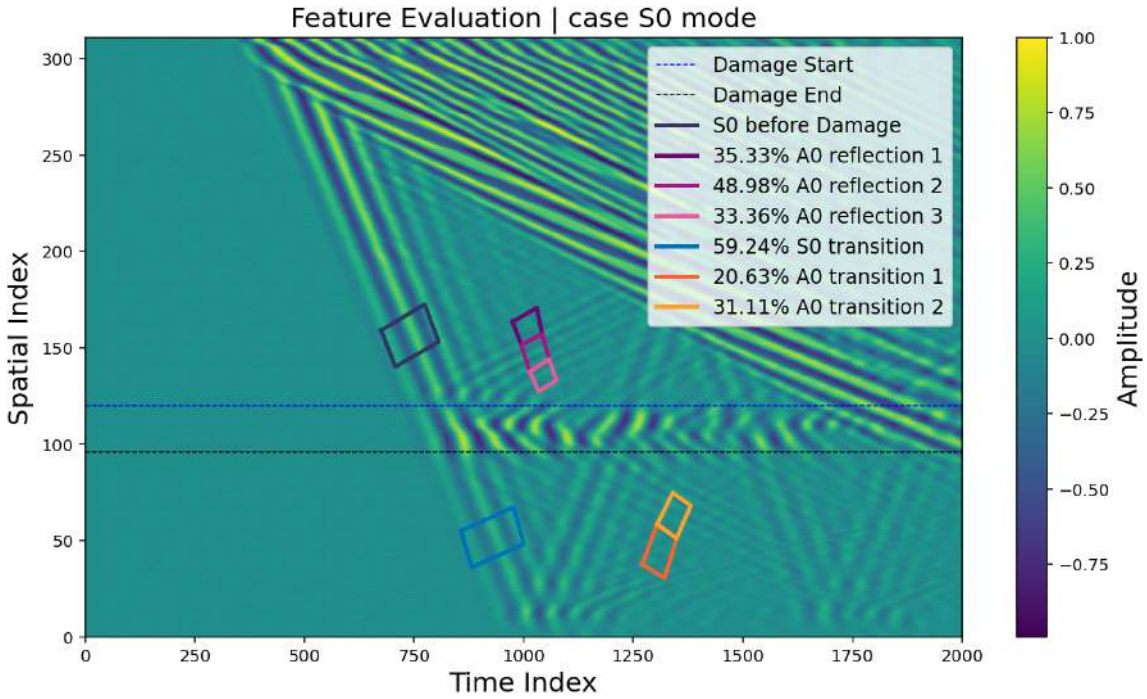


Figure 3.11: common areas defined for S0 feature extraction.

3.2 The Design of Experiment (DoE)

The Design of Experiments (DoE) is a systematic methodology employed to plan, execute, and analyze experiments in order to identify and quantify the influence of multiple input variables. Unlike traditional one-factor-at-a-time approaches, DoE enables the simultaneous variation of multiple factors, thus facilitating the detection of not only the main effects but also potential interactions among variables. A typical DoE framework includes several essential components: the factors, which are the independent variables under investigation; the levels, which define the specific values or categories assigned to each factor; and the responses, which are the output metrics of interest. The combination of factor levels across experimental runs is designed according to predefined schemes, enabling both efficient sampling of the parameter space and statistically robust inference.Several types of experimental designs exist, ranging from full factorial and fractional factorial designs to more advanced configurations such as latyn hypercubes and Box-Behnken designs, often used in the context of response surface methodology. The choice of the particular design clearly depends on the nature of the problem, the number of factors involved, the complexity of their interactions, and the available experimental res ources.

The core of this section is to present how the DoE principles have been applied to explore the behavior of the two fundamental Lamb modes as a function of pseudo-damages configurations. The motivation behind employing this approach is , in one hand, to acquire as much information as possible, testing a vast design space of pseudo-damage configurations. In the other hand, the idea to use this DoE is to evaluate, among the candidates, what is the best pseudo-damage configurations in terms of similarities with BVIDs.

The DoE design has to take into account the main requirements that a good pseudo-damage needs to have, that are:

- To be quasi-reversible, in the sense that it has to be cost-effective and easy to apply and remove onto the specimen in order to simulate the damage along different positions without producing residual effect after the removal.
- To have correlations in terms of wave trasmission, reflections and mode conversion energies with the one of a real impact damage.

Thus in order to prepare a poll of different pseudo-damaged strategies one must take into account how the bound between the PD and the specimen is made, how the mass interact locally with the specimen and how that interaction influences the Lamb modes propagation.

Figure 3.12: different materials and dimensions for the pseudo-damages.

The aim of this study is to investigate the correlation between the B-scan signatures of both A0 and S0 modes and the characteristics of the pseudo-damages applied to the specimen. To this end a DoE based on a full factorial design with three main factors:

- Material composition of the pseudo-damage: four different materials are considered for this factor, namely Polyvinyl Chloride (PVC), Acrylic Glass (RESIN), High-Density Polyethylene (THERMO) and Aluminum (ALL). The material composition will be taken into account by means of the relative density of the material: in this sense we are considering the effect of the composition without correlating it with its dimensions.
- Diameter of the pseudo-damage: this is the only geometric parameter varied in the study [10]. Five different levels are employed, ranging from 20 mm to 40 mm in increments of 5 mm.
- Equivalent shear modulus of the adhesive technique: ten distinct levels are examined, each corresponding to a different adhesive technique. These range from simple solutions (e.g., wax, vacuum tape) to more advanced methods (e.g., ClaroCit combined with polymeric mesh). Among all the considered factors, the adhesive technique has the most significant impact on the quasi-reversibility of the pseudo-damage, making its characterization particularly important in this investigation.

[10]The pseudo-damages used in this framework are cylindrical with a fixed height of 2 cm.

Level	ID Code	Adhesive Type & Description
1	WAX	Thermoplastic mounting wax
2	2SIDED TAPE	Thin double-sided adhesive tape
3	VACUUM	Elastomeric vacuum bagging sealant
4	S50-RT	Low-viscosity cyanoacrylate on masking tape
5	S50-MT	Low-viscosity cyanoacrylate on aluminum tape
6	S77-RT	High-viscosity cyanoacrylate on masking tape
7	S77-MT	High-viscosity cyanoacrylate on aluminum tape
8	2CG-RT	Structural bicomponent epoxy on masking tape
9	2CG-MT	Structural bicomponent epoxy on aluminum tape
10	CC-PM	Cold-curing acrylic resin with polymer mesh

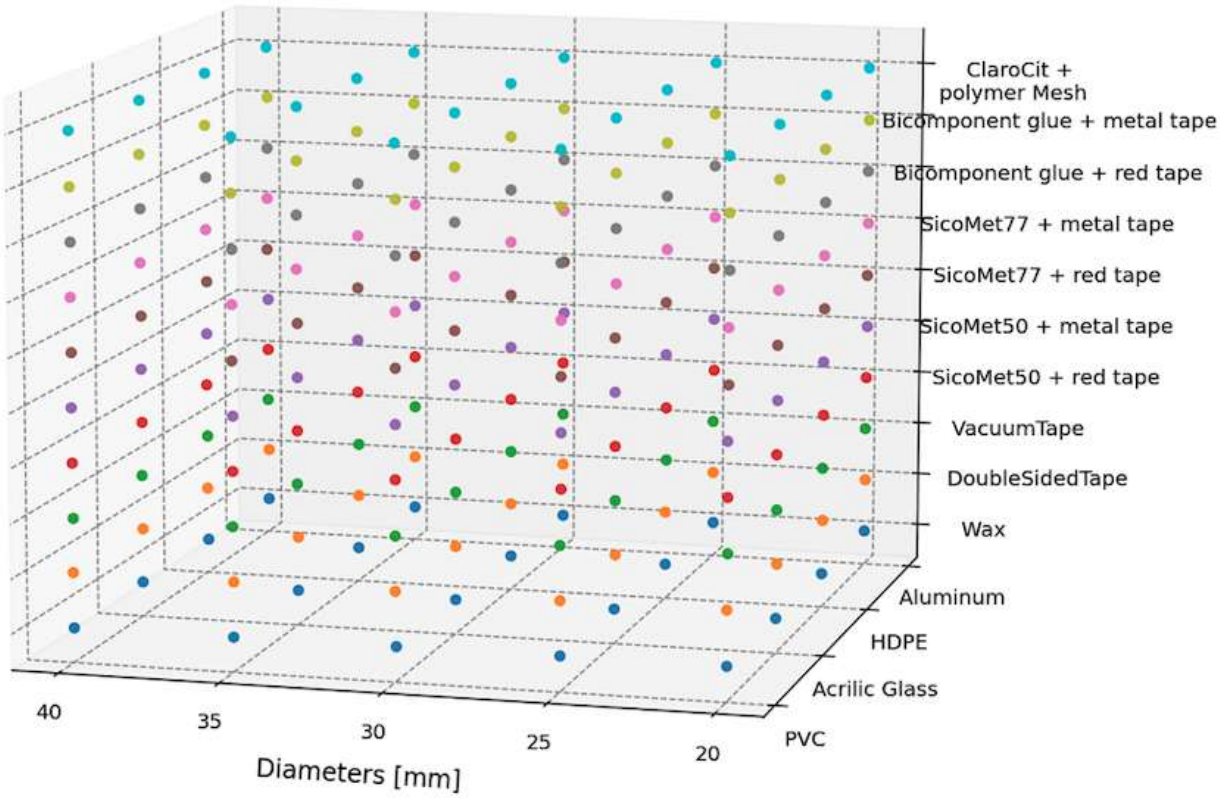


Figure 3.13: Fully Factorial Design of Experiments, visually.

The three-dimensional Design space explored is depicted in Figure 3.14: each one is a configuration that has to be tested by exciting both the A0 and S0 fundamental modes. Clearly, in order to assess a numerical analysis of the interactions, the tested pseudo-damaged configuration has to refer to numerical values that depicts its inertial and elastic properties: for what stands for the densities and diameters,both representative of the inertia, the evaluation is fairly straight by knowing the material used for crafting the pseudo-damage and its dimension. More complex is the G modulus evaluation, for which a set of sub-experiments is conducted.

3.3 Experimental characterization of adhesive shear Properties

To accurately capture the interaction between pseudo-damage and the specimen, the shear properties of the selected adhesive strategies must be characterized. Since the experimental setup only provides the off-plane displacement component of the Lamb modes, the characterization must consider the interaction in a manner that accounts for the preparation of feature labels.

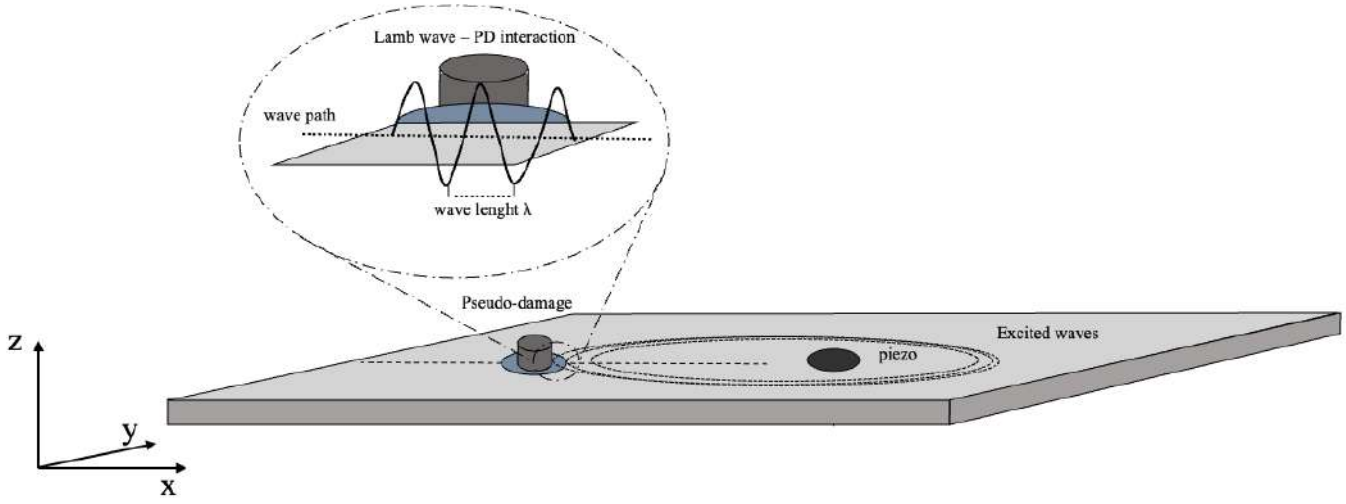


Figure 3.14: Lamb modes and PD interaction, zoomed.

Looking at Figure 3.14, one can note that if the wavelength is bigger than the thickness of the adhesive layer, the main elastic response to the deformation imposed onto the adhesive by the elastic modes of the specimen is a shear deformation[11]. Consequently, the primary elastic response of the adhesive layer is assumed to be shear deformation.

Given the broad range of materials tested, a single experimental characterization method was not feasible. Due to the drastic differences in stiffness and rheological behavior, the adhesives were categorized into two distinct groups:

1. *Structural Adhesives* (e.g., Epoxies, Cyanoacrylates): characterized by high stiffness and linear elastic behavior up to failure.

2. *Soft Adhesives* (e.g., Wax, Tapes): characterized by low shear modulus and significant viscoelasticity, requiring a specialized low-load approach.

[11]For the A0 and S0 modes, the wavelengths are respectively $\lambda_{A0} \approx 15\,\text{mm}$ and $\lambda_{S0} \approx 24\,\text{mm}$.cfr. Cap.2, Dispersion curves. Since $\lambda \gg t_{adhesive}$, the adhesive layer acts primarily as a shear transfer interface between the oscillating structure and the pseudo-damage inertia.

3.3.1 Structural Adhesives: Single Lap Shear Test

The seven structural adhesive configurations (S50-RT, S50-MT, S77-RT, S77-MT, 2CG-RT, 2CG-MT, and CC-PM) were characterized using the Single Lap Shear (SLS) test method, in accordance with the *FOR 3022 DIN EN 1465* standard.

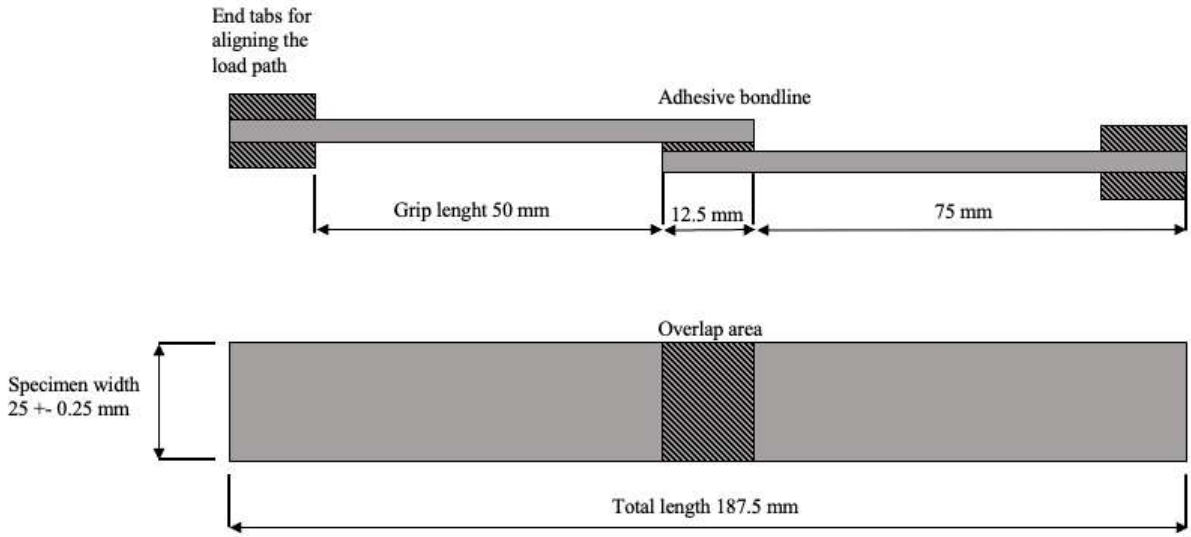


Figure 3.15: Single Lap Shear Specimen dimensions.

Specimens were prepared using Glass fiber (GFK) smooth adherends to minimize substrate deformation. The dimensions of the bonded overlap area were set to a length of $12.5mm$ and a width of $25mm$. For each type of adhesive tecnique several specimen are prepared and the specific preparation process involved first a face of surface preparation: the surfaces were degreased with acetone and lightly abraded to ensure optimal mechanical interlocking. Then the adhesive was applied to the overlap area, and the joint was cured under controlled pressure and temperature conditions specific to each adhesive type.

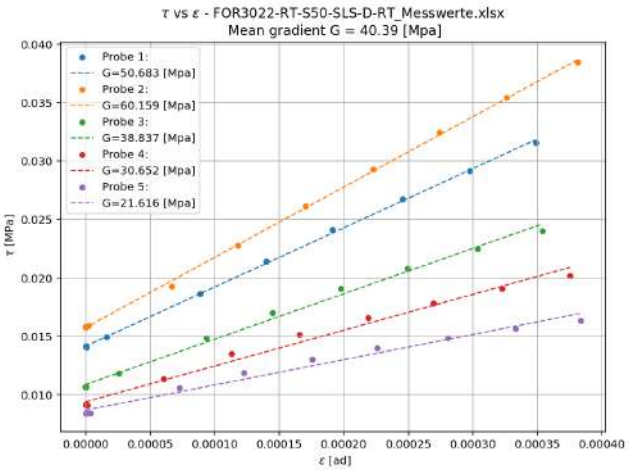
The tests were conducted using a Universal Testing Machine (UTM) equipped with a load cell of 10kN. The crosshead speed was maintained constant at 0.01mm/min until failure occurred.

The shear strength, τ, was calculated as the load F_{max} registered by the load cells increased by dividing by the bond area A:

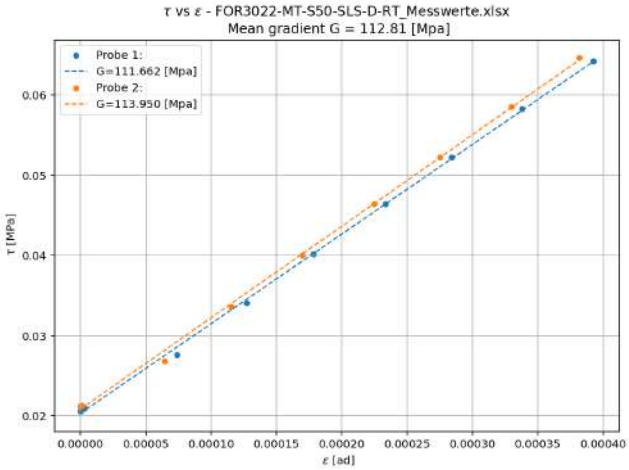
$$\tau = \frac{F_{max}}{l_o \cdot w} \tag{3.3}$$

The shear modulus (G) was subsequently derived from the linear elastic region of the stress-strain curve, obtained for each specimen, for each adhesive.

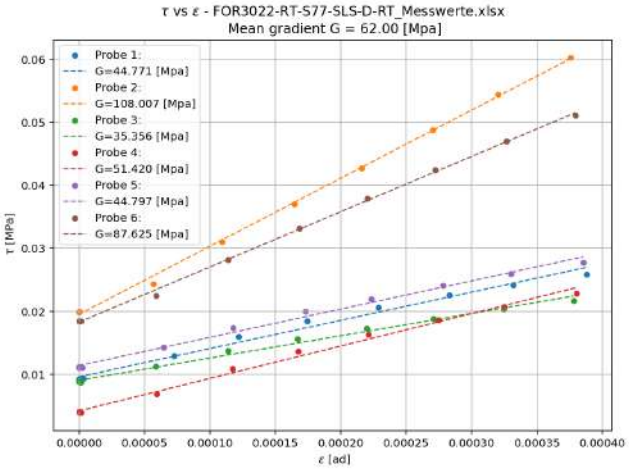
Figure 3.16: Single Lap Shear Test Setup



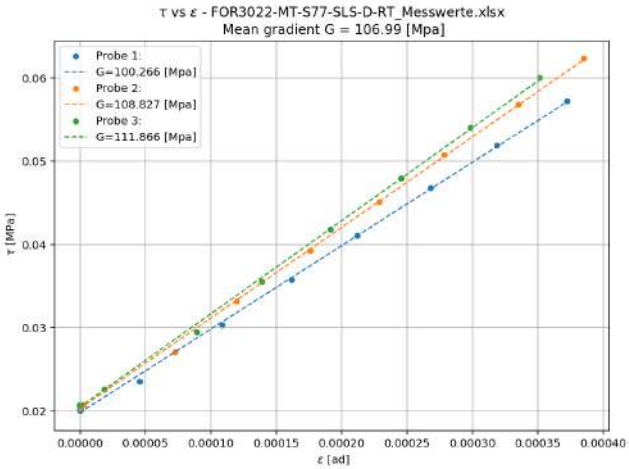
(a) S50-RT



(b) S50-MT

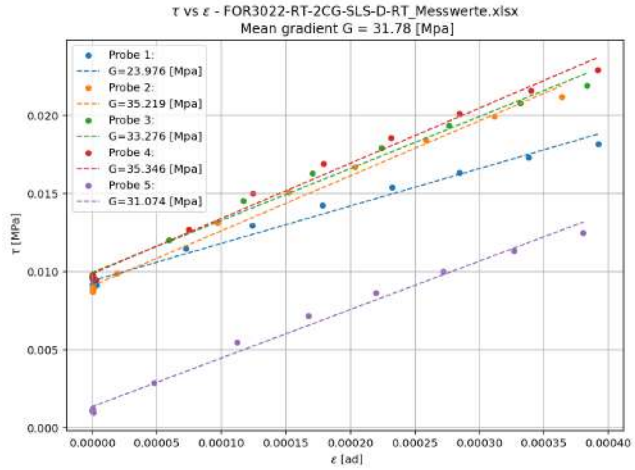


(c) S77-RT

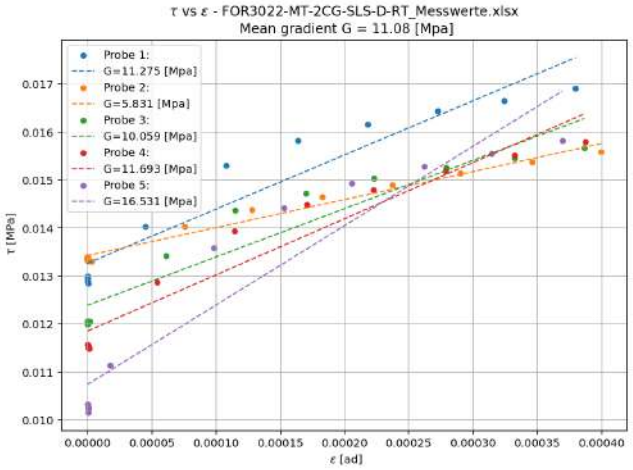


(d) S77-MT

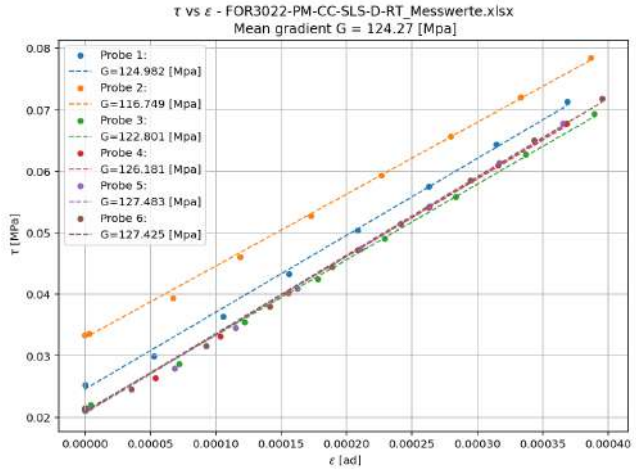
Figure 3.17: Stress-Strain curves for Cyanoacrylate adhesives (SicoMet 50 and 77).



(a) 2CG-RT



(b) 2CG-MT



(c) CC-PM

Figure 3.18: Stress-Strain curves for Bicomponent Epoxy (2CG) and Acrylic (CC) adhesives.

3.3.2 Soft Adhesives: equivalent stiffness characterization

The three compliant anchoring strategies, namely the WAX, 2SIDED TAPE, and VACUUM, exhibit viscoelastic behavior and low shear moduli that make the standard DIN EN 1465 test unsuitable. For these materials, significant creep and substrate misalignment occur before shear failure, invalidating the standard procedures for evaluating the shear modulus.

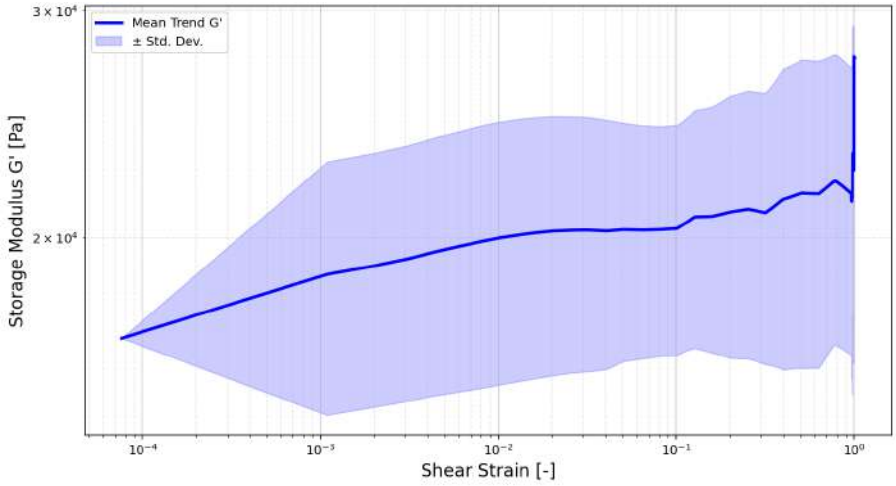
When an adhesive exhibits viscoelastic behavior, it displays both elastic (solid-like) and viscous (liquid-like) properties simultaneously. Its mechanical response can be conceptualized as a combination of a spring and a dashpot: it stores energy in the form of elastic strain (characterized by the Storage Modulus, G') while also dissipating energy as heat due to internal friction (characterized by the Loss Modulus, G''). This particular behavior leads to a lack of a clear "breaking" point; under a static load, these materials undergo continuous deformation, or creep.

Because of this load-dependent flow, the static DIN EN 1465 test is inapplicable. Instead, oscillatory rheometry is used to decouple the material's response. At very low strain amplitudes, the adhesive structure remains intact, defining the Linear Viscoelastic Region (LVR) where the moduli are constant. To establish a parameter comparable to the static shear modulus of the structural adhesives, the equivalent shear modulus G_{eq} for the soft adhesives is defined as the average value of the Storage Modulus G' strictly within this linear plateau:

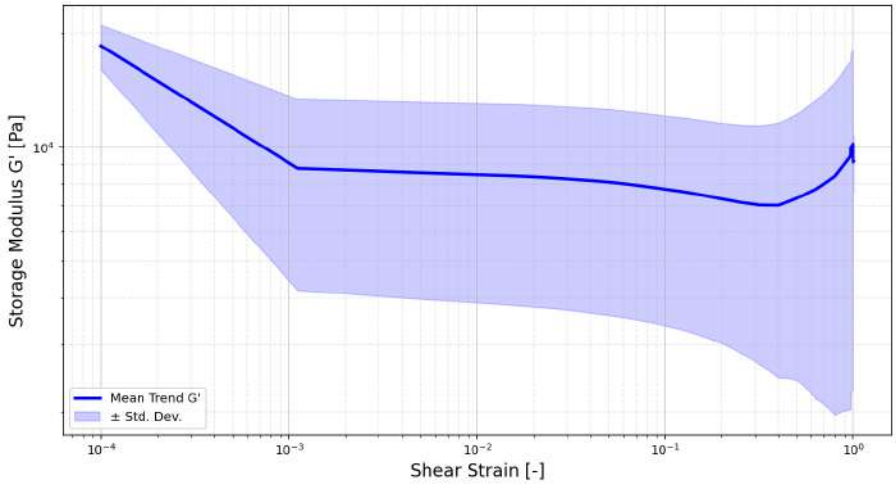
$$G_{eq} = E(G'_{LVR}) \tag{3.4}$$

By postprocessing the rheological analysis for each one of the samples[12], one obtains the trends of the equivalent shear modulus G_{eq} for the three viscoelastic adhesives.

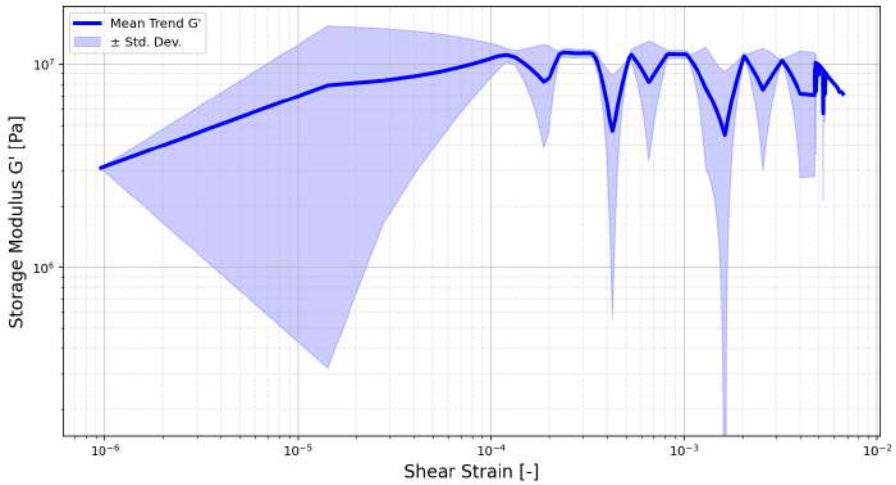
[12]For each one of the three adhesives several samples are produces, thus obtaining a distribution of G' with respect both of the sampled values and with the elastic displacement: the equivalent shear modulus is the average upon those variations.



(a) 2SIDED TAPE



(b) VACUUM



(c) WAX

Figure 3.19: Storage modulus (G') statistical trends, obtained via rheometry.

Doing the aftermath of the experimental shear modulus evaluation, a feasible way to numerically quantify the effect of the adhesive in the interaction between the pseudo-damage has been obtained:.

Level	**ID Code**	**Adhesive Type & Description**	G_{eq} **[MPa]**
1	WAX	Thermoplastic mounting wax	8.46
2	2SIDED TAPE	Thin double-sided adhesive tape	$7.89 \cdot 10^{-3}$
3	VACUUM	Elastomeric vacuum bagging sealant	$2.14 \cdot 10^{-2}$
4	S50-RT	Low-viscosity cyanoacrylate on masking tape	55,42
5	S50-MT	Low-viscosity cyanoacrylate on aluminum tape	112.81
6	S77-RT	High-viscosity cyanoacrylate on masking tape	64.11
7	S77-MT	High-viscosity cyanoacrylate on aluminum tape	106.98
8	2CG-RT	Structural bicomponent epoxy on masking tape	31.78
9	2CG-MT	Structural bicomponent epoxy on aluminum tape	11.08
10	CC-PM	Cold-curing acrylic resin with polymer mesh	124.27

Table 3.1: Summary of the investigated anchoring strategies and their equivalent shear modulus evaluated experimentally.

3.4 Reproducibility and B-scan alignment

This section addresses the reproducibility of the experimental campaign, specifically focusing on mitigating the inherent variations among the ten CFRP plates. To reliably utilize the entire plate batch for pseudo-damage testing, it is imperative to ensure that the extracted features are not biased by the natural baseline discrepancies of each individual specimen. Minor variations in manufacturing tolerances, material properties, and sensor installation can induce slight phase and time-of-flight shifts in the propagating Lamb waves and thus altering the feature extracted from the B-scans.

To systematically account for these baseline differences, a temporal alignment procedure was implemented prior to any DoE evaluation. Plate 1 was selected as the universal reference pristine baseline. The calibration process is structured as follows:

1. *ROI Definition:* A specific Region of Interest (ROI) is defined within the pristine B-scans to isolate Lamb wave modes.

2. *Cross-Correlation:* To evaluate the temporal misalignment of each plate with respect to the reference, the mean signal profile within the ROI is extracted. A discrete cross-correlation is then computed between the reference profile ($x_{ref}[n]$) and the target plate profile ($y_{target}[n]$):
$$R_{xy}[k] = \sum_{n=-\infty}^{\infty} x_{ref}[n] \cdot y_{target}[n+k] \tag{3.5}$$

3. *Time Shifting:* The optimal time shift Δk required to align the signals is identified by finding the lag k that maximizes the cross-correlation function. The entire B-scan of the target plate is subsequently shifted by this calculated delay.

Applying this temporal alignment is a critical prerequisite for the subsequent B-scan post-processing. By neutralizing the intrinsic pristine variability, we ensure that any variations detected later are exclusively attributable to the pseudo-damage. Without this rigorous alignment protocol, the dynamic features evaluated during the feature definition phase (such as amplitude drops or phase shifts) could not be tracked consistently or reliably across different plates.

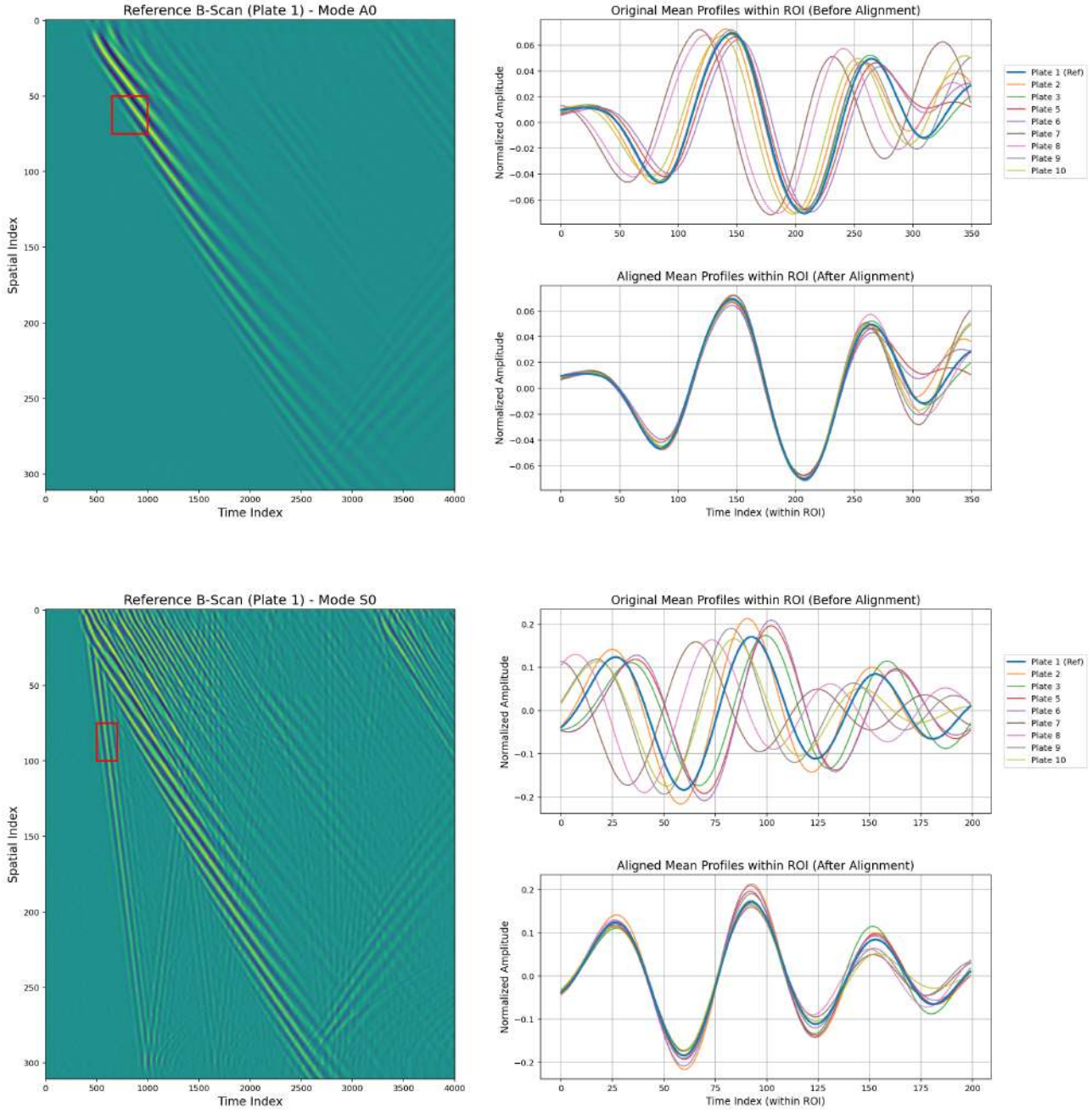


Figure 3.20: Left: Reference pristine B-scan with the selected Region of Interest (ROI) highlighted. Right: Original and aligned mean profiles within the ROI.

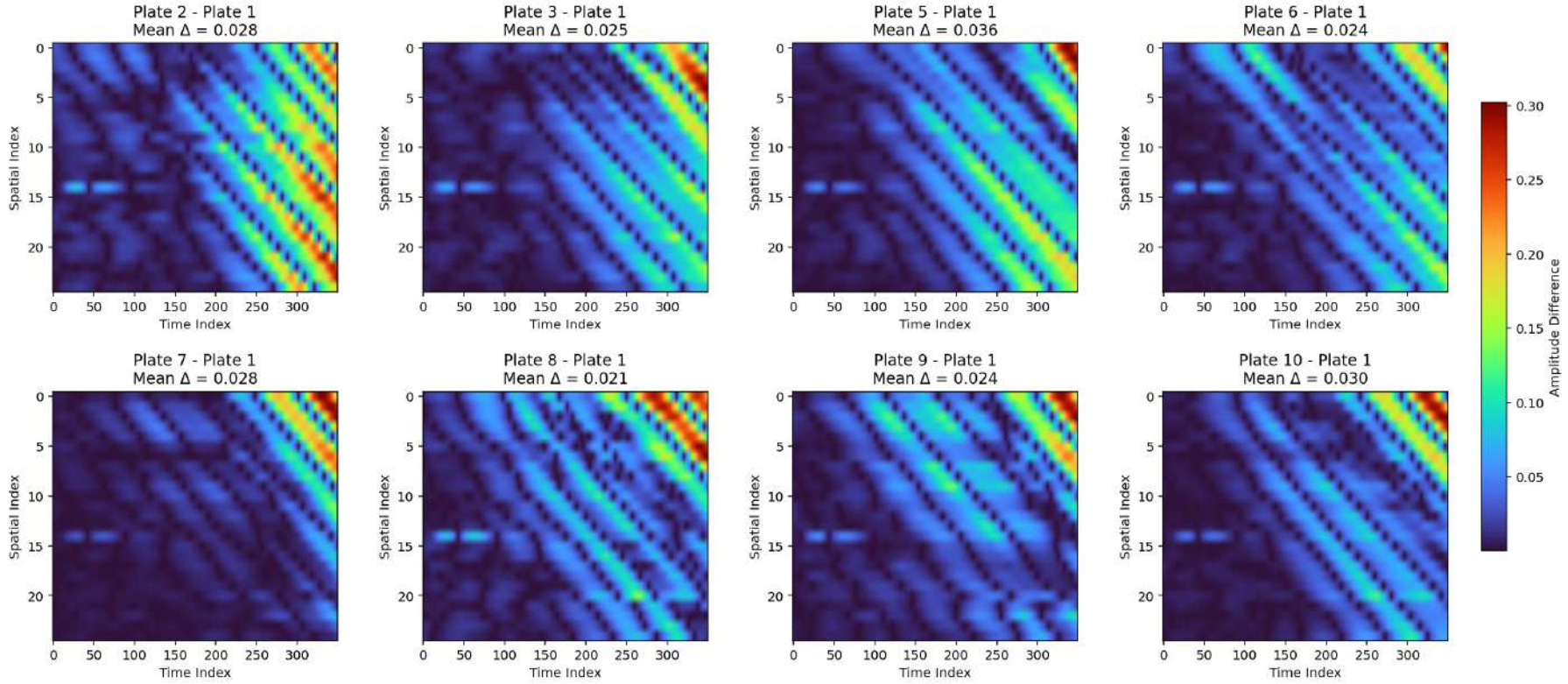


Figure 3.21: Heatmaps representing the absolute amplitude differences between the aligned pristine plates and the reference baseline (Plate 1) within the A0 mode ROI.

3.5 Feature extraction

To systematically evaluate the influence of different pseudo-damage configurations on the propagation of Lamb waves, a feature extraction pipeline was needed: the description of the algorithm tailored to solve this task is the topic of this section.
This algorithm processes the raw ultrasonic B-scans produced in the experimental campaign to isolate, quantify, and compile the specific wave interactions associated with the damage zone. The procedure operates through the following logical phases:

1. Signal Preprocessing and Alignment: Initially, the raw binary ultrasonic data is loaded, reshaped into a spatial-temporal matrix, labeled to its corresponding DoE set and normalized against its maximum absolute amplitude. To ensure consistency across the entire dataset, each B-scan is temporally aligned to the reference baseline, as evaluated by the cross-correlation analysis briefly described in Section 3.4.

2. Dynamic Damage Localization and Feature Tracking: Because the pseudo-damage diameter varies across the test pool of configurations prescribed via DoE, the algorithm dynamically calculates the spatial boundaries of the damage for each specific scan. By extracting the diameter directly from the dataset metadata, the spatial indices corresponding to the start and end of the pseudo-damage are automatically defined.

   Subsequently, discrete spatio-temporal windows are mapped onto the B-scans to isolate the specific wave packets previously defined in Section 3.1.2. Since the temporal arrival of these boundary interactions depends strictly on the physical size of the anomaly, the temporal and spatial coordinates of these bounding boxes are dynamically shifted as a function of the extracted damage diameter.

3. Quantification and Data Compilation: Within each dynamically positioned bounding box, the algorithm extracts the maximum absolute amplitude. To account for slight variations in the initial excitation energy across different experimental runs, all extracted features are normalized against the incident baseline wave ($A0_{\text{before}}$ and $S0_{\text{before}}$) of that specific scan. This yields a relative percentage for each feature:

$$x^{PD}_{LambMode_i} = \left(\frac{A_{\text{ith-box}}}{A_{\text{incident}}} \right) \times 100 \tag{3.6}$$

   where $A_{\text{ith-box}}$ is the maximum amplitude within the specific feature window and A_{incident} is the maximum amplitude of the reference incident wave. Finally, the algorithm generates an annotated B-scan image for visual validation and compiles the quantified features vector $\underline{x^{PD}_{A0}}$ and $\underline{x^{PD}_{S0}}$, along with their corresponding material, adhesive stiffness and diameter metadata, into a structured dataset.

This automated pipeline is used for analyzing the behavior of the defined feature, which are representative of the reflected, transmitted and eventually converted energies of A0 and S0 Lamb modes, across the 200 different configuration: a total of 400 different configurations[13] is so analyzed.

[13]For which a total of 1o different features are collected.

Chapter 4

Characterization of Pseudo-damaged configurations

4.1 Feature analysis & modeling

The primary objective of this section is to establish mathematical relationships between the physical characteristics of the Pseudo-Damage configurations and the resulting wave scattering and attenuation phenomena, in order to understand the relationships between those patterns and eventually exploit them for understanding how to simulate a real damage. Clearly, the modeling must be:

- Simple enough: the features are spanning into a relatively dense DoE which is essentially referring upon 3 variables, so the law underlying the features with respect to them must be able to catch the fundamental relations with the three of them.
- Representative: The relationships obtained must accurately reflect the actual behavior of the features, ensuring that any extrapolated conclusions are not misleading.

The experimental space generated by the DOE is governed by three independent physical variables, the diameter d, stiffness G and density ρ across which the 3D Fully-Factorial DoE is spanning. The dependent variables are the sets of normalized features extracted in the previous chapter:

$$\underline{x}_{S_0} = \begin{pmatrix} x_{S_0 transition1} \\ x_{A_0 reflection1} \\ x_{A_0 reflection2} \\ x_{A_0 reflection3} \\ x_{A_0 transition1} \\ x_{A_0 transition2} \end{pmatrix}^T \quad \text{and} \quad \underline{x}_{A_0} = \begin{pmatrix} x_{A_0 transition1} \\ x_{A_0 reflection1} \\ x_{A_0 reflection2} \\ x_{A_0 reflection3} \end{pmatrix}^T$$

To quantify the relationship between the DOE variables and the responses, a progressive regression analysis was implemented by evaluating the Least Square Method regression curves.

Generally speaking, if one has n samples $\hat{\underline{y}}$ of the generic function $\underline{y}$ which is depending upon p variables, a viable way to obtain a linear model of $\underline{y}$ is to set it as:

$$\underline{y} = \underline{\underline{X}}\underline{\beta} + \underline{\epsilon} \tag{4.1}$$

Where $\underline{\underline{X}}$ is the matrix that contains the independent variables, $\underline{\beta}$ is the unknown vector of coefficients, and $\underline{\epsilon}$ is the vector of the residuals. A feasible way to obtain the model in Equation (4.1) is to use the n samples of $\hat{\underline{y}}$ and to evaluate, again from (4.1), the residual vector $\underline{\epsilon}$. These residuals can be used to define a scalar cost function $J(\underline{\beta}) = \underline{\epsilon}^T\underline{\epsilon}$[1] where $\epsilon = \underline{y} - \hat{\underline{y}}$ from the variable sampled values . The $\hat{\beta}$ that minimize the cost function is the one that reflect that:

$$\hat{\underline{\beta}} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\hat{\underline{y}} \tag{4.2}$$

By changing the form of $\underline{\underline{X}}$, different regression models can be obtained. For instance, assuming a first-order relationship between the defined energy features and two independent variables of the DoE —such as the pseudo-damage diameter d and the material density ρ—the feature is modeled as a 2D plane. In this scenario, for n experimental samples, the design matrix $\underline{\underline{X}}_{lin}$ is constructed with a column of ones to account for the intercept β_0:

$$\underline{\underline{X}}_{lin} = \begin{bmatrix} 1 & d_1 & \rho_1 \\ 1 & d_2 & \rho_2 \\ \vdots & \vdots & \vdots \\ 1 & d_n & \rho_n \end{bmatrix} \qquad \Longrightarrow \qquad \underline{\beta}_{lin} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \tag{4.3}$$

Conversely, if the physical interaction demands a second-order polynomial regression to capture curvature and interaction effects between the variables, the geometry of the fitted surface becomes a paraboloid or a more complex curved surface. To achieve this without altering the linear solving algorithm in Equation (4.2), the design matrix is artificially augmented: new columns are computed directly from the original inputs to represent the squared and cross-product terms. In the scenario of using :

$$\underline{\underline{X}}_{poly} = \begin{bmatrix} 1 & d_1 & \rho_1 & d_1^2 & \rho_1^2 & (d_1 \cdot \rho_1) \\ 1 & d_2 & \rho_2 & d_2^2 & \rho_2^2 & (d_2 \cdot \rho_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & d_n & \rho_n & D_n^2 & \rho_n^2 & (D_n \cdot \rho_n) \end{bmatrix} \qquad \Longrightarrow \qquad \underline{\beta}_{poly} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} \tag{4.4}$$

Because the non-linear terms (D^2, ρ^2, $D \cdot \rho$) are treated simply as additional independent columns within $\underline{\underline{X}}_{poly}$, the system remains mathematically linear with respect to the unknown coefficients vector $\underline{\beta}$ and so, the procedure just explained is still usable.

To quantitatively assess the validity of the regression models across the different adhesive strategies, the Coefficient of Determination R^2, was utilized. The R^2 score represents the proportion of the information in the dependent variable that is predictable from the independent variables. An R^2 value approaching 1.0 indicates that the feature is highly sensitive and strictly

[1] This basically represents the squared distance of the actual sampled values, from which the method takes the name

governed by the variables upon which the PD was crafted, implying a stable and reliable representation from the model. Conversely, a low or negative R^2 suggests high data scattering, indicating that the feature is either bad captured by the model or heavily influenced by other, unknown variables.
Consequently, analyzing the R^2 scores across the different adhesive allows for an objective ranking of both the most interesting acoustic features and their relations with the DoE independent values.

In order of simplicity, the initial analysis conducted focused on the analysis of features across the five different diameters and four densities, while disregarding the effects of stiffness. For each adhesive used, a linear regression model was employed to elucidate the dependencies of the features with respect to diameters, across the various materials.

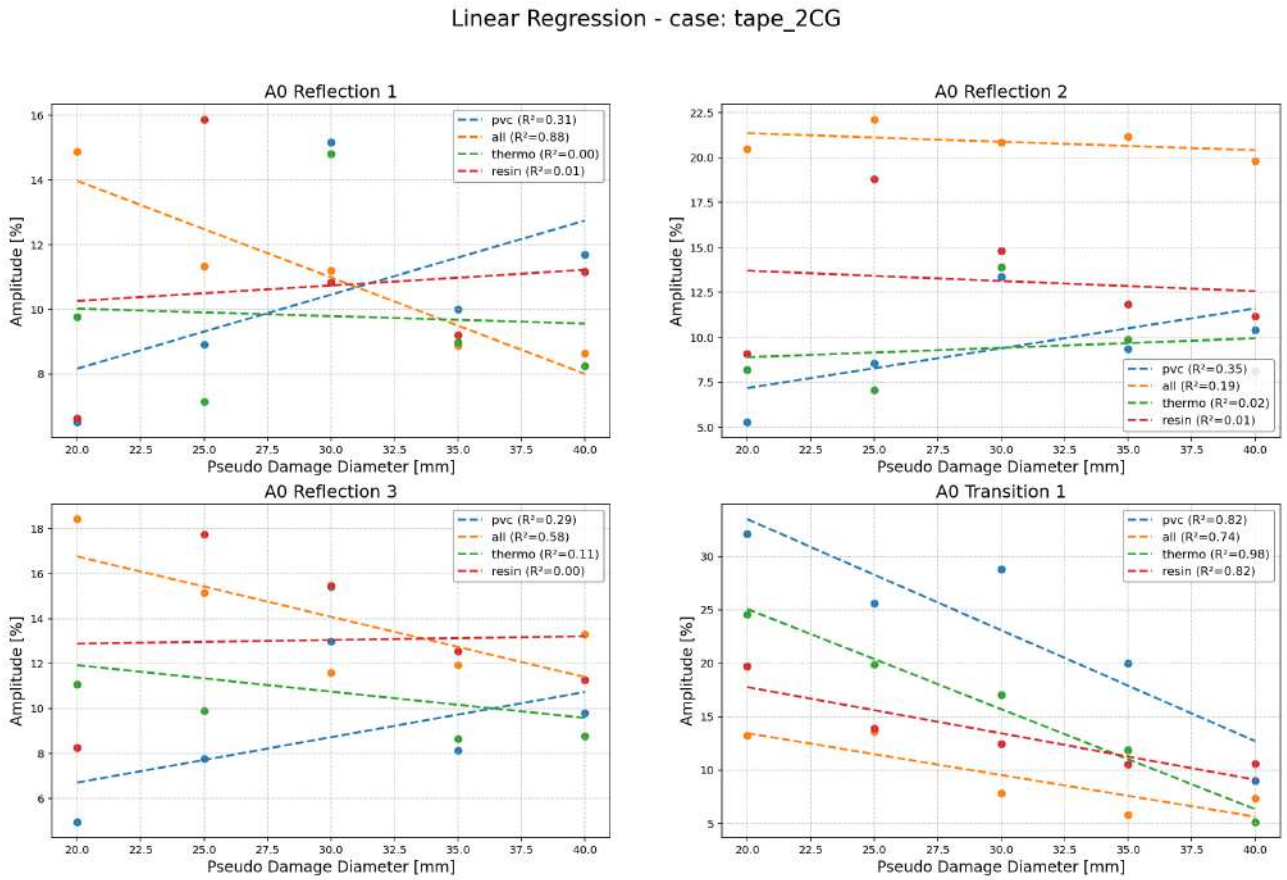


Figure 4.1: 2D regression models for A0 features, differences across different densities.

For each type of material used for the PD configuration, one can note how the diameter affect the reflection and transmission energies: the bigger the diameter, the lower the transmission energy with the slope, and the R^2, that are also dependent upon the different densities and different adhesive used. The same reasoning can be applied also for the feature defined for the S0 Lamb mode: the bigger the diameter, the bigger the features referring to the reflected energy.

Clearly those models provide simple linear relationship that enable a qualitative analysis of both the independent variables of the DoE and the phisical-based features: as long as their

mutual relationship is strong, the R^2 will be big and the bigger the R^2 evaluated for the linear regressor, the closer their relationship is to a linear one. But what is more important is that in those 2D linear models all the features values, spreading across DOE, are not taken into account, thus excluding from the analysis any consideration about the effect of different adhesives[2]. Also those model, as already noted, are not able to model the cross-influences of different diameters and densities in the features values.

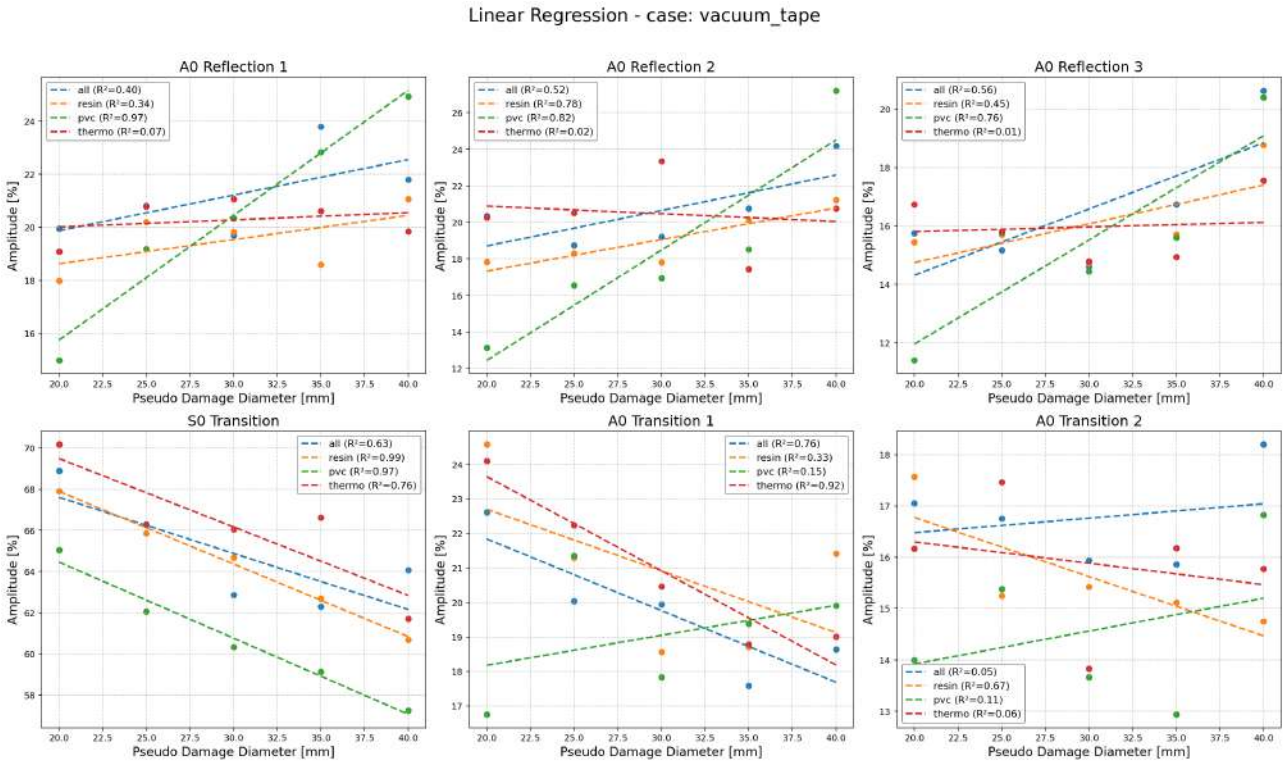
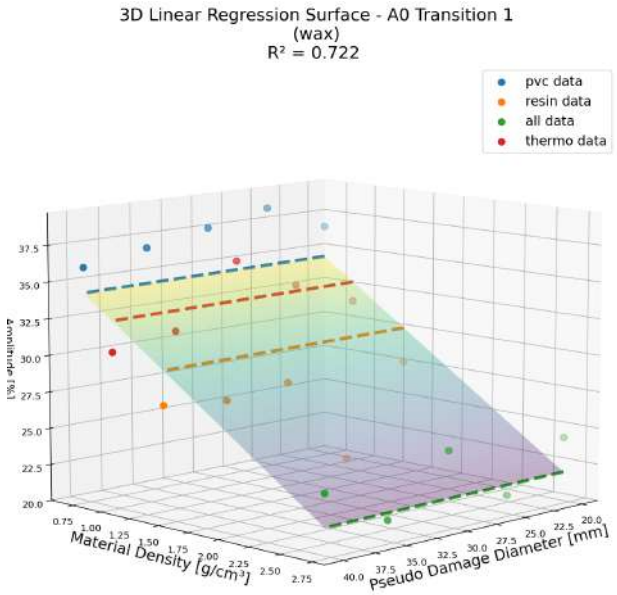


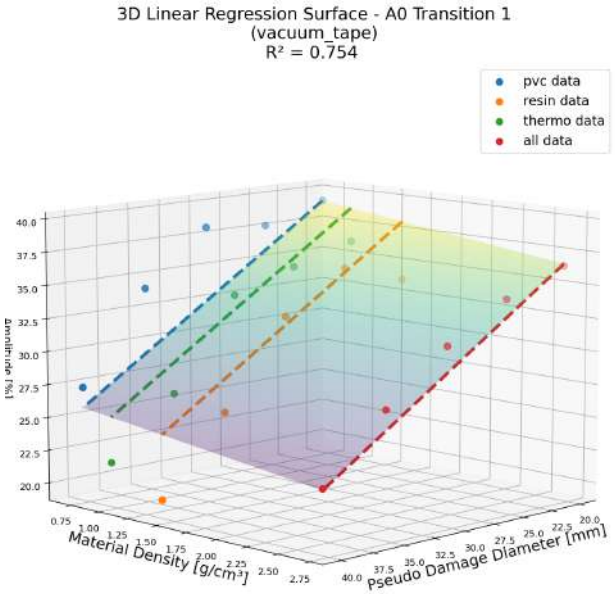
Figure 4.2: 2D regression models for S0 features, differences across different densities.

To gain insights into the influences of both diameter and density, linear regression response surfaces are generated for each feature defined, for each adhesive utilized. The regression type remains linear, and the strength of the dependency is quantified by the R^2 metric. Notably, for each regression plane obtained for a defined feature, a distinct trend emerges, reflecting the mediation of interaction by the adhesive used. For instance, in the A0 case, analyzing the transmitted energy for two soft adhesives, such as wax and vacuum tape, reveals that for the wax adhesive, the primary influence is attributed to the effect of density change. Conversely, for the vacuum tape adhesive, the influence is strongly related to the diameter of the PD, with a diminished influence across diameter. (Of course, this is related to the linear response surface, but it is still insightful to understand how these differences can be observed while merely altering the adhesive.) The distinct topological trends between the two figures unequivocally demonstrate that the adhesive interface dictates the scattering behavior, while the numerical R^2 metric ensures that the response surface accurately represents the overall trends.

[2]As the plots title suggests, for each feature at fixed density only 5 samples are taken into account, that are those referring to the same adhesives used.

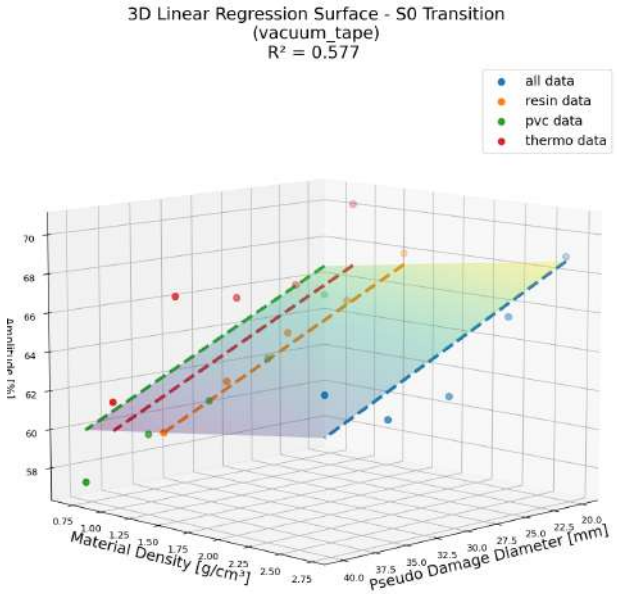


(a)

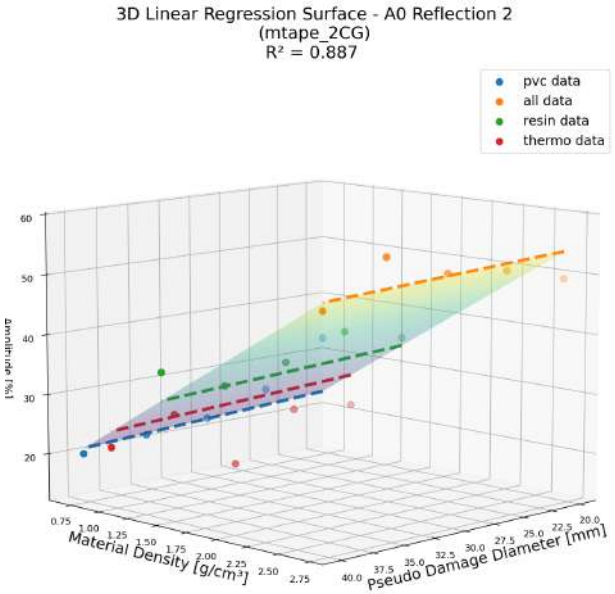


(b)

Figure 4.3: Comparison of the 3D quadratic regression surfaces for the same feature plotted against PD diameter d and material density ρ, for different adhesives.



(a)



(b)

Figure 4.4: Comparison of the 3D plane regression surfaces for the same feature plotted against PD diameter d and material density ρ, for different adhesives.

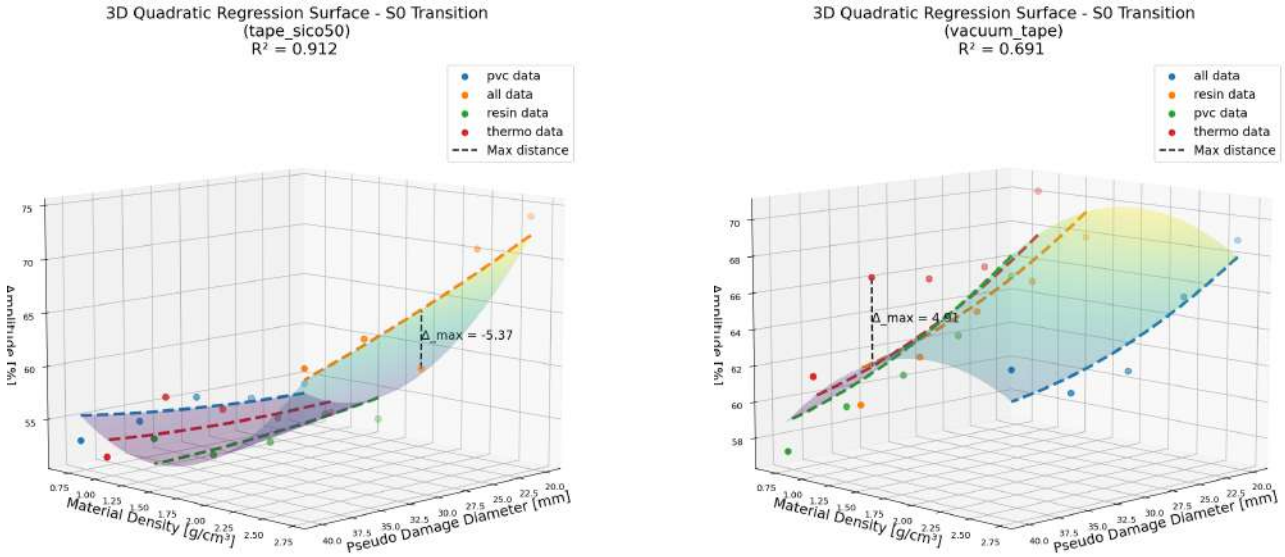
The initial linear models provided a baseline understanding of feature sensitivity, but they inherently assume a constant rate of change over the DoE. However, the physical interaction of ultrasonic Guided Waves with structural heterogeneities is a highly complex, non-linear phenomenon.

Specifically, the wave scattering cross-section is geometrically related to the area of the Pseudo-Damage rather than its mere 1D diameter, suggesting a proportional dependence on d^2. Furthermore, phenomena such as mode conversion, local resonances, and trapped wave modes do not scale linearly with the inertial mass, represented by the density ρ.

To mathematically capture these physical non-linearities, a second-order polynomial regression was implemented. The feature response x_i^{PD} is no longer constrained to a flat plane, but is instead modeled as a 3D curved surface defined by:

$$x_i^{PD}(d, \rho) = \beta_0 + \beta_1 d + \beta_2 \rho + \beta_3 d^2 + \beta_4 \rho^2 + \beta_5 (d \cdot \rho) \tag{4.5}$$

This formulation introduces pure quadratic terms (d^2 and ρ^2) to model the spatial and inertial curvatures. Crucially, it also introduces the cross-product interaction term that allows the model to capture synergistic effects, acknowledging that the scattering caused by a Pseudo-Damage that is simultaneously large and dense may differ significantly from the linear superposition of its individual geometric and material properties. This non-linear behavior, unseen from the linear response surfaces, is now modeled and so the R^2 is typically higher than the linear ones[3].



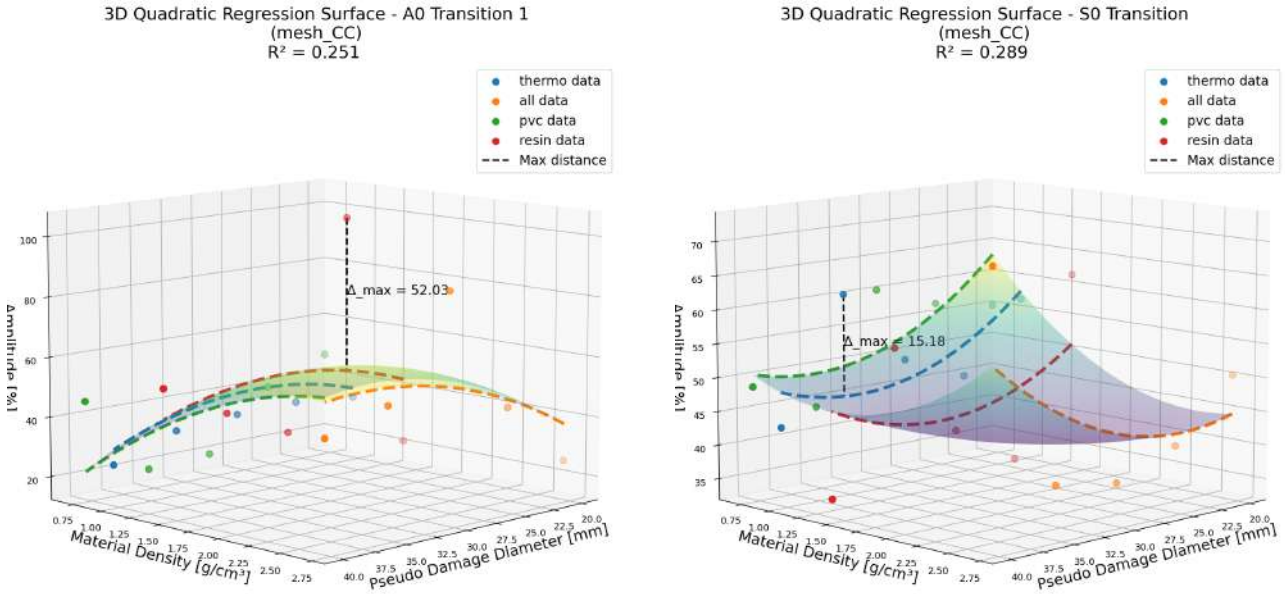
(a) (b)

Figure 4.5: Comparison of the 3D quadratic regression surfaces for the same feature plotted against PD diameter d and material density ρ, showing better correlation than linear models.

However, it is important to emphasize that an increase in model fidelity is not universally guaranteed for every feature or adhesive configuration. As clearly illustrated in Figure 4.6,

[3]As a direct example, the reader can compare Figure 4.4a and Figure 4.4b

certain features exhibit significant data dispersion and poor correlation. Rather than indicating an experimental flaw, this weak fit reveals a physical insight: for these specific Lamb mode- PD interactions feature, the scattering is not only primarily governed by the geometric footprint d or the inertial mass ρ of the Pseudo-Damage. Instead, the high unmodeled variance strongly suggests that the system is dominated by the third, until now hidden, variable of the DoE: the shear stiffness of the adhesive layer. This confirms that, in certain regimes, the bonding interface dictates the energy transfer mechanism, rendering any bivariate model that omits the adhesive stiffness inherently incomplete.



(a) (b)

Figure 4.6: Comparison of the 3D quadratic regression surfaces for the same feature plotted against PD diameter d and material density ρ, showing low correlation.

For a fixed feature, the mean value of the R^2 and its variability across different adhesive techniques can serve as an indicator of the feature's sensitivity to diameter and density, as well as its invariance to the adhesive choice.[4] This line of reasoning can thus be utilized as a foundation for a preliminary feature sensitivity analysis regarding the parameters of the DoE.

The global statistical benchmark, summarized in the ranking of mean R^2 and standard deviations, provides profound insights into the underlying wave-damage interaction mechanisms with A0 and S0 Lamb modes.

For the A0 mode, The primary transmitted wave energy (*A0_transition_1*) emerges as the most robust and sensitive diagnostic feature, exhibiting an exceptional mean R^2 of 0.820 with a minimal standard deviation across all bonding strategies. From a physical perspective, forward scattering is predominantly governed by the macroscopic "shadowing" effect of structural

[4]It is important to acknowledge that this statistical interpretation relies on the assumption of high experimental repeatability. Despite strict laboratory protocols during the Pseudo-Damage installation, manual bonding inherently introduces a degree of random variability. Consequently, the unmodeled variance reflected in the R^2 metric is not exclusively driven by the adhesive's physical shear mechanics, but is inevitably convoluted with experimental noise.

Table 4.1: Global mean R^2 scores and standard deviations for the extracted A_0 features across all evaluated bonding strategies.

A0 Feature	**Mean R^2**	**Standard Deviation (σ)**
A_0 Transition 1	0.820	0.109
A_0 Reflection 2	0.559	0.222
A_0 Reflection 3	0.437	0.136
A_0 Reflection 1	0.272	0.123

heterogeneity. The transmitted energy is strongly attenuated by the spatial footprint d and the inertial mass ρ of the PD, while remaining largely insensitive to the specific shear transfer properties of the intermediate adhesive layer.

Conversely, the backscattered energies (*A0_reflection_1, 2, 3*) exhibit significantly lower mean correlations (ranging from 0.272 to 0.559) and high relative variance. This statistical behavior substantiates that wave reflection is highly sensible to the adhesive type and, so, critically influenced by the interaction of the anti-symmetric Lamb mode and the adhesive at the leading edge of the defect. Modifications in the adhesive stiffness and composition significantly alter the local acoustic impedance mismatch and the shear coupling. Consequently, the reflected energy is heavily contaminated by the bonding strategy.

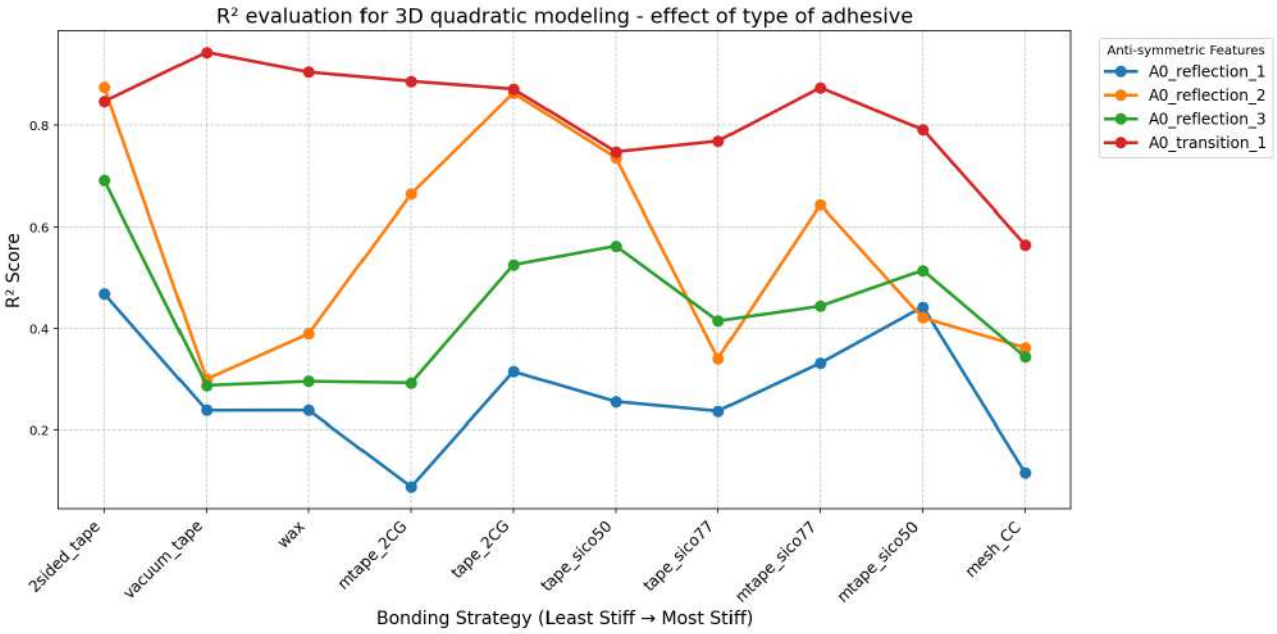


Figure 4.7

The global statistical analysis of the S_0 incident wave regime reveals a different scattering behavior compared to the antisymmetric case. The most striking observation is the significant overall reduction in the R^2 scores accompanied by substantially wider standard deviations across all evaluated features.

This statistical degradation perfectly aligns with the elastodynamic nature of the S_0 mode. The S_0 mode is dominated by in-plane displacements with a nearly uniform stress profile across

Table 4.2: Global mean R^2 scores and standard deviations for the extracted S_0 features across all evaluated bonding strategies.

S0 Feature	**Mean R^2**	**Standard Deviation (σ)**
A_0 Transition 2	0.571	0.207
S_0 Transition	0.569	0.242
A_0 Reflection 3	0.566	0.226
A_0 Reflection 1	0.515	0.322
A_0 Reflection 2	0.511	0.350
A_0 Transition 1	0.402	0.183

the plate thickness. Consequently, this mode exhibits an intrinsic low sensitivity to purely surface-mounted heterogeneities, such as the applied Pseudo-Damages. For the S_0 wave to effectively interact with the surface mass, the interfacial adhesive must facilitate a strong shear stress transfer.

Because the bonding strategy, and thus the shear stiffness varied significantly across the DoE, the interaction fluctuated wildly. This is mathematically reflected in the high standard deviations, confirming that the adhesive layer dictates the physics of the S_0 scattering far more heavily than the actual footprint d or mass ρ of the defect. Despite this heightened complexity,

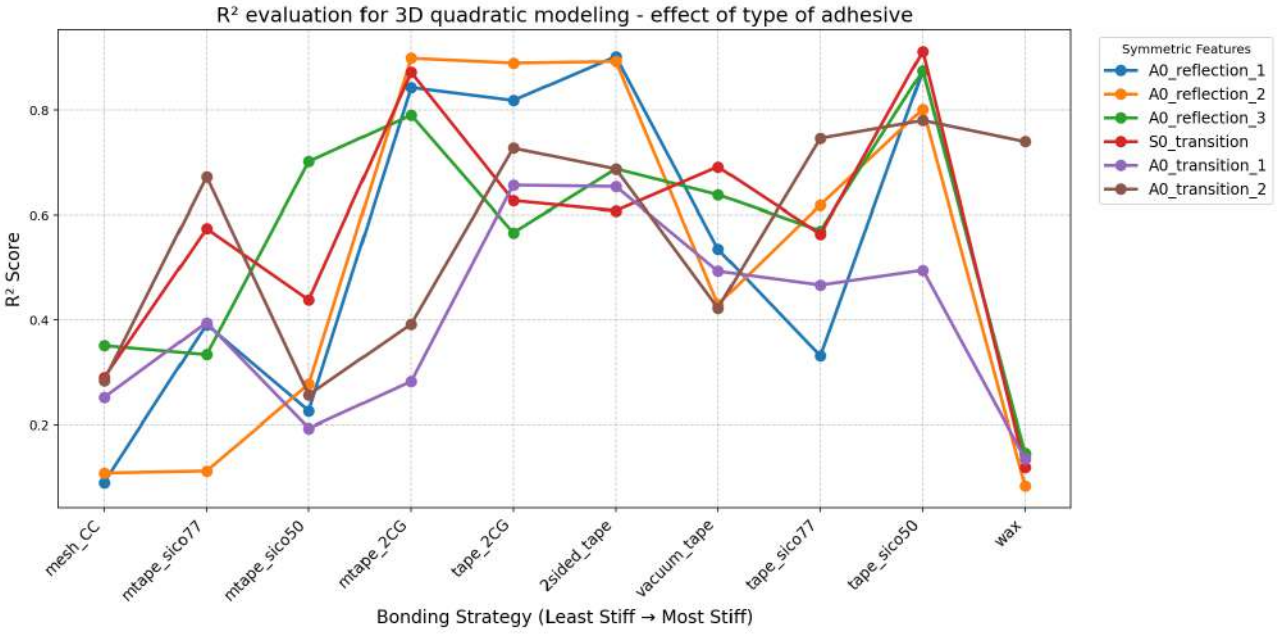


Figure 4.8

the forward-scattering features, specifically the mode-converted *A0_transition_2* and the direct *S0_transition*, retain the highest diagnostic robustness. The prominence of *A0_transition_2* as the top-ranking feature (Mean $R^2 = 0.571$) highlights that the asymmetric mode conversion phenomenon ($S_0 \rightarrow A_0$), triggered by the geometric asymmetry of the surface damage, scales more predictably with the defect's size and density than any backward reflection. Conversely, the reflected energies prove to be highly chaotic and strictly dependent on the specific PD conditions, reiterating their unsuitability for reliable, generalized damage quantification.

Let us undress from the mathematical framework built until this point and let us reconnect to the physical scenario that with the experiments has been reproduced. the specimen is the mean by which A0 and S0 Lamb modes, respectively flexural and extensional waves, travel and eventually encounter the Pseudo-Damage.

The A_0 mode produces an out-of-plane displacement through the specimen. When it encounters the Pseudo-Damage, it essentially forces the mass to move up and down, acting as a flexural wave. In this kind of dynamics, whether the adhesive is stiff or compliant plays a secondary role. The physics of the scattering phenomenon, and so what is transmitted and reflected, is primarily governed by the geometrical footprint d and the inertial resistance ρ, which together define the effective mass of the PD. As a result, the bivariate quadratic models based solely on diameter and density are highly accurate: the R^2 values exhibit a high mean and a low standard deviation across all bonding strategies, proving that these two geometrical and material variables strictly rule the A_0 acoustic features.

For the S_0 mode, the physics changes completely. The S_0 wave propagates with an in-plane displacement, stretching and compressing the plate horizontally[5] without inducing a direct out-of-plane movement. Consequently, the only way the S_0 wave "sees" the surface-mounted PD is through the shear stress transfer at the interface just under the defect. If the adhesive is soft, the plate simply slips beneath the PD; the wave passes through with minimal energy loss. Conversely, if the adhesive is stiff, the specimen pulls violently on the PD's base, generating high interfacial shear stresses and significant energy scattering.

This shear-dependent behavior perfectly explains the statistical results for the S_0 excitation: because the S_0 interaction is heavily dictated by the adhesive stiffness, attempting to model it using only diameter and density yields a low mean R^2 and a remarkably high standard deviation, highlighting how differently the wave behaves just by changing the adhesive tape.

Furthermore, a closer look at the extracted metrics reveals that, apart from the direct S_0 transmission, all the other parameters are mode-conversion-related features. When the in-plane S_0 wave hits the asymmetric structural discontinuity created by the PD, the structural imbalance forces a mode conversion, transforming part of the extensional energy into flexural A_0 waves. The efficiency of this conversion is entirely bound to the mechanical coupling provided by the adhesive. The high variance in the R^2 scores visually confirms this: mode-converted features are fundamentally a function of the interface characteristics, rendering them highly sensitive to the bonding strategy rather than just the PD's inherent weight and size.

[5] A helpful analogy is visualizing the A_0 mode as a transversal sea wave, and the S_0 mode as a longitudinal accordion wave.

Chapter 5

Barely Visible Impact Damage analysis

5.1 Impact testing

A Barely Visible Impact Damage (BVID) is a certain type of damage that occurs in thin plate-like composites when subjected to impact strikes with objects of relatively low velocity and masses. BVID represents a critical threat in composite structures: while the external surface may show only a microscopic indentation, the internal layers suffer extensive delamination and matrix cracking. When a BVID occurs, from the point of contact and throughout the whole specimen, a cascade of coned-shape defects are formed[1]. Although invisible to visual inspection, these defects constitute a major threat to the structural integrity of CFRP elements, drastically reducing their residual compression strength due to the onset of local instabilities and premature buckling of the fibers in the delaminated regions.[2].

The central objective of this experimental phase is the generation of controlled BVID on the composite specimens to establish a rigorous benchmark for the Pseudo-Damage investigation.

In this direction, understanding the interaction between Lamb waves and this specific defect morphology is crucial for evaluating the fidelity of the artificial surrogates.

The inherent asymmetry of the conical damage relative to the laminate's neutral axis forces a redistribution of the wave stress field exibit during the Lamb modes excitation, leading to significant mode conversion phenomena where incident symmetric modes S_0 are partially converted into antisymmetric packets A_0[3].

Furthermore, the delaminations effectively separate the laminate into thinner sub-layers that can act as local waveguides: as the ultrasonic pulse traverses this region, a portion of the energy becomes wrapped into these sub-laminates, reverberating between the boundaries. This phenomenon, known as the "trapped mode" effect, delays the energy transmission and manifests in the received signal as a combination of amplitude attenuation and phase distortion.

To induce controlled structural damage, three of the ten CFRP plates, cleaned from the pseudo-damage test campaign, are selected. The experimental setup utilized a mobile impactor manufactured by Ingegnieurtechnitsche Dienleistunger (ID)-Lindner.

[1]The stress waves arising from the point of impact generate a series of micro-delaminations through the thickness, the area of which increases with the distance from the contact point.

[2]Under compressive loads, the delaminations reduce the effective buckling load of the sub-laminates, leading to global structural failure at loads significantly lower than the design limit.

[3]Recalling Figure 3.9.

Figure 5.1: ID Lindner Impactor 013.

This device is a specialized pneumatic gas gun designed to deliver highly-controlled energy, localized impacts to test specimens. Such impactors are primarily employed to evaluate the damage tolerance of advanced composite materials and bonded joints. By accelerating a guided projectile using compressed air, the device ensures that the kinetic energy transferred to the target is strictly regulated by the input pressure, eliminating the variability inherent in manual striking methods.

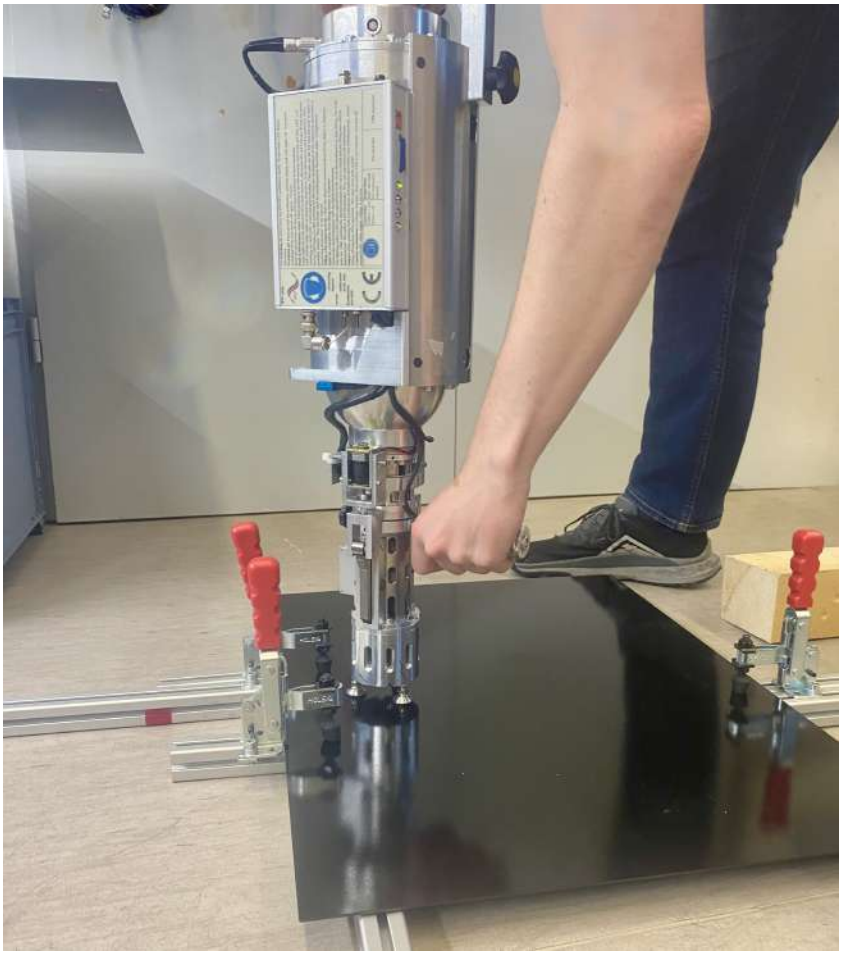


Figure 5.2: Impact Testing, pratically.

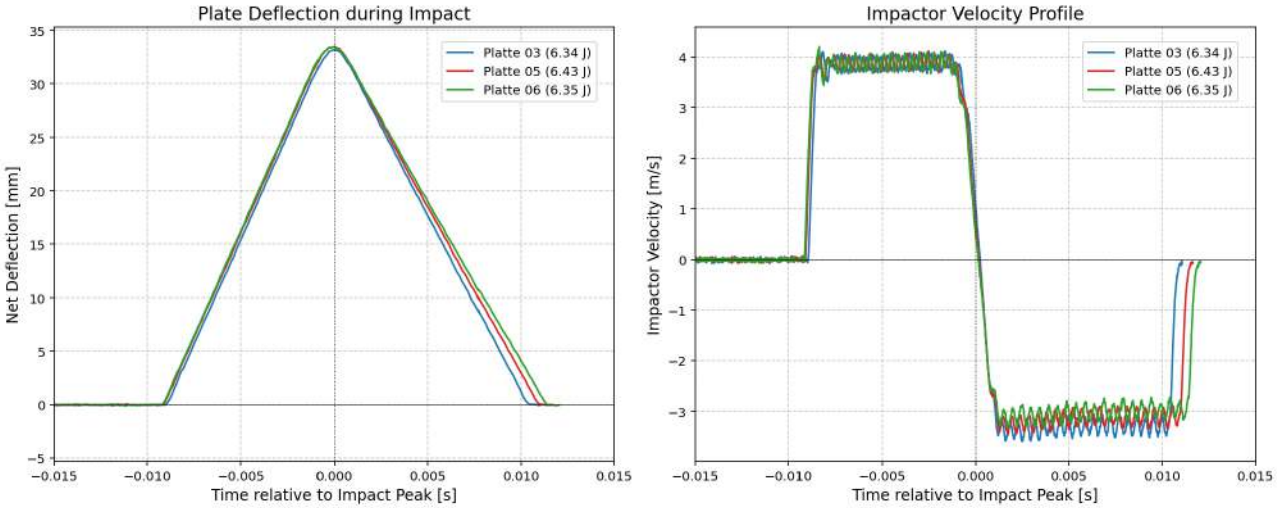


Figure 5.3: Impact Testing, deflection and velocity profiles.

The recorded kinematic data, presented in Figure 5.3, is representative of the three impact tests. The physics of the collision can be analyzed through two distinct phases visible in the plot

1. *Triangular deflection profile:* The deflection history in Figure 5.3 exhibits a distinct triangular profile, characterized by a linear ascent to a peak displacement of approximately 33 mm, followed by a similarly linear descent. Mathematically, a linear variation of displacement over time ($s \propto t$) implies a constant velocity. The dynamics are dominated by the projectile mass: the laminate is deflected at a near-constant velocity until the point of maximum potential energy, where an instantaneous reversal of motion occurs, typical of brittle-elastic behavior.

2. *Velocity step and energy absorption:* The velocity profile on the right in Figure 5.3, mirrors the deflection behavior. The impactor maintains a nearly constant velocity of $\approx 4\,\mathrm{m/s}$ during the approach and indentation phase. At the point of maximum deflection ($t = 0$), the velocity undergoes a rapid reversal. Crucially, the magnitude of the rebound velocity ($|v_{out}| \approx 3.2\,\mathrm{m/s}$) is distinctively lower than the impact velocity ($v_{in} \approx 4.0\,\mathrm{m/s}$).

The discrepancy between the inbound and outbound velocities provides quantitative evidence of the damage formation. The loss in kinetic energy, ΔE, can be estimated as:

$$\Delta E = \frac{1}{2} m_{head}(v_{in}^2 - v_{out}^2) \tag{5.1}$$

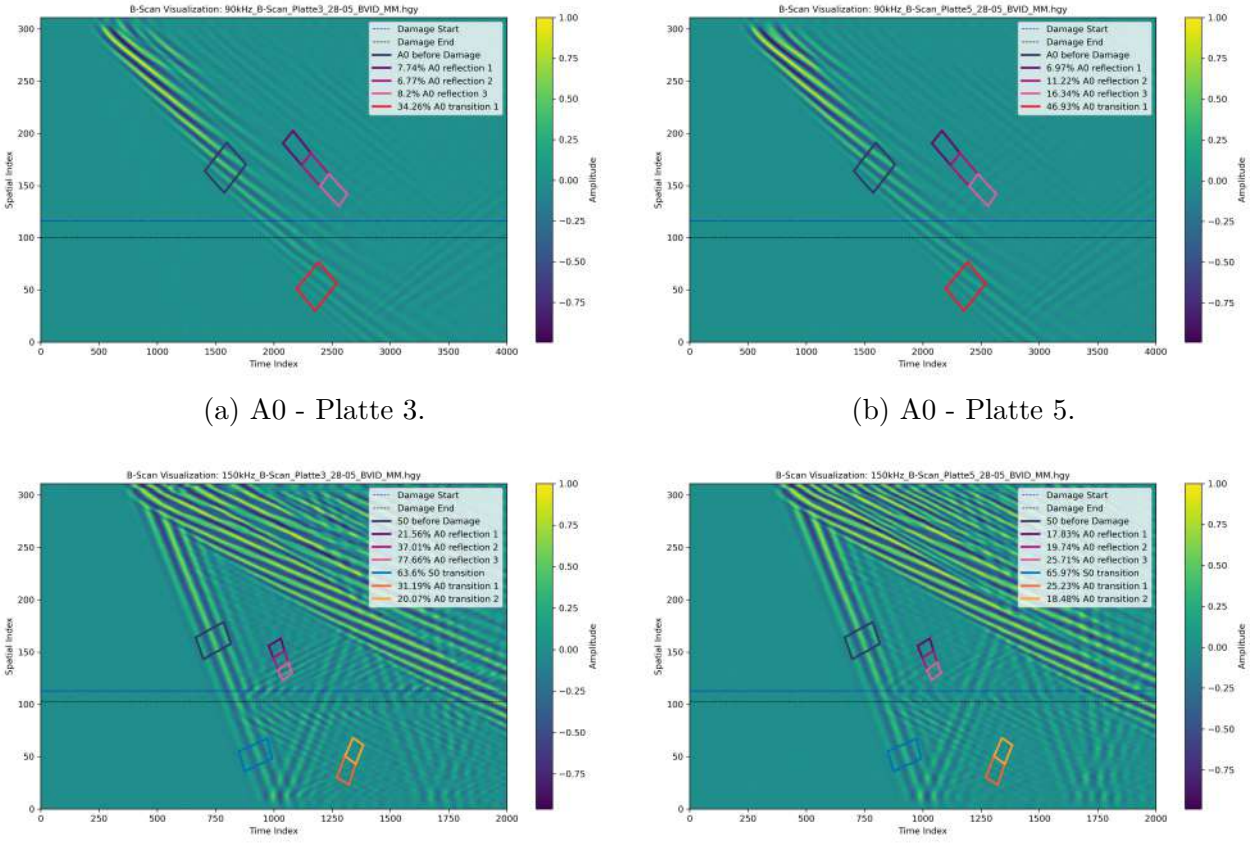
Since friction and viscous damping alone are insufficient to account for this drop in such a short impulse, the missing energy is attributed to the irreversible work done to fracture[4] the internal matrix and delaminate the plies, confirming the successful generation of BVID within the three samples.

[4]Note in the plot on the left how the slopes changes once the maximum deformation is reached.

5.2 B-scans of BVID configurations

5.2.1 Feature extraction

With the physical presence of Barely Visible Impact Damage (BVID) confirmed by the kinematic impact data, the next critical step is to quantify its effect on the propagation of ultrasonic guided waves. To ensure a rigorous and unbiased comparison, the identical feature extraction algorithms developed for the Pseudo-Damage analysis were deployed on the impacted specimens.



(a) A0 - Platte 3.

(b) A0 - Platte 5.

(c) S0 - Platte 3.

(d) S0 - Platte 5.

Figure 5.4: B-scans of two of the three BVID cases (Platte 3 and Platte 5) highlighting the interaction with the fundamental A0 (top row) and S0 (bottom row) Lamb modes.

The complex, and poorly known, vibroacustic signature of the defect is translated into the same feature space used for the PD configurations, populated by the DoE approach.

By collecting $\underline{x}_{A0}^{BVID}$ and $\underline{x}_{S0}^{BVID}$ for the three impacted plates and so by matching the BVID characteristics into the predefined feature space, it becomes possible to objectively assess their similarities. The fact that the impact energies and testing conditions were tightly controlled with minimal variations allows for a robust statistical assessment of the extracted features.

As a quantitative baseline for the genuine BVID, these statistical results are summarized in Table 5.1. The extracted parameters are categorically divided according to the fundamental Lamb wave mode used for the excitation due to their distinct physical interactions with the asymmetric "Pine Tree" damage morphology.
In an optimistic world, the acoustic signature of a BVID should be a deterministic property, identical impact energies applied to nominally identical plates are expected to produce consistent and identical scattering behaviors. However, the experimental data summarized in Table 5.1 explicitly contradicts this idealized assumption: the generation of three nominally identical impacts resulted in a highly scattered acoustic response, as evidenced by the significant standard deviations observed. This highlights the inherently stochastic nature of real impact damages, where unpredictable microscopic variations in matrix cracking and delamination boundaries heavily heavily perturb the wave interaction.

Table 5.1: Statistical summary of the extracted features.

Feature vector	**Extracted feature**	**Mean μ**	**Std. Dev. σ**
$\underline{\mathbf{x}}_{\mathbf{S0}}^{\mathbf{BVID}}$	A0 Reflection 1	19.71 %	1.87 %
	A0 Reflection 2	25.30 %	10.15 %
	A0 Reflection 3	39.62 %	33.34 %
	S0 Transition	66.12 %	2.60 %
	A0 Transition 1	27.13 %	3.52 %
	A0 Transition 2	20.07 %	1.59 %
$\underline{\mathbf{x}}_{\mathbf{A0}}^{\mathbf{BVID}}$	A0 Reflection 1	7.39 %	0.38 %
	A0 Reflection 2	9.57 %	2.44 %
	A0 Reflection 3*	12.15 %	4.08 %
	A0 Transition 1	40.92 %	6.36 %

* A standard deviation of 33.34% over a mean of 39.62 % of the reflected energy was observed, not due to a flaw in the defined feature, but rather an external factor.

Acknowledging this physical variability, a purely deterministic comparison between real and artificial damage is unfeasible. Therefore, the similarity assessment must be grounded in a statistical framework. By averaging the extracted features to define a "mean BVID" centroid, and utilizing the standard deviation as a direct metric of the natural physical scatter, a probabilistic baseline is constructed. Only through this statistical lens is it possible to perform a reliable comparative evaluation against the pseudo-damages of the DoE, eventually determining which specific combination of mass, diameter, and adhesive stiffness most accurately replicates the complex, non-deterministic phenomena just characterized.

5.3 Defect sensitivity of S0 mode

As it was already reported in the statistical summary 5.1, the feature corresponding to the A0-to-S0 converted reflected energy exhibits a significant standard deviation across the three identically impacted specimens. Specifically, while Plates 5 and 6 yielded highly consistent values for this converted mode, Plate 3 registered a substantially higher converted energy. This

variance is not indicative of a flaw in the feature itself, but rather highlights the high sensitivity of guided wave scattering to the exact through-thickness morphology of the BVID.

To adress this phenomenon, the underlying physics of Lamb wave mode conversion must be considered. In a pristine, symmetric waveguide, the symmetric (S_0) and anti-symmetric (A_0) Lamb wave modes are mathematically orthogonal. Consequently, an incident A_0 wave will propagate without converting into an S_0 wave, as the boundary conditions at the upper and lower surfaces remain perfectly balanced.

Table 5.2: Comparison of the $A_0 \rightarrow S_0$ reflected energy across the three identically impacted CFRP specimens.

Specimen	**Converted Reflected Energy - A0 Reflection 3**
Plate 3	77.61 %
Plate 5	25.71 %
Plate 6	15.49 %

When an incident A_0 wave interacts with a defect, the wave must partition its energy to satisfy the new traction-free boundary conditions created by the crack faces. If a delamination is located exactly at the the neutral plane of the laminate, the geometrical symmetry is preserved, and the A_0 wave will primarily scatter as an A_0 wave.

However, if the delamination is offset from the midplane, the through-thickness symmetry is broken, effectively dividing the plate into two asymmetric sub-laminates. To satisfy the boundary conditions across this asymmetric discontinuity, the purely anti-symmetric stress field of the incident A_0 wave can no longer be solely supported, forcing a portion of the energy to convert into the symmetric S_0 mode. The magnitude of this mode-converted reflected energy is directly proportional to the degree of geometrical asymmetry.

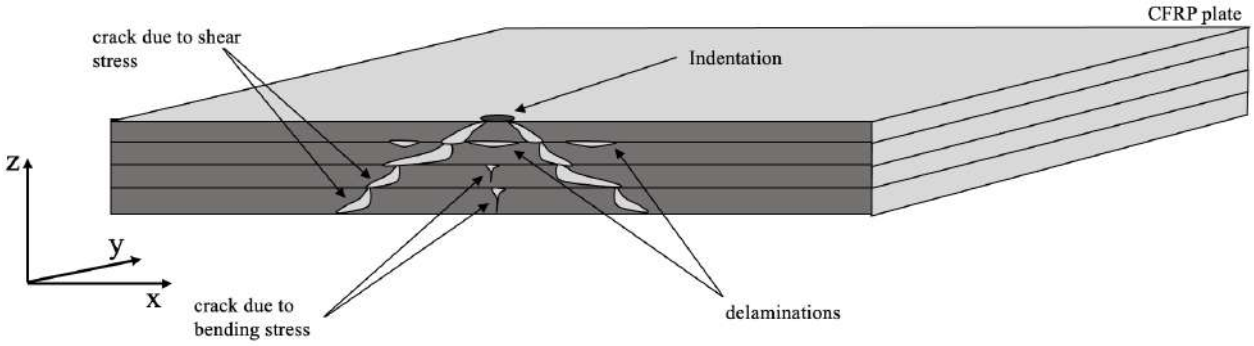


Figure 5.5: BVID internal defect formation.

BVID in composite laminates typically exhibits a complex, conical "pine-tree" morphology, comprising a network of matrix cracks and delaminations. Despite subjecting the plates to identical impact energies, the local heterogeneity of the composite material can dictate variations in the damage depth profile.

To validate the hypothesis that the higher converted energy in Plate 3 is due to an offset delamination, ultrasonic C-scans were performed on the specimens.

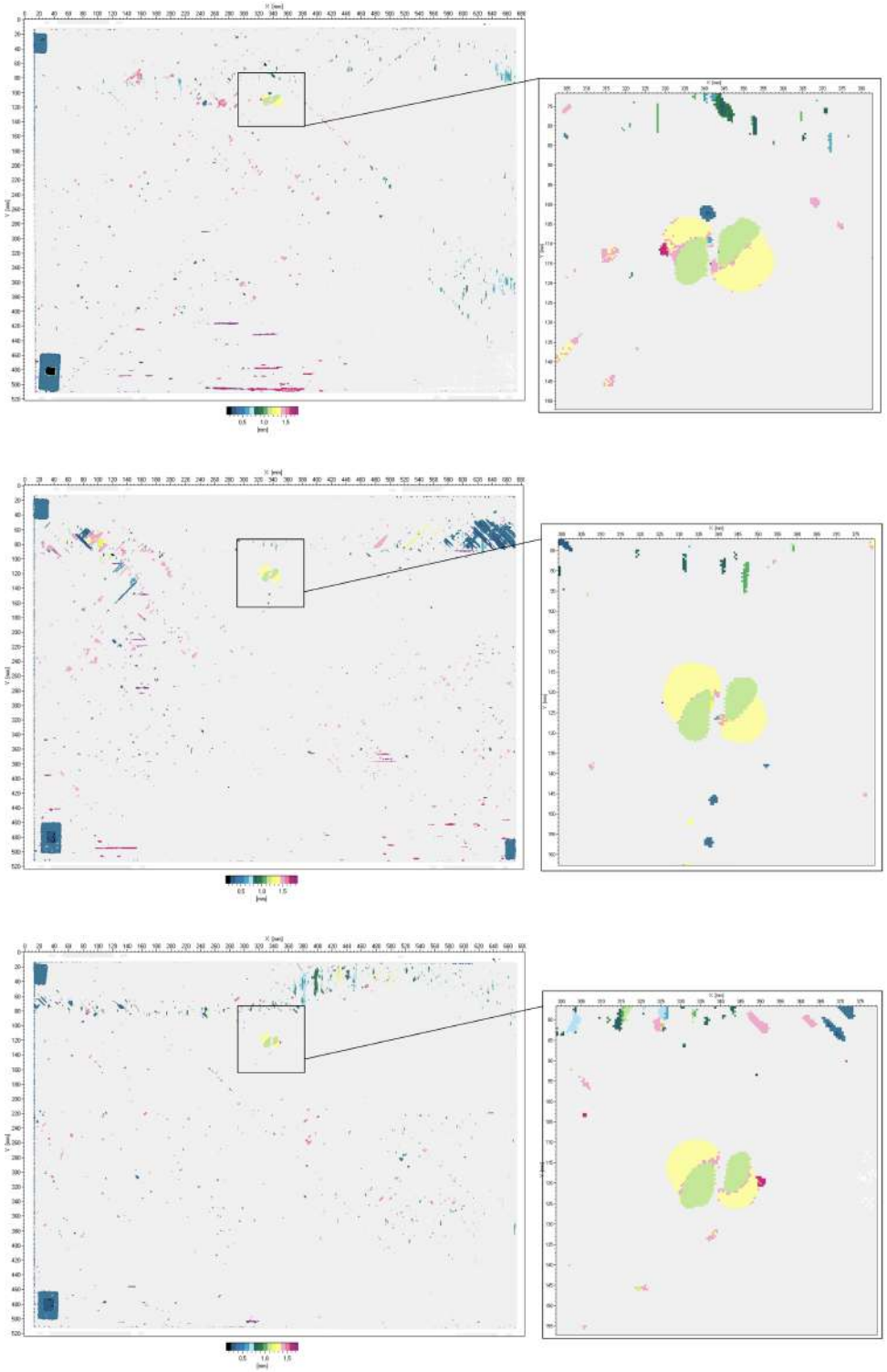


Figure 5.6: Ultrasonic C-scans mapping the delamination depth. Top: Profile of Plate 3, revealing a primary delamination significantly offset towards the surface. Middle & Low: Representative profiles of Plate 5 and Plate 6, showing damage centered near the midplane.

C-scan results, presented in Figure 6.2, provide clear visual confirmation. The scans map the depth of the primary reflective interfaces within the laminates. For Plates 5 and 6, the largest delaminations are distributed relatively symmetrically around the midplane of the plate. In contrast, the C-scan for Plate 3 reveals that the impact energy caused the primary delamination to open at a ply interface significantly offset towards the impacted surface.

This pronounced geometrical asymmetry in Plate 3 directly breaks the through-thickness symmetry to a greater extent than in Plates 5 and 6, thereby dictating a stronger mode conversion and perfectly explaining the higher amplitude of the reflected S_0 energy captured in the feature extraction phase.

Chapter 6

Analysis in the Feature Space

6.1 Metric Definition & Similarity assessment

To quantitatively and objectively evaluate which artificial pseudo-damage configuration best replicates the scattering behavior of a real barely visible impact damage, a multivariate statistical framework was built.

Let $N_{features} = 10$ be the total number of extracted features, comprising both A_0 mode-based and S_0 mode-based features. Any specific experimental acquisition inside the experimental space prescribed by the DoE can be mathematically represented as a multi-dimensional row vector in the feature space $\underline{x}_i^{PD} \in \mathbb{R}^{N_{features}}$:

$$\underline{x}_i^{PD} = [\underline{x}_{A0_i}^{PD} \underline{x}_{S0_i}^{PD}] = [f_{1,i}, f_{2,i}, \ldots, f_{10,i}] \tag{6.1}$$

where each element $f_{k,i}$ represents the normalized amplitude of the k-th feature for that specific experiment.

To define the target fingerprint each of the three real BVID experiments can be seen again as a row vector in the same feature space $\underline{x}_i^{BVID} \in \mathbb{R}^{N_{features}}$: those vectors are three different vectors in $\mathbb{R}^{N_{features}}$, as noted in the last chapter, a possible way to tackle the similarity assessment is to define the mean vector in the feature space.

Let $N_{BVID} = 3$ be the total number of tested impact damages: the centroid of this distribution, representing the expected mean response of a real impact, is defined by the sample mean vector $\underline{\mu}^{BVID} \in \mathbb{R}^{N_f}$:

$$\underline{\mu}^{BVID} = \frac{1}{N_{BVID}} \sum_{i=1}^{N_{BVID}} \underline{x}_i^{BVID} \tag{6.2}$$

The same reasoning can be used to find the centroid of the PD distribution, which is defined by the sample mean vector $\underline{\mu}^{PD} \in \mathbb{R}^{N_f}$:

$$\underline{\mu}^{PD} = \frac{1}{N_{PD}} \sum_{i=1}^{N_{PD}} \underline{x}_i^{PD} \tag{6.3}$$

Furthermore, the natural experimental dispersion and the physical cross-correlation between the different acoustic features across the DoE are captured by the sample covariance matrix:$\underline{\underline{\Sigma}}^{PD} \in \mathbb{R}^{N_f \times N_f}$:

$$\underline{\underline{\Sigma}}_{PD} = \frac{1}{N_{PD} - 1} \sum_{i=1}^{N_{PD}} (\underline{x}_i^{PD} - \underline{\mu}^{PD})(\underline{x}_i^{PD} - \underline{\mu}^{PD})^T \tag{6.4}$$

How one can define how close a PD configuration must be to act "as" a real BVID configuration? The reasoning up to this point led to the definition of a 10-dimensional feature space, where each of the PD configurations defined in the DoE is associated with a specific feature vector. It therefore seems reasonable to define a mathematical norm in this space to evaluate how close the PD feature vector is to the baseline real impact.[1]

The most intuitive metric to evaluate this spatial proximity is the standard Euclidean distance D_E. For a generic Pseudo-Damage vector $\underline{x}_{PD}$, its geometric distance from the centroid of the real impact distribution $\underline{\mu}_{BVID}$ is defined as:

$$D_E(\underline{x}^{PD}) = \sqrt{(\underline{x}^{PD} - \underline{\mu}^{BVID})^T (\underline{x}^{PD} - \underline{\mu}^{BVID})} \tag{6.5}$$

While geometrically straightforward, D_E relies on the strict assumption that the feature space is isotropic and orthogonal. It treats all acoustic features as statistically independent and equally scaled. However, the extracted ultrasonic features exhibit vastly different absolute magnitudes and intrinsic experimental variances, as seen in Chapter 4. In such an anisotropic environment, the Euclidean metric is inherently biased: it disproportionately penalizes deviations along features with naturally larger numeric scales, while ignoring the cross-correlation between features.

To rigorously address this limitation and rank the Pseudo-Damage configurations based on their true statistical similarity to the real impacts, the Mahalanobis distance D_M was utilized. This metric acts as a powerful generalization of the Euclidean distance. For a generic Pseudo-Damage feature vector $\underline{x}_{PD}$, it is computed as:

$$D_M(\underline{x}^{PD}) = \sqrt{(\underline{x}^{PD} - \underline{\mu}^{BVID})^T \underline{\underline{\Sigma}}_{PD} (\underline{x}^{PD} - \underline{\mu}^{BVID})} \tag{6.6}$$

where $\underline{\underline{\Sigma}}_{PD}$ is the inverse of the PD sample covariance matrix, also known as precision matrix.

[1] Ideally, no matter the norm, the optimal solution would be to find across the DoE at least one configuration $\underline{x}^*_{PD}$ such that $\underline{x}^*_{PD} - \underline{\mu}_{BVID,i} = \underline{0}$. Although optimistic, this line of reasoning proves simplistic just by looking at the intrinsic statistical scatter of the features across the BVID tests.

Mathematically, if the features were perfectly uncorrelated and had unit variance $\underline{\underline{\Sigma}}^{-1}_{BVID} = \underline{\underline{I}}$, the Mahalanobis distance would perfectly collapse into the standard Euclidean distance.[2]

Rank	Material	Adhesive	Diameter	Δ Features vs BVID centroid										$D_m(\underline{x}^{PD})$
				$\Delta R1_{A0}$	$\Delta R2_{A0}$	$\Delta R3_{A0}$	$\Delta T1_{A0}$	$\Delta R1_{S0}$	$\Delta R2_{S0}$	$\Delta R3_{S0}$	ΔT_{S0}	$\Delta T1_{S0}$	$\Delta T2_{S0}$	
1°	PVC	Wax	40 mm	+2.81	−1.21	−3.87	−3.75	−2.90	−1.90	−15.20	−3.20	−6.48	−4.45	0.99
2°	PVC	Vacuum Tape	30 mm	−2.06	−4.48	−7.62	+0.11	+0.61	−7.73	−23.73	−5.91	−9.30	−6.41	1.07
3°	PVC	Wax	25 mm	+2.44	−1.54	−4.56	−4.85	−4.21	−7.08	−21.54	−3.72	−11.93	−6.23	1.14
4°	PVC	Wax	30 mm	+2.46	−1.53	−4.27	−1.97	+1.58	−1.37	−21.45	−4.41	−7.04	−6.40	1.17
5°	PVC	Vacuum Tape	25 mm	+2.01	−0.83	−4.03	−1.50	+0.33	−7.40	−25.53	−4.41	−5.78	−4.69	1.18

Table 6.1: Top 5 pseudo-damage configurations ranked by Mahalanobis distance and their Δ Features.

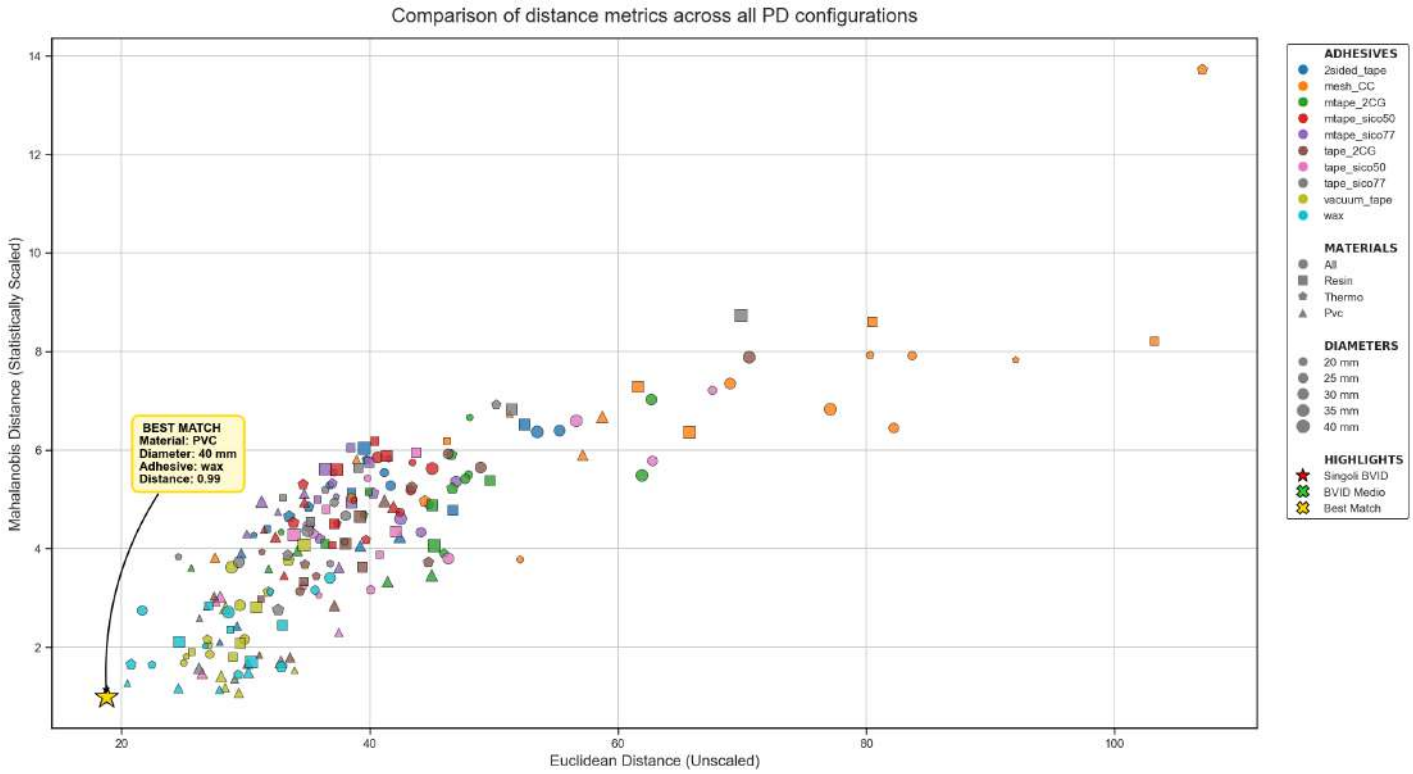


Figure 6.1: Biplot of Mahalanobis vs Euclidean Distance for each PD of the DoE.

[2]Dear collegue of Srinivasa Ramanujan, the Indian statistician Prasanta Chandra Mahalanobis in 1936 introduced the Mahalonobis Distance for the first time,. The same Mahalonobis saw the Mahalanobis distance as a way to "obtain" the Euclidean distance after a "whitening process", that is a trasformation that developes standardization and decorrelation.

The precision matrix $\underline{\underline{\Sigma}}^{-1}_{BVID}$ normalizes and rotates the space, penalizing spatial deviations in directions of tight, low-variance BVID clustering while tolerating deviations in directions of high experimental scatter. The "Best Match" is the PD configuration that minimizes D_M across the 10-dimensional space.

By plotting the results for each one of the PD configuration adressed in the DoE, one can see why the Mahalanobis metric was needed for this kind of analysis: if the data was isotropic with respect to each one of the ten defined features,looking at 6.1, the behavior of the cloud of points would be perfectly linear.

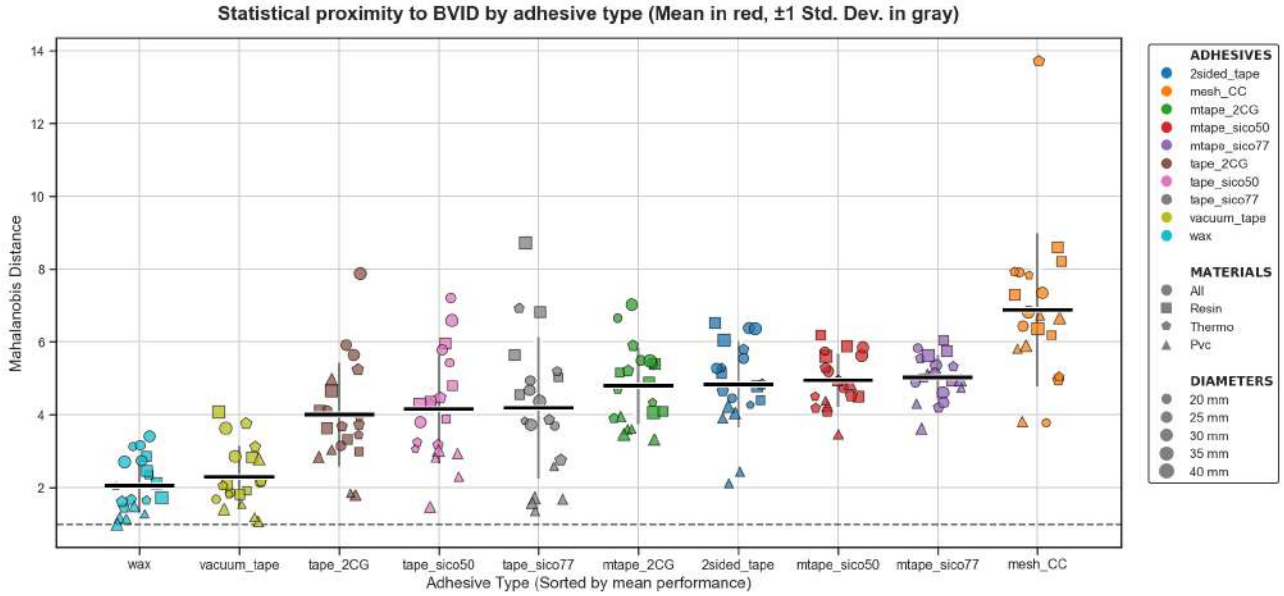


Figure 6.2: Statistical distribution of Mahalanobis distance, for each adhesive used.

As established in Table 6.1, the statistical evaluation within the physical feature space successfully isolated the pseudo-damage configuration comprising a PVC insert of 40 mm of diameter and bonded with wax as the absolute "Best Match". While the Mahalanobis metric minimizes the multi-dimensional discrepancy across the ten extracted features, it fundamentally remains a statistical abstraction that serves purely to identify the closest PD from the DoE to the mean BVID centroid. To definitively validate whether this selected candidate genuinely replicates the complex acoustic scattering induced by a real impact, it is imperative to return to the fundamental source of the data: the ultrasonic B-Scans.

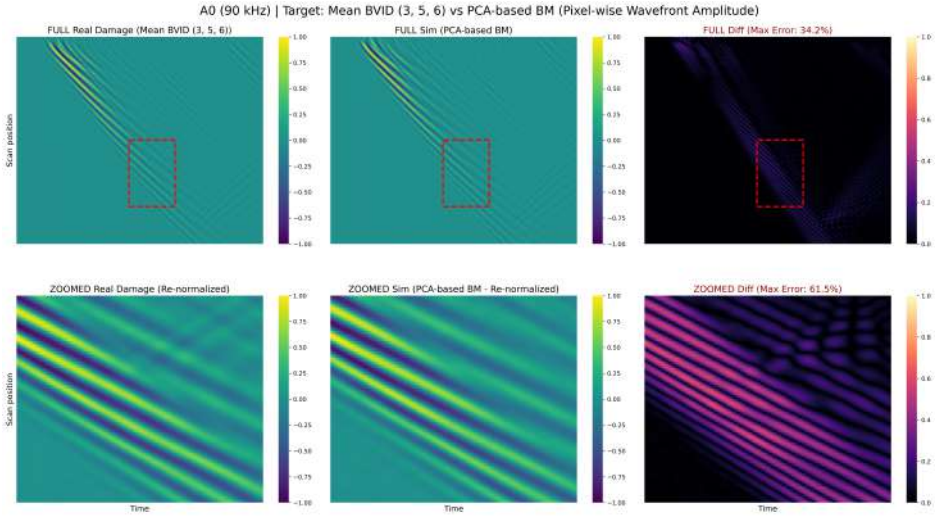
Consequently, this section transitions from a feature-space ranking to a direct, spatial comparative analysis. The acoustic signature of the identified Best Match is superimposed and subtracted from the experimental B-Scans of the real BVIDs: specifically, Plates 3, 5, and 6, as well as their computed mean. To comprehensively assess the morphological and energetic fidelity of the pseudo-damage, the evaluation is conducted through two distinct and complementary diagnostic lenses: a more general one that takes all the B-scan dimensions and a more concentrated one, which is based upon a ROI that highlight the scattering of Lamb modes that rises in the vicinity of the BVIDs and PDs.

As a first analytical step, the B-Scan comparison is performed on a pixel-wise basis, evaluating the raw wavefront amplitudes differences. A fundamental strength of the proposed is the rigorous initial calibration procedure described in Chapter 3: by meticulously aligning the coordinates of the incident wavefields across all acquisitions, the baseline signals are perfectly

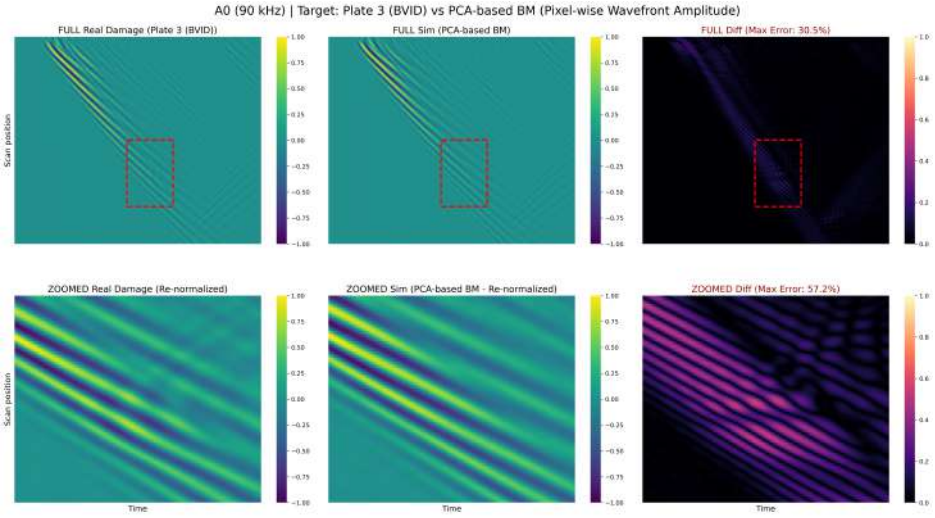
synchronized. Theoretically, this precise calibration should allow for a direct, point-by-point subtraction between the experimental BVID and the pseudo-damage surrogate to isolate the differential scattering.

Unfortunately, observing the pixel-wise residual error maps reveals substantial mathematical discrepancies. This phenomenon is deeply rooted in the complex elastodynamic nature of ultrasonic guided waves. Lamb waves are inherently dispersive; their phase and group velocities are highly sensitive to local frequency-thickness variations. When an incident wave packet interacts with a real BVID, it encounters a highly irregular, three-dimensional network of micro-cracks and multi-layer delaminations. This interaction induces severe local velocity alterations, multiple internal scattering, and consequent microscopic phase shifts. In contrast, a pseudo-damage configuration alters the local stiffness and thickness through a much more regular boundary condition. Consequently, even if the pseudo-damage successfully replicates the macroscopic scattering amplitude of the real impact, microscopic phase mismatches between the two oscillating signals are inevitable, expecially in the vicinity of the BVIDs and PDs positions, in which the different topologies of the two lead to a different, local Lamb modes behavior. Thus, in a pixel-wise difference, subtracting two slightly out-of-phase waves leads to severe artificial interference, both constructive and destructive. This generates high residual error peaks that are mere mathematical artifacts[3], rather than genuine indicators of physical dissimilarity.

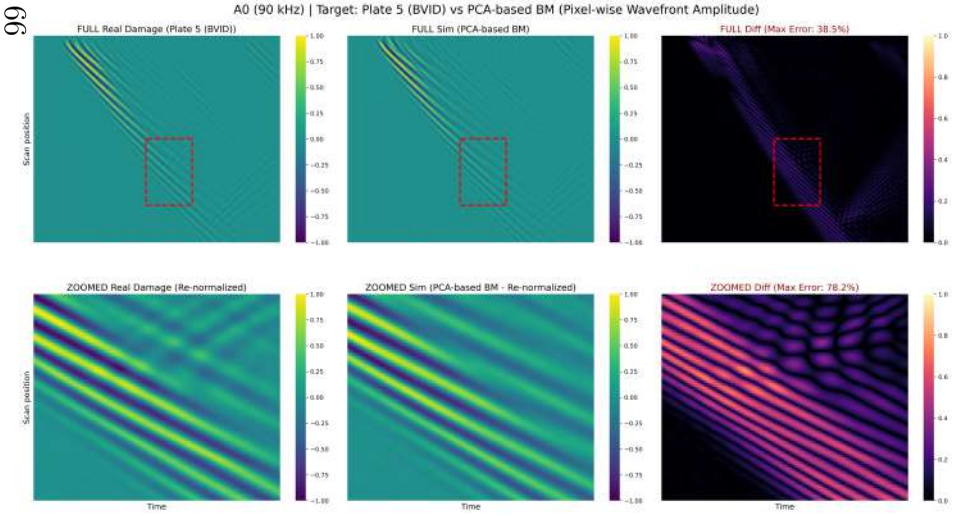
[3]See the titles in 6.5 and 6.6, which highlights the max percentage differences.



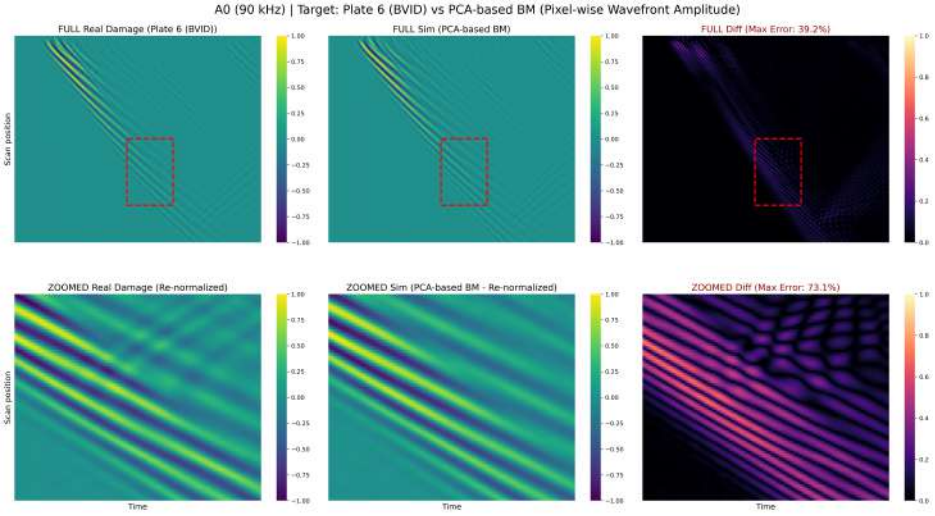
(a) Mean of the BVIDs.



(b) Plate 3 BVID.

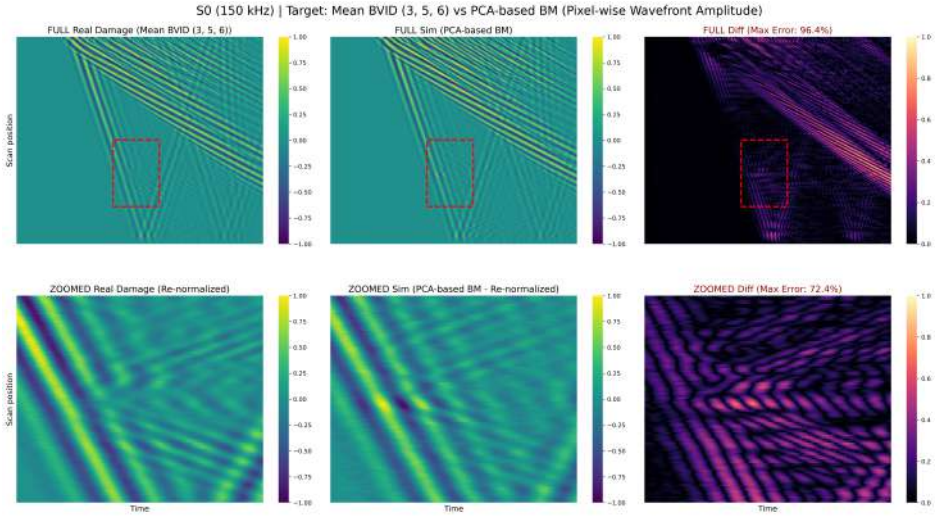


(c) Plate 5 BVID.

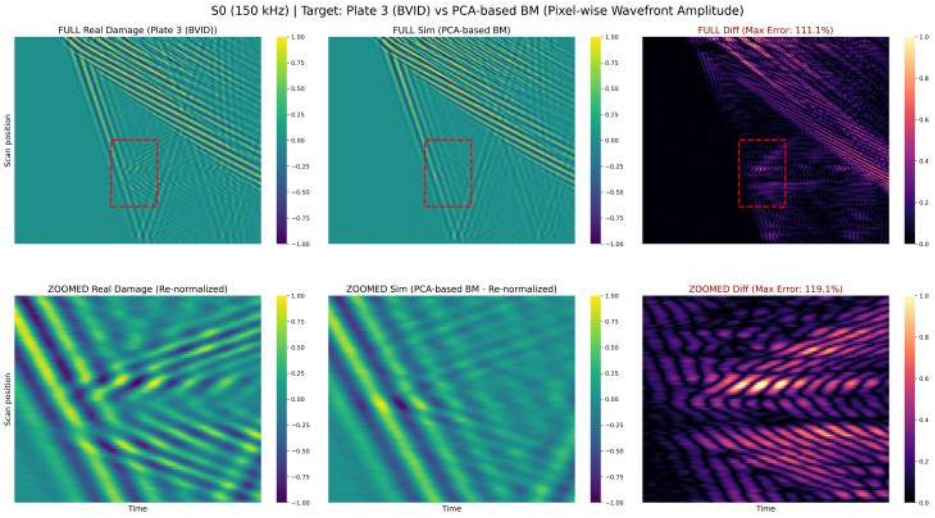


(d) Plate 6 BVID.

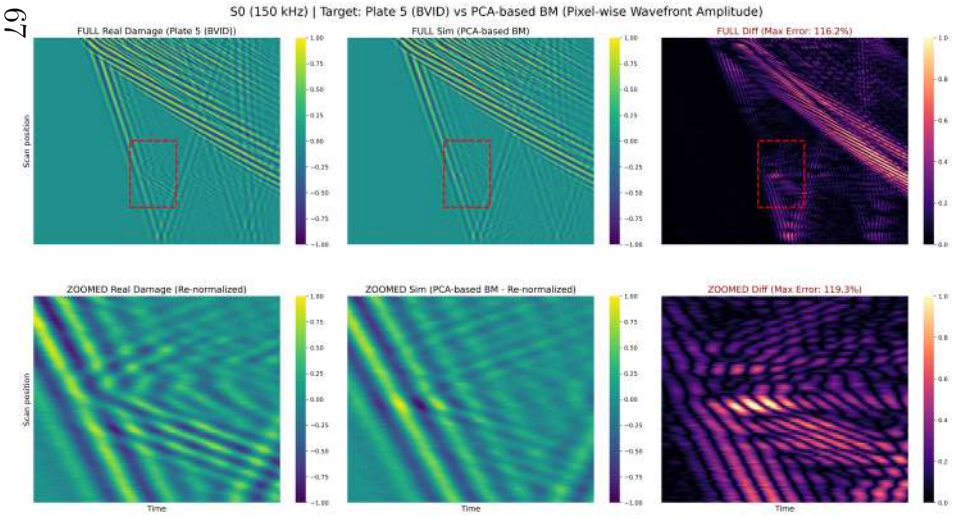
Figure 6.3: Pixel-wise comparative analysis for the A0 mode (90 kHz) between the PCA-based Best Match and the BVIDs.



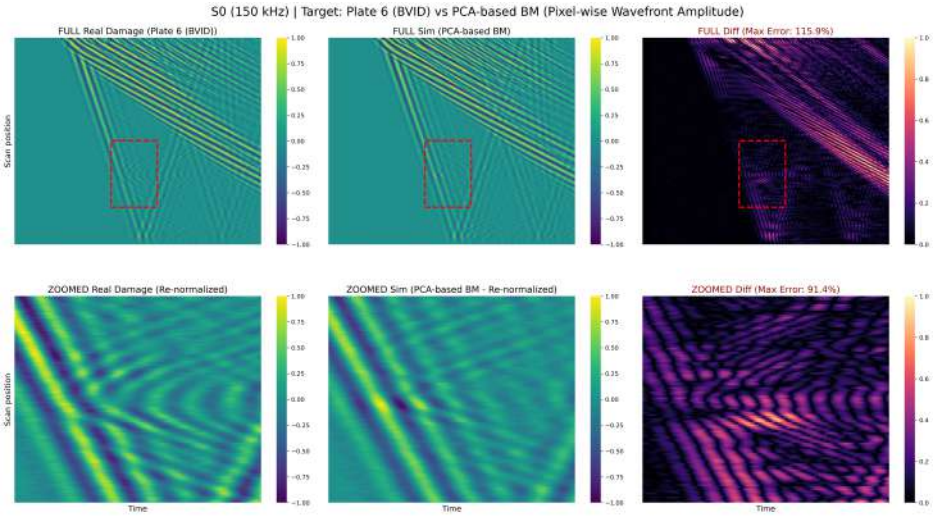
(a) Mean of the BVIDs.



(b) Plate 3 BVID.



(c) Plate 5 BVID.



(d) Plate 6 BVID.

Figure 6.4: Energy-wise comparative analysis for the S0 mode (150 kHz) between the PCA-based Best Match and the BVIDs.

To overcome these limitations the comparative analysis goes to an *energy-wise* strategy. The primary objective of this approach is to decouple the macroscopic scattering behavior of Lamb modes, acquired through the B-scans, from the highly volatile microscopic pixel-level phase shifts. To achieve this, the normalized ultrasonic amplitude $A(x,t)$ of each B-Scan is first squared, yielding the instantaneous acoustic energy density $E(x,t) = A(x,t)^2$. This mathematical operation effectively rectifies the signal, transforming the alternating positive and negative wave phases into a strictly positive energy representation.

However, the squared signal inherently retains high-frequency spatial and temporal ripples corresponding to the fundamental periods of the original Lamb wave modes. To extract the macroscopic energy envelope, a two-dimensional Gaussian smoothing filter is applied via convolution. The Gaussian window acts as a spatio-temporal low-pass filter, mathematically governed by its standard deviations σ_x and σ_t . This 2D convolution averages the local energy over a defined neighborhood, deliberately blurring the exact zero-crossings and phase crests while rigorously preserving the global energy distribution.

Mathematically, the smoothed macroscopic energy envelope $E_{env}(x,t)$, is obtained through the two-dimensional continuous convolution integral between the instantaneous energy density $A(x,t)^2$ and the Gaussian kernel. This is formally expressed as:

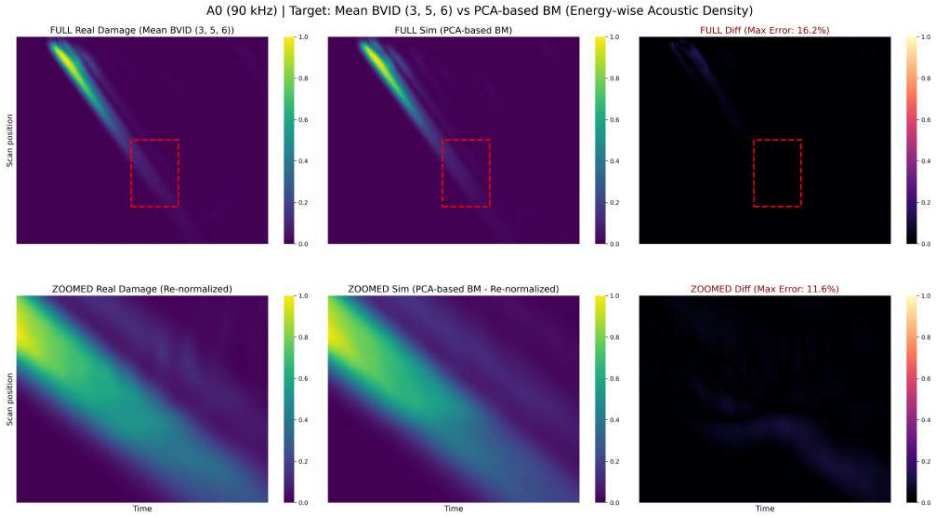
$$E_{env}(x,t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(x-\xi, t-\tau)^2 \cdot \frac{1}{2\pi\sigma_x\sigma_t} e^{\left(-\frac{\xi^2}{2\sigma_x^2} - \frac{\tau^2}{2\sigma_t^2}\right)} d\xi \, d\tau \tag{6.7}$$

where:

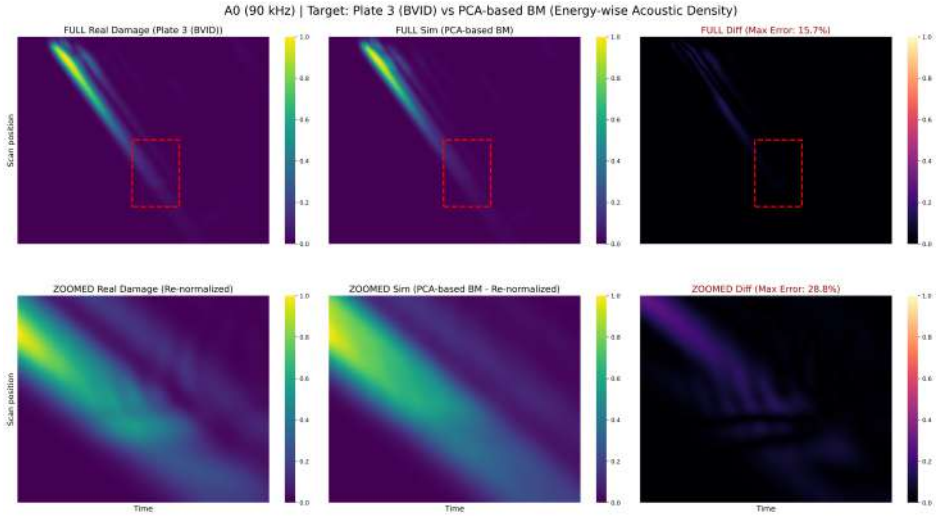
- $A(x-\xi, t-\tau)^2$ represents the shifted instantaneous acoustic energy;
- ξ and τ are the spatial and temporal integration variables for the convolution operation;
- σ_x and σ_t dictate the spatial and temporal spread of the Gaussian window, actively controlling the cut-off frequency of the equivalent low-pass filter.

An asymmetric window ($\sigma_t = 12$, $\sigma_x = 4$) was strategically selected because of the defined B-scans dimensions. This configuration effectively bridges the temporal wave crests to form a cohesive energy envelope, whilst strictly restricting spatial blurring. Consequently, the intricate multiple scattering signatures of the Lamb modes and the lateral contours of the damage are meticulously preserved. Fundamentally, this filtering strategy neutralizes the erratic phase interferences that previously corrupted the pixel-wise evaluation. By bypassing these microscopic temporal misalignments, the comparative framework is re-centered on the actual acoustic phenomena of interest: the magnitude of the reflected wavefield, the residual trapped energy, and the severity of the transmission shadow.

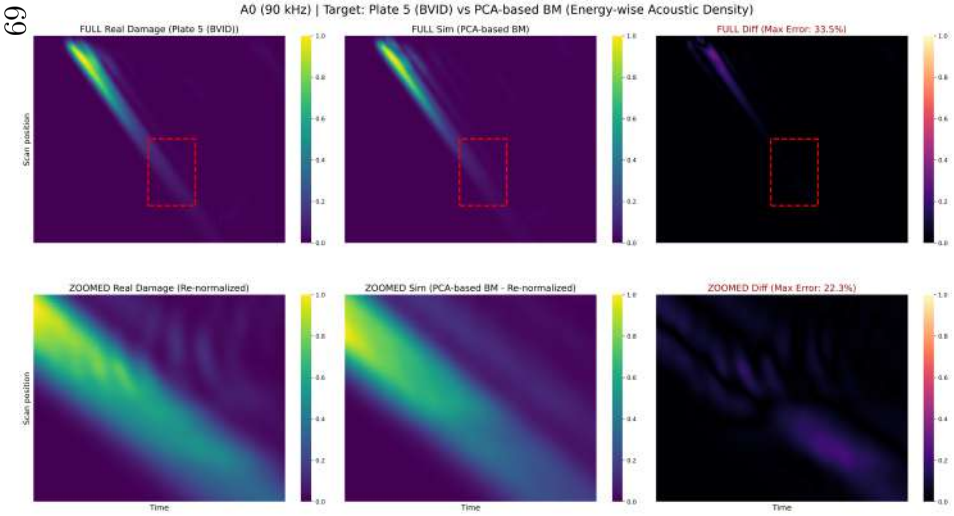
Having established this robust, energy-centric metric, the analysis now shifts to the visual and quantitative examination of the resultant residual error heatmaps. The following comparative plots map the spatial distribution of the acoustic energy differences between the physical-based optimal surrogate and the experimental BVID configurations, providing a definitive assessment of the pseudo-damage's physical fidelity.



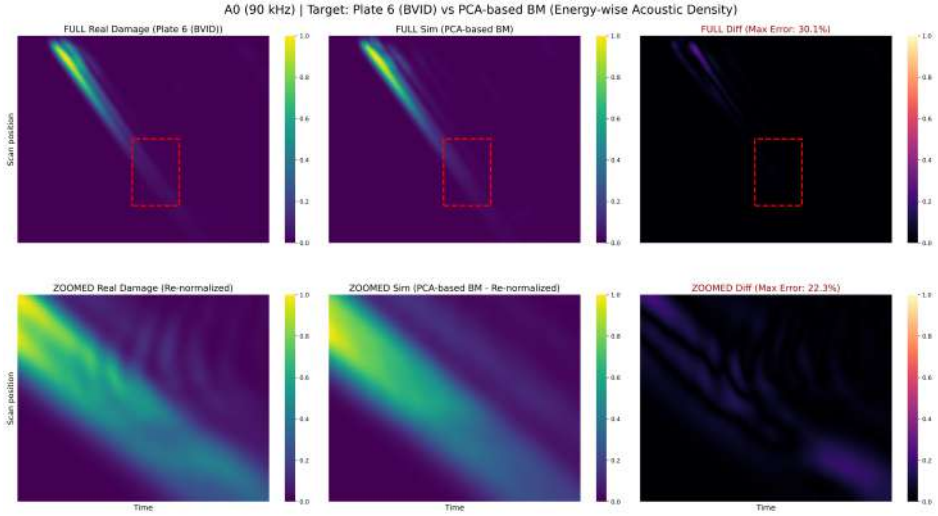
(a) Mean of the BVIDs.



(b) Plate 3 BVID.

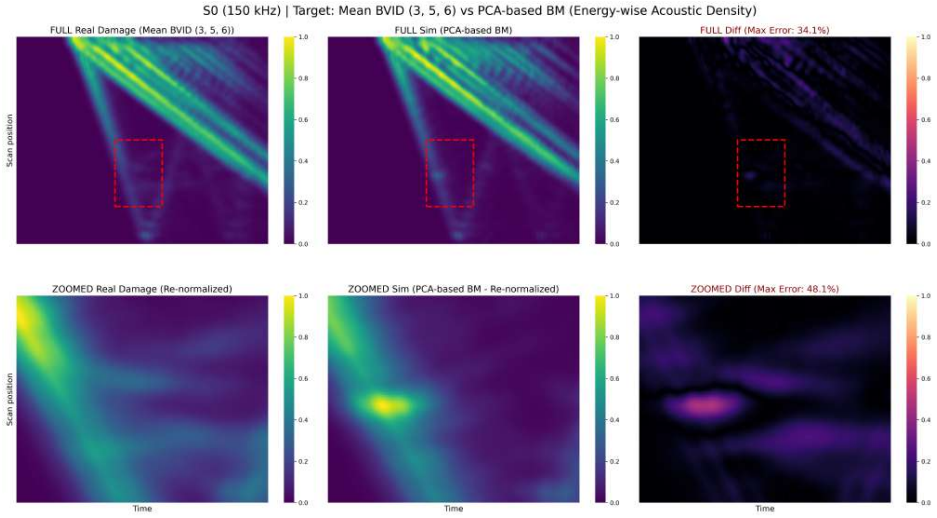


(c) Plate 5 BVID.

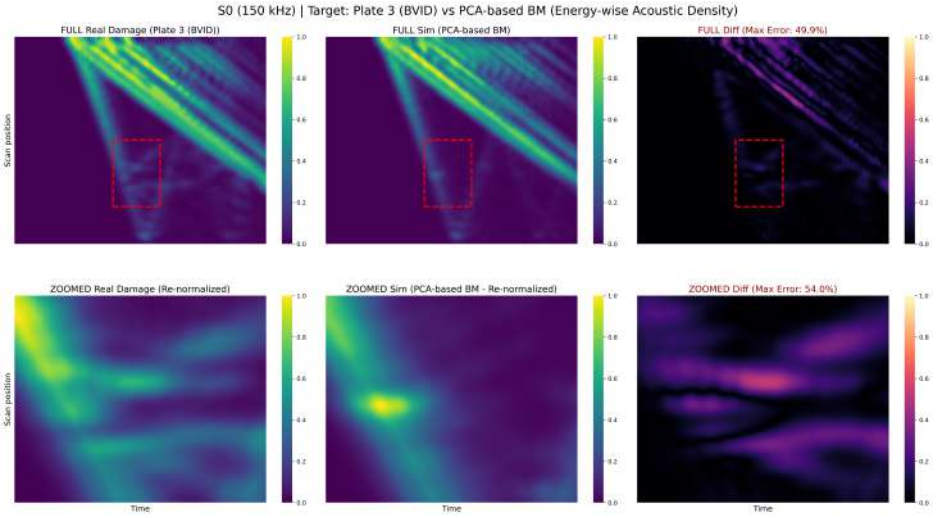


(d) Plate 6 BVID.

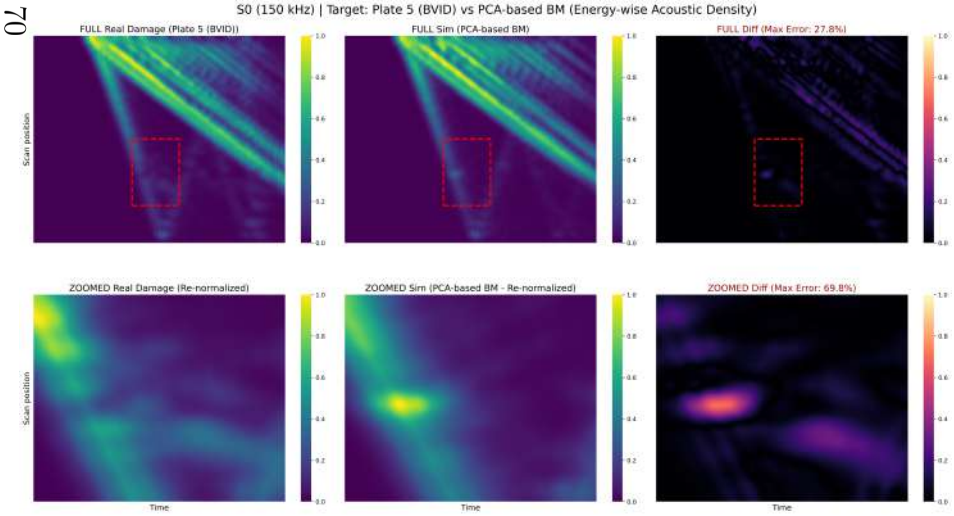
Figure 6.5: Pixel-wise comparative analysis for the A0 mode (90 kHz) between the PCA-based Best Match and the BVIDs.



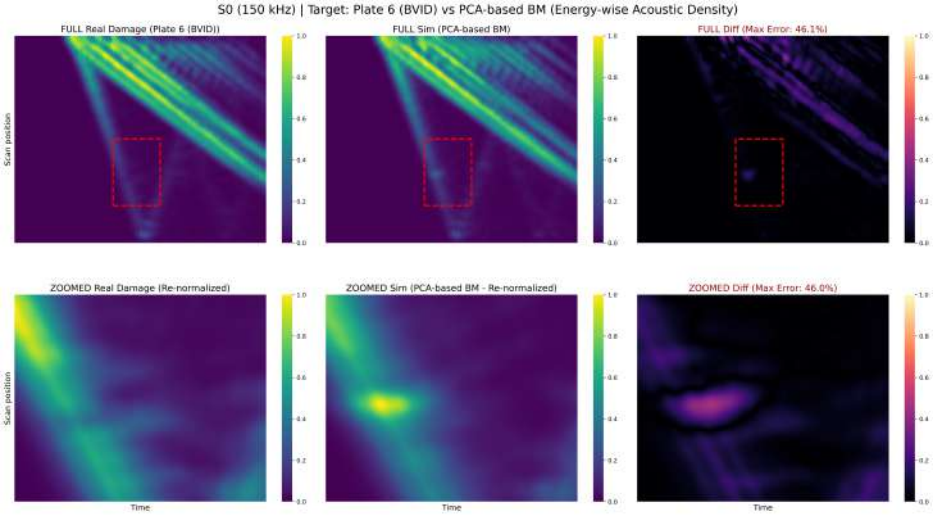
(a) Mean of the BVIDs.



(b) Plate 3 BVID.



(c) Plate 5 BVID.



(d) Plate 6 BVID.

Figure 6.6: Energy-wise comparative analysis for the S0 mode (150 kHz) between the PCA-based Best Match and the BVIDs.

For the sake of clarity, the consideration made in this context will take as a reference the heatmap for the best in class with respect to the mean BVID B-scan. Analyzing the comparative heatmaps for the A0 mode, the asymmetric Gaussian smoothing successfully eradicates the phase-induced Moiré artifacts. The topological agreement between the simulated pseudo-damage and the experimental mean BVID is remarkable, with the maximum localized residual error dropping significantly to 11.6%. The surrogate accurately replicates the overall transmission shadow and the macroscopic energy envelope, proving its validity at lower frequencies. This strong correlation can be physically justified by the fundamental mechanics of the A0 Lamb mode: being predominantly flexural and characterized by out-of-plane displacements, the A0 mode is highly sensitive to local variations in transverse inertia. The macroscopic addition of the PVC and wax patch effectively mimics the localized inertial alteration induced by a real impact, leading to a highly accurate representation of the A0 energy field.

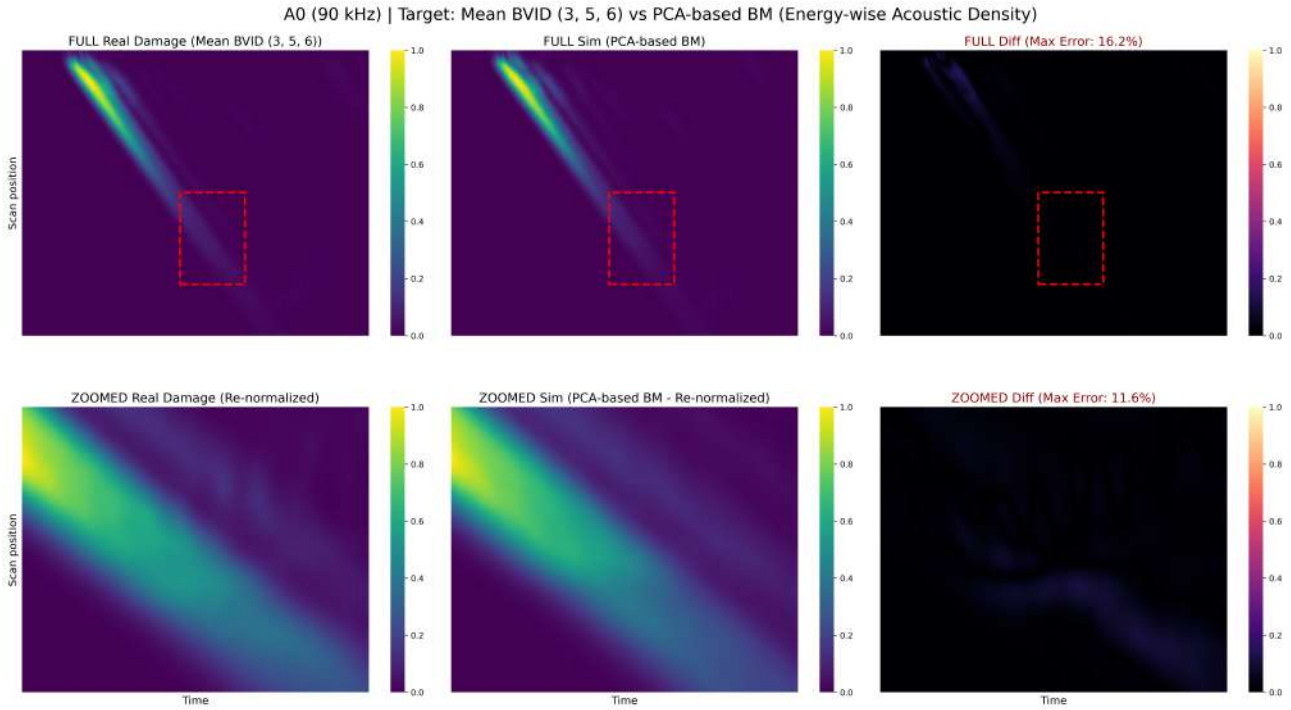


Figure 6.7: Highlight on energy-based, case A0 — Mean of the BVIDs.

Conversely, the S0 mode analysis exposes a critical limitation of the physical-based candidate. The energy map of the simulated surrogate reveals a highly localized, intense acoustic hotspot, followed by different scattering tendencies. Because the S0 mode is predominantly extensional and driven by in-plane displacements, its propagation is exquisitely sensitive to local in-plane stiffness and axial impedance mismatches. The localized hotspot indicates that the uniform PVC and wax configuration acts as a "hard" acoustic boundary: the rigid coupling completely reflects the extensional wave, causing a highly concentrated energy reflection. In stark contrast, the real BVID dissipates the S0 wavefield over a much broader spatial region. A real impact generates a complex internal network of micro-delaminations and shattered fibers—air-filled voids that act as a "soft" acoustic boundary, gradually degrading the structural stiffness and promoting diffuse multiple scattering.

This fundamental inability of the simplified surrogate to replicate the spatial energy diffusion of a real impact drives the residual error up to 48.1%.

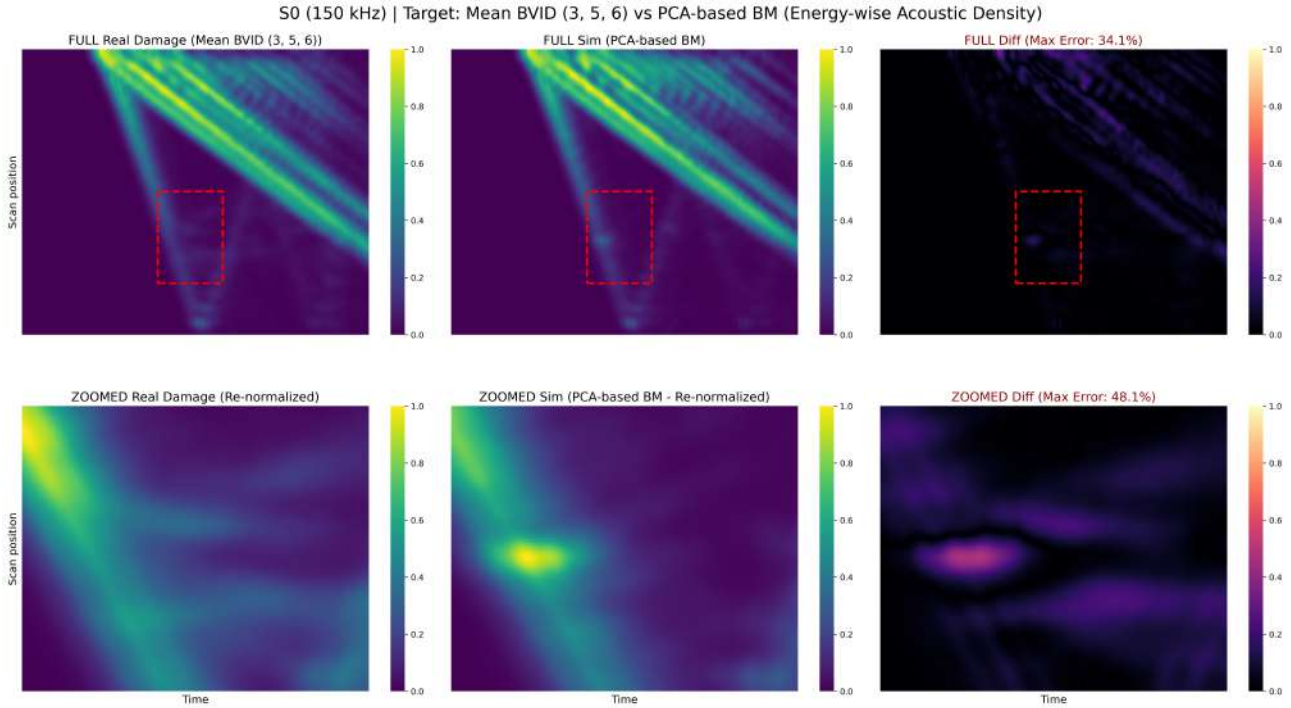


Figure 6.8: Highlight on energy-based, case S0 — Mean of the BVIDs.

these results demonstrate that while this physical-based surrogate successfully captures the inertial properties governing low-frequency macroscopic behavior, its heuristic feature selection fundamentally fails to emulate the stiffness degradation and high-frequency, multi-layered scattering of a true delamination. This profound physical limitation explicitly justifies the necessity of evaluating the best surrogate for BVIDs via a different approach.

6.2 Dimensionality Reduction via Principal Component Analysis

To visually interpret the complex 10-dimensional spatial relationship between the numerical pseudo-damage configurations and the three different experimental BVID, Principal Component Analysis (PCA) was employed.

The process begins with the raw data matrix $\underline{\underline{X}}_{PD}$ of size $n \times m$, where $n = 200$ represents the total number of samples and $m = 10$ represents the extracted acoustic features aligned.

$$\underline{\underline{X}}_{PD} = \begin{bmatrix} x^1_{A0_{R1}} & x^1_{A0_{R2}} & \cdots & x^1_{S0_T} & \cdots & xx^1_{A0_{T2}} \\ x^2_{A0_{R1}} & x^2_{A0_{R2}} & \cdots & x^2_{S0_T} & \cdots & x^2_{A0_{T2}} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x^n_{A0_{R1}} & x^n_{A0_{R2}} & \cdots & x^n_{S0_T} & \cdots & x^n_{A0_{T2}} \end{bmatrix} \tag{6.8}$$

Because these features may possess vastly different absolute magnitudes and units of measurement, calculating variance directly would erroneously bias the model towards features with larger numerical scales. To prevent this, the data matrix is standardized to form the matrix $\underline{\underline{Z_{PD}}}$. Each element is centered to a zero mean and scaled to a unit variance:

$$Z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \tag{6.9}$$

where μ_j and σ_j are the mean and standard deviation of the j-th feature, respectively.

To extract the directions of maximum variance within the standardized dataset, the Eigenvalue Decomposition (EVD) of the covariance matrix indentifies those directions. The covariance matrix $\underline{\underline{Q}}_{PD}$ of the standardized dataset is computed as $\underline{\underline{Q}}_{PD} = \frac{1}{n-1}\underline{\underline{Z}}^T_{PD}\underline{\underline{Z}}_{PD}$ [4]. This $m \times m$ symmetric matrix captures the linear correlations between all pairs of features, containing the variance of the i-th feature on its diagonal, and the cross-correlations on its off-diagonal elements.
The search for eigenvalue and eigenvector is represented by the classical equation:

$$\underline{\underline{Q}}_{PD}\underline{v}_k = \lambda_k \underline{v}_k \tag{6.10}$$

Here, λ_k are the eigenvalues, which are also representing the magnitude of the variance explained by each corresponding eigenvector $\underline{v}_k$. When sorted in descending order ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$), the eigenvectors to them associated can be also sort as directions in the feature space related to the portion of variance "inside" that direction[5].

The ultimate goal is to reduce the dataset from 10 dimensions to a visually interpretable 2D or 3D sub-space while retaining the maximum possible variance. The original standardized

[4] $\underline{\underline{Q}}_{PD} = \underline{\underline{\Sigma}}^{-1}_{PD}$ previously defined.

[5] These direction can be used as a way to reduce the dimensionality of the dataset, i.e. reducing the number of features considered. that is why these vectors are often called Principal Components.

dataset $\underline{\underline{z}}_i$ is linearly transformed into the reduced space $\underline{\underline{y}}_i$ by projecting it onto the subspace spanned by the top principal components:

$$\underline{\underline{y}}_i = \underline{\underline{V}}_{sub}^T \underline{\underline{z}}_i \tag{6.11}$$

where $\underline{\underline{V}}_{sub} = [\underline{v}_1, \underline{v}_2]$ for a 2D projection, $\underline{\underline{V}}_{sub} = [\underline{v}_1, \underline{v}_2, \underline{v}_3]$ for a 3D one.

To geometrically assess which specific acoustic features are driving the separation between the PD configurations and the real BVID, the projection of the 200 PD features inside the 2D Principal Component (PC) space is constructed, also called biplot.

The biplot simultaneously displays both the projected observations $\underline{\underline{y}}_i$ and the original variables on the same reduced subspace.

To geometrically assess the similarity between the PD configurations and the real BVID, a biplot representation is constructed. The formulation of the Biplot is mathematically rooted in the dual nature of principal component analysis, which allows for the simultaneous display of both the projected observations $\underline{\underline{y}}_i$ and the original features on the same reduced subspace. The projections of the original features in the PC space are often defined as *loadings*.

This dual representation is possible due to the algebraic properties of the principal components. By definition, the PCs are the eigenvectors of the covariance matrix of the standardized dataset, forming a new orthonormal basis. This ensures that the new axes represent mutually uncorrelated directions that sequentially maximize the data variance. Because the transformation is a linear transformation from sub-spaces, the original variables can be directly projected onto the new PC subspace.

Geometrically, the loadings represent the coordinates of the original variables in the new subspace. Mathematically, the loading l_{jk} of the j-th original variable on the k-th principal component is defined by scaling the elements of the eigenvectors $\underline{v}_k$ by the square root of the corresponding eigenvalue λ_k:

$$l_{jk} = v_{jk}\sqrt{\lambda_k} \tag{6.12}$$

where v_{jk} is the j-th element of the eigenvector $\underline{v}_k$. Since the initial dataset $\underline{\underline{z}}_i$ is standardized, the loadings perfectly coincide with the Pearson correlation coefficients [6] between the original acoustic features and the principal components.

To avoid visual clutter and focus purely on the physical phenomena driving the clustering topology, only the 4 most informative features are selected for the graphical representation. The selection criterion relies on the R^2 analysis done in Chapter 4 all over the DoE. For a 2D projection spanned by PC1 and PC2, the magnitude M_j of the j-th feature's vector is calculated as:

$$M_j = \sqrt{l_{j1}^2 + l_{j2}^2} \tag{6.13}$$

Variables with the highest magnitude strictly dominate the variance explained in that specific plane. In the Biplot, these 4 phisical features are plotted as directional vectors originating from the centroid of the PD configurations. The length of the vector visually quantifies the variable's importance, while the angle between different vectors reveals their underlying physical

[6] The Pearson correlation coefficient $r \in [-1, 1]$ quantifies the strength and direction of the linear relationship between two continuous variables, defined as $r = \frac{\mathrm{Cov}(x,y)}{\sigma_x \sigma_y}$. Its square represents the coefficient of determination, $R^2 \in [0, 1]$, already presented in Chapter 4.

correlation: highly correlated features point in similar directions, whereas orthogonal vectors identify entirely independent physical behaviors.

Let us first analyze the 200 different PD configurations and the three BVID configurations inside the 2D Principal Component Space.

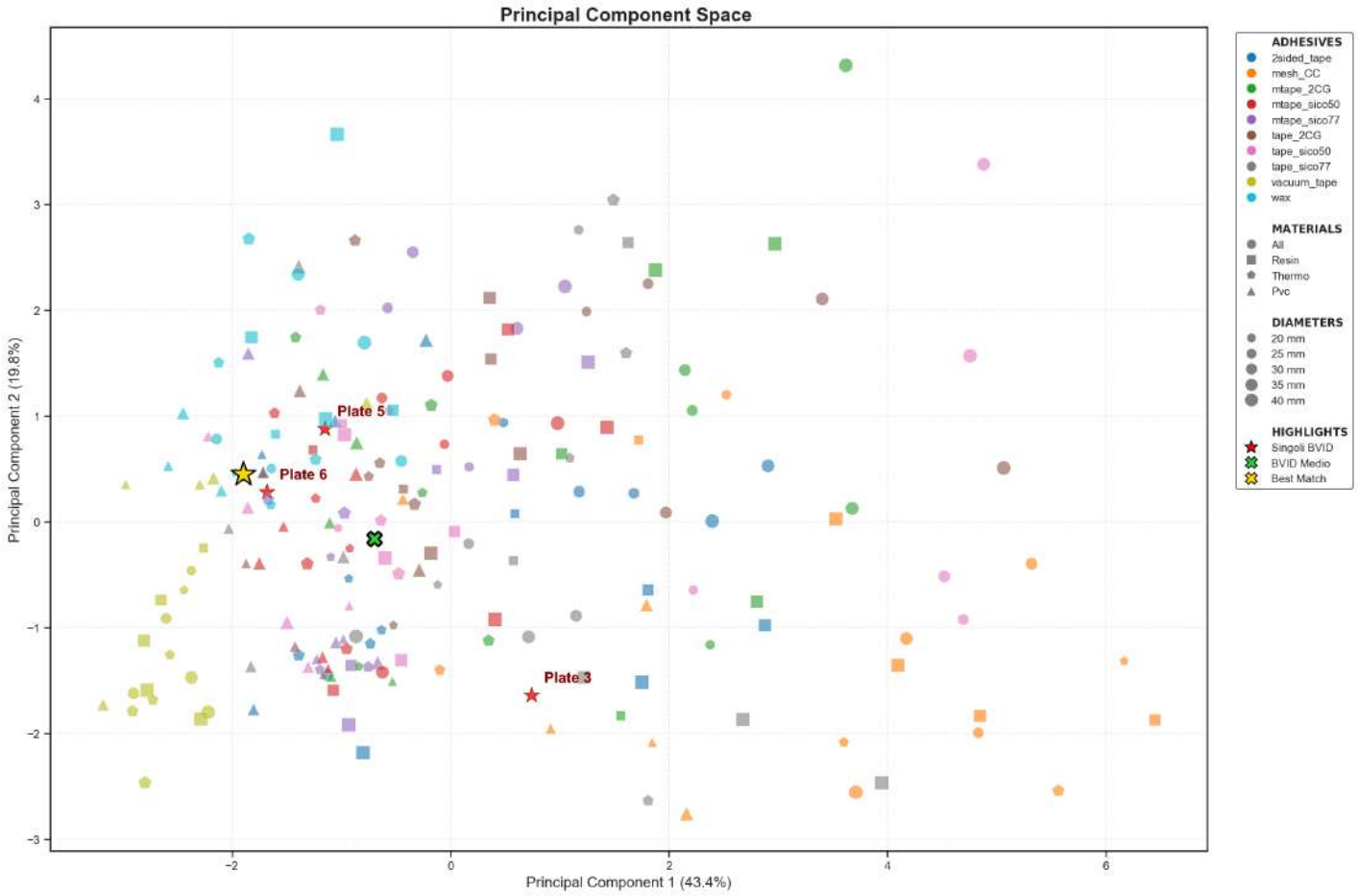


Figure 6.9

The PCA biplot reveals a high degree of dispersion and overall mixing in the dataset. Unlike distinct, isolated clusters based on material or adhesive type, the pseudo-damage configurations are distributed across a broad, continuous cloud. This widespread distribution suggests a highly coupled acoustic response, where the multiple geometric, inertial and stiffness variables interact simultaneously to shape the ultrasonic signal. Notably, the BVIDs do not form a separate region in the principal component space; instead, they are embedded within this broader PD distribution. This spatial overlap constitute itself a promising experimental result, as it geometrically demonstrates that the PD setups effectively replicate the fundamental acoustic signatures and variance of actual BVIDs.

However, a critical aspect to consider is the pronounced dispersion among the individual BVID samples. Instead of clustering tightly together to define a single, uniform "reference damage" signature, the real BVIDs are widely scattered across the latent space. Physically this sparsity could highlights the inherently chaotic and stochastic nature of real impact events from the features point of view. Even when composite plates are subjected to similar nominal

impact energies, the resulting internal damage morphology varies significantly from plate to plate. Consequently, the Lamb waves scatter, mode-convert, and attenuate uniquely for each specific impact. Therefore, the large distance between the BVID points in the PCA space is not a failure of the feature extraction, but rather an accurate representation of the high physical variability inherent to real impact damage.

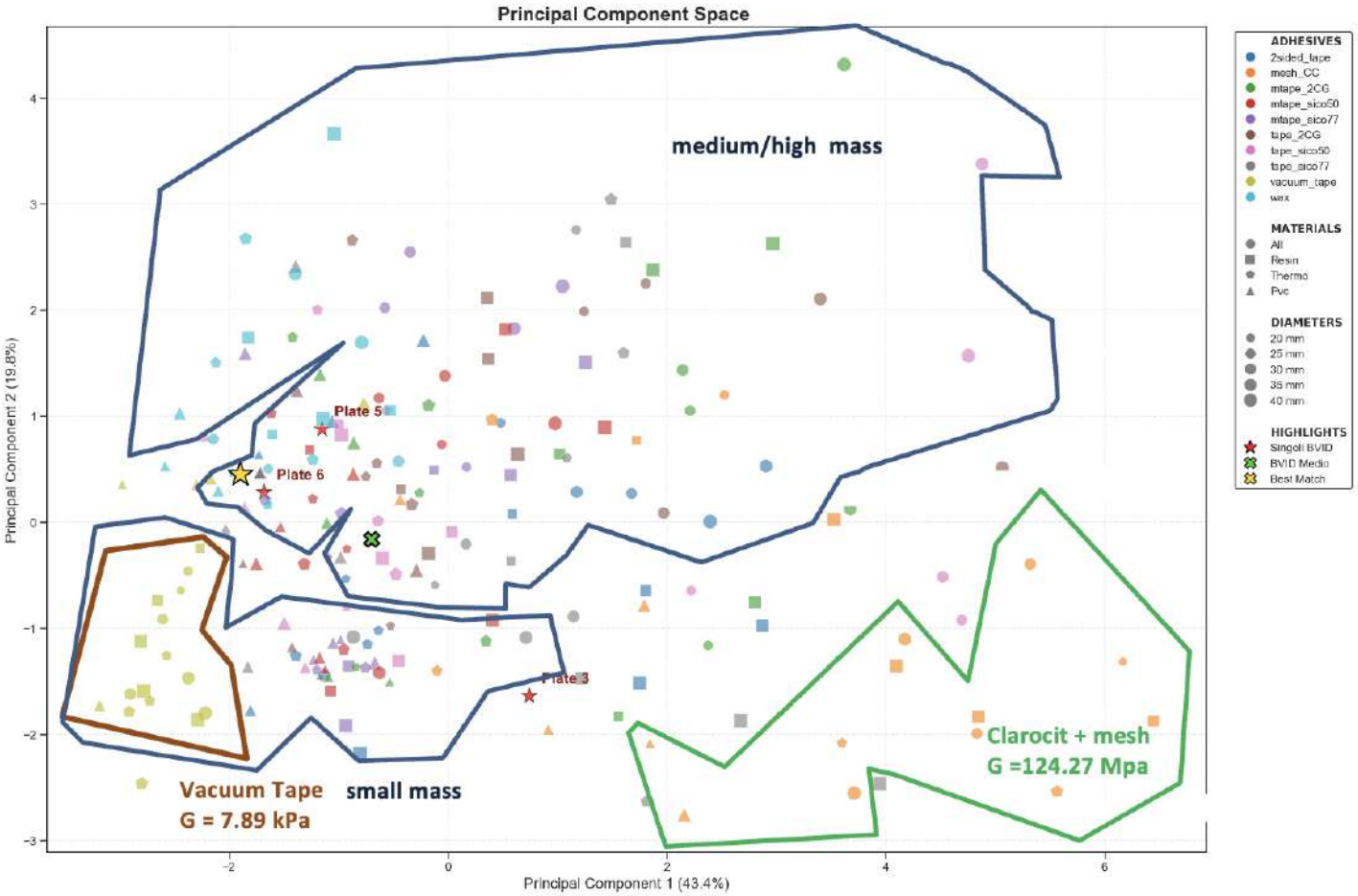


Figure 6.10: Relationship of PC1 and PC2 with the DoE independent variables.

The main advantage of the PCA is the possibility to visualize the dataset from a point of view that reduces the dimensionality while maximizing the retained information. However, this advantage comes at a cost: while the 10 defined features are based on the scattering phenomena that Lamb modes exhibit while interacting with both PD and BVID configurations, the principal components PC1 and PC2 used in the biplot have no such direct physical meaning.

Nevertheless a careful observation of the clustering topology in the reduced space allows for a phenomenological interpretation of these latent axes. Specifically, the first principal component PC1 appears to be strongly correlated with the structural stiffness of the adhesive layer. This is distinctly evident from the pronounced horizontal separation of the clusters: the configurations employing Vacuum Tape, which is the second softest adhesive tested, characterized by a very low shear modulus ($G = 7.89$ kPa), are densely clustered on the far-left side of the spectrum. Conversely, configurations utilizing Clarocit coupled with a polymeric mesh, the stiffest adhesive condition, with a shear modulus of 124.27 MPa, are on the extreme right.

The second principal component PC2 exhibits an inertial correlation: Analyzing the vertical distribution of the dataset, a macroscopic transition is visible from configurations defined in the lower region to those characterized by a medium/high mass in the upper region. From a physical perspective, this vertical dispersion captures the localized inertial resistance offered by the pseudo-damage against the propagating Lamb waves. This mass-like effect is fundamentally governed by the combination of the glued material's volumetric density and its geometric diameter, both parameters of the DoE. Consequently, while mathematically abstract, the PCA space elegantly unravels the three main drivers governing the ultrasonic scattering interaction and the PD testing campaign.

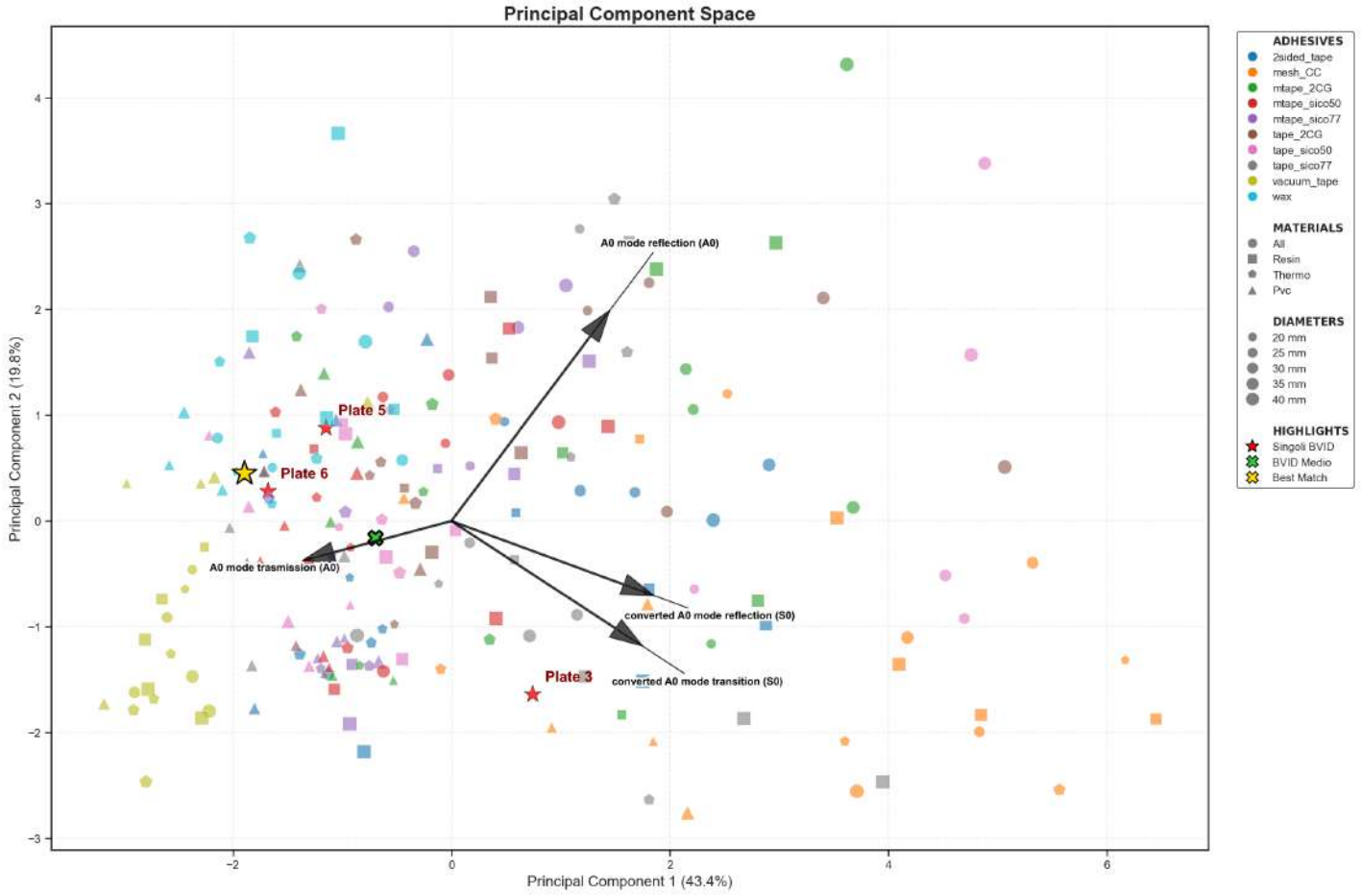


Figure 6.11: A0 and S0 most interesting features, inside the principal component space.

To further investigate the physical drivers governing this spatial distribution, the biplot in 6.11 is enriched by superimposing the loading vectors of the four dominant acoustic features: by projecting these specific features onto the PC plane, they effectively bridge the gap between the abstract principal component space and the previously discussed relationship between physical features and the variable of stiffness and inertia controlled by the DoE. The magnitude of each vector visually quantifies its variance representation within this reduced subspace, while its angular orientation unveils its underlying dependencies. A comprehensive analysis of the feature loadings within the principal component space provides crucial insights into the wave interaction mechanisms.

Regarding the antisymmetric fundamental mode (A0), the analysis focuses on the reflected and transmitted energy features. Geometrically, both the transmission and reflection vectors exhibit their most significant components along the PC2 axis. As illustrated by the vertical distribution of the clusters in Figure 6.10, this alignment confirms that the A0 mode scattering is predominantly governed by the inertial properties of the pseudo-damage configurations [7]. Furthermore, the angular relationship between these two vectors reflects the fundamental principle of energy conservation: an increase in the transmitted energy naturally corresponds to a proportional decrease in the reflected energy scattered by the defect[8].

Conversely, the features associated with the symmetric fundamental mode (S0), the energies of the mode-converted reflections and transmissions, display a distinctly different topological behavior. While their vector magnitudes are comparable to those of the primary A0 features, their spatial orientation is nearly orthogonal to them. these mode-conversion vectors are strongly aligned with PC1 and this alignment unequivocally demonstrates that the mode conversion phenomenon is primarily driven by the shear stiffness G of the adhesive layer rather than the inertial mass of the surrogate defect [9]. Interestingly, both the reflected and transmitted converted energies share a closely aligned directional vector. This positive correlation reveals a coupled generation mechanism: an increase in the reflected converted energy is accompanied by a proportional increase in the transmitted converted energy [10].

Building upon the phenomenological interpretation of the PCA space by means of the loadings, a deeper examination of the BVID sparsity reveals a highly significant insight. Spatially, the data points corresponding to Plate 5 and Plate 6 are positioned in relative proximity, whereas Plate 3 is isolated as a clear spatial outlier. This topological distribution perfectly mirrors the internal morphological observations detailed in Chapter 3. As previously confirmed by C-scan inspections, Plate 3 exhibited a distinctively asymmetric damage profile compared to the other samples. This structural asymmetry was hypothesized to be the primary driver behind its anomalous mode-conversion behavior, particularly affecting the S0-to-A0 converted energy features, as recalled in 5.2. A striking geometric evidence supporting this hypothesis emerges when analyzing the relative positioning of these plates alongside the feature loadings. The spatial vector connecting the Plate 5 and Plate 6 cluster to the isolated Plate 3 is almost perfectly parallel to the loading vector that define the direction in the principal component space that maximizes the S0-to-A0 converted energy[11]. This alignment is remarkably consistent: it provides a rigorous, data-driven validation of the morphological assumptions made during the preliminary BVID analysis.

[7] Consistent with the analytical findings in Chapter 4, the flexural nature of the A0 mode renders it highly sensitive to out-of-plane inertial blockages, exhibiting a comparatively lower dependence on interfacial stiffness.

[8] This inverse correlation strictly aligns with the standard energy partitioning behavior of an incident wave interacting with a change in the mean of traveling

[9] This geometric evidence again corroborates the deductions presented in Chapter 4, highlighting the high sensitivity of in-plane extensional waves to the shear coupling at the damaged interface.

[10] A possible hypothesis for this trend is that both converted modes simultaneously drain acoustic energy from the primary, non-converted S0 transmitted wave.

[11] Check for A0 mode reflection (S0) loading vector.

Chapter 7

Convolutional Neural Network based Feature analysis

7.1 What is Machine Learning?

> "Over the past few years, artificial intelligence (AI) has garnered significant media attention. Machine learning, deep learning, and AI have become prominent topics in numerous articles, often appearing in non-technology-focused publications. The future is envisioned as one of intelligent chatbots, self-driving cars, and virtual assistants, a future that is sometimes depicted in a bleak light and other times as utopian, where human jobs will become scarce and most economic activity will be managed by robots or AI agents." [1]

This chapter provides the essential theoretical context regarding artificial intelligence, machine learning, and deep learning. The primary aim is to understand how, and most importantly why, the specific architecture of Convolutional Neural Networks (CNNs) is employed to characterize the interaction of Lamb wave modes with pseudo-damages and barely visible impact damages. The foundational concepts and architectural frameworks discussed herein are primarily based upon the comprehensive methodologies outlined by F. Chollet in [1].
The chapter introduces the fundamental paradigms of Machine Learning and its evolution into Deep Learning. It then narrows to Convolutional Neural Networks, analyzing the ten digits classification problem and explaining the mathematical operations of convolution and pooling for spatial feature extraction. The concept of Transfer Learning is explored, highlighting how pre-trained networks can be used as powerful, automatic feature extractors. Finally, the proposed methodology is presented, detailing how these pre-trained CNN architectures are applied to extract high-dimensional latent features from ultrasonic B-Scans, justifying the transition from hand-crafted energy features to a fully automated, data-driven approach that characterizes Lamb modes across the DoE variables while validating a pseudo-damage configuration as a mean to simulate a damage.

Artificial intelligence (AI) emerged in the 1950s when pioneers in computer science questioned whether computers could "think." This question remains relevant today.
AI automates human intellectual tasks, encompassing machine learning and deep learning, but

[1] This is the incipit of [1].

also other approaches without learning. Early chess programs, for example, used hardcoded rules and didn't qualify as machine learning. Experts believed that human-level AI could be achieved by programming explicit rules for knowledge manipulation, known as symbolic AI. This approach dominated AI from the 1950s to the late 1980s, peaking during the expert systems boom. Symbolic AI solved well-defined, logical problems like chess but struggled with complex, fuzzy problems like image classification, speech recognition, and language translation. Machine learning emerged as a new approach to replace symbolic AI.

The conceptual roots of artificial intelligence can be traced back to Victorian England, specifically to the pioneering collaboration between Charles Babbage and Lady Ada Lovelace. Babbage's Analytical Engine, conceptualized in the mid-19th century, is widely recognized as the first mechanical precursor to the modern computer. However, its original scope was strictly deterministic: it was engineered to automate tedious mathematical calculations rather than to serve as a general-purpose cognitive machine. Recognizing this functional limitation, Lovelace famously noted in 1843 that the Engine possessed no inherent capability to originate anything, arguing that it could only execute tasks that humans explicitly knew how to instruct it to perform [1]. Over a century later, this strictly deterministic perspective was profoundly challenged by Alan Turing. In his seminal 1950 paper, Computing Machinery and Intelligence, Turing addressed what he formally termed "Lady Lovelace's objection". He hypothesized that computing machines could, in fact, transcend explicit human instructions to exhibit originality and learning capabilities. Turing's inquiry laid the philosophical and practical groundwork for the field of Machine Learning (ML). The fundamental engineering question shifted from whether a machine could execute complex instructions to whether it could autonomously deduce how to perform a task by observing empirical data. This conceptual leap represents a radical departure from the classical programming paradigm, historically associated with symbolic AI. As illustrated by the fundamental contrast between the two paradigms, in a traditional algorithmic approach, human engineers provide the computational system with both the input data and the explicit logical rules required to generate the desired answers. Conversely, the Machine Learning paradigm reverses this operational logic: the algorithm is provided with vast amounts of input data alongside the expected outcomes. Through iterative optimization, the machine autonomously maps the input to the output, effectively discovering the underlying mathematical rules. Once extracted, these inferred rules can be generalized and applied to novel, unseen data to predict future behaviors or classifications.

7.2 Defining ”Deep” Learning

To operationalize a machine learning algorithm, three fundamental prerequisites must be established. First, a comprehensive dataset of input features is required. Second, the corresponding expected outputs, often referred as labels, must be provided for supervised tasks. Third, a quantitative metric is strictly necessary to evaluate the algorithm's performance. This metric calculates the discrepancy between the model's predicted output and the true expected output, serving as a feedback mechanism to iteratively adjust the system's internal parameters. In the mathematical framework of machine learning, this continuous optimization process is the very definition of ”learning”. The central objective of this optimization is to discover meaningful transformations, or representations, of the input data. A representation is essentially a specific encoding of information that simplifies a given computational task.

To intuitively grasp what these learned representations entail in a deep architecture, consider the classic computer vision task of handwritten digit classification [1]. The raw input data consists of a two-dimensional matrix of pixel intensities. As this raw grid is processed through a sequence of interconnected layers, the network systematically transforms the image into internal representations that bear decreasing visual resemblance to the original input, yet become increasingly informative for the final classification. This process can be formally conceptualized as a multistage information distillation. At each sequential layer, the input data passes through mathematical filters that progressively ”purify” the signal: early layers might extract low-level spatial features such as edges or simple gradients, while deeper layers combine these into highly abstract, latent structures that directly correlate with the target classes the digits from 0 to 9.

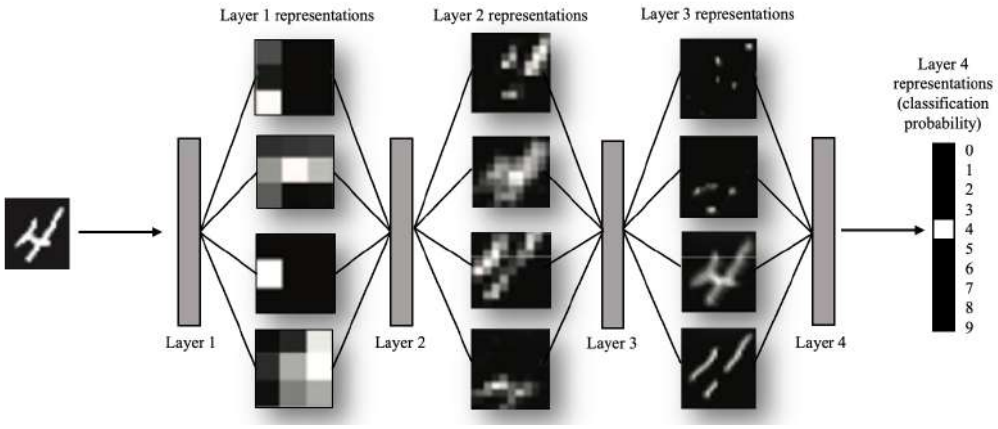


Figure 7.1: Convolutional Neural Network (CNN) as a neural network example, qualitatively.

Building upon this concept of information distillation, Deep Learning emerges as a highly specialized subfield of machine learning. It is crucial to clarify that the term ”deep” does not imply a profound cognitive understanding of the data; rather, it refers strictly to the architectural topology of the models. It denotes the strategy of cascading multiple, successive layers of representations to achieve this hierarchical abstraction. How many layers contribute to a specific model defines its depth. While traditional machine learning techniques typically rely

on extracting one or two transformations of the data, deep learning architectures construct a robust feature space through tens or even hundreds of sequentially stacked layers. Each layer refines the output of the previous one, learning increasingly complex and meaningful features automatically from the raw training data. Ultimately, deep learning is fundamentally a scalable framework: through the chained application of relatively simple parametric functions, rigorously optimized via a feedback signal, the system becomes capable of modeling extraordinarily complex, non-linear mappings.

In the specific domain of spatially structured data, such as images, these hierarchical representations are universally implemented via Convolutional Neural Networks (CNNs). Although the early conceptual inspiration for CNN architectures was drawn from the neurobiology of the human visual cortex, equating these algorithms with actual biological vision is a severe misconception. Rather than attempting to mimic biological perception, a CNN operates as a rigorous, deterministic mathematical pipeline. It is explicitly engineered to process multi-dimensional arraysby alternating two fundamental operations: convolutional layers and spatial maxpooling layers. Instead of connecting every input pixel to every single neuron, as in traditional fully connected networks, the convolutional layer employs a set of small, trainable weight matrices known as kernels or filters. During the forward pass, each kernel systematically slides across the spatial dimensions of the input matrix. At every position, an element-wise multiplication is performed between the kernel weights and the corresponding local region of the input, followed by a summation. Mathematically, the discrete convolution operation for a two-dimensional input matrix I and a kernel K yields a feature map S, expressed as:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i-m, j-n)K(m,n)$$

This localized operation ensures that the network strictly evaluates spatial relationships. By doing so, it acts as a pattern recognizer, extracting fundamental geometric primitives such as edges, gradients, or, in the context of ultrasonic testing, specific localized acoustic scattering. Following the convolution, a non-linear activation function[2] applied element-wise to the feature map. This non-linearity allows the network to model highly complex, non-linear phenomena within the data. To complement the convolutional layers, CNN architectures periodically interleave pooling layers. The primary objective of pooling is to progressively reduce the spatial dimensions of the feature maps, condensing the extracted information. The most prevalent technique utilized in modern architectures is Max Pooling. Similar to the convolutional process, a pooling window slides over the feature map. Instead of computing a weighted sum however, it simply extracts the maximum numerical value within that local window and discards the rest. From a diagnostic and engineering standpoint, this operation serves two critical purposes: first, it drastically reduces the computational burden by decreasing the number of parameters propagating through the network. Second, and more importantly, it introduces spatial translation invariance[3]. If a specific feature shifts by a few pixels due to acquisition noise or minor sensor misalignments, the max pooling operation ensures that the highest energy activation is still reliably captured and forwarded to the subsequent layers.

[2]Tipically Rectified Linear Unit (ReLu) is a common use.

[3]The biological heritage is here: as every creature is able to recognize patterns regardless their position in the field of view.

7.3 Feature Extraction via Pre-Trained CNN Architectures

Building upon the fundamental operations of discrete convolution and spatial pooling, it becomes evident how a Convolutional Neural Network constructs a hierarchical understanding of the input data. Convolutions extract localized geometric patterns, while pooling layers compress this information, granting spatial translation invariance. However, as the demand for deeper and more accurate representations grew, modern CNN architectures evolved beyond these simple sequential blocks.

State-of-the-art networks incorporate highly complex architectural innovations to solve issues related to deep optimization, such as the vanishing gradient problem. For instance, models like ResNet (Residual Networks) introduce skip connections that allow gradients to bypass certain layers, enabling the training of networks with hundreds of layers. Similarly, architectures like Inception utilize parallel convolutional filters of varying sizes within the same module, while EfficientNet employs a systematic scaling of network width, depth, and resolution.

Rather than training these massive architectures from scratch[4] this work leverages Transfer Learning. Specifically, the CNNs, pre-trained on massive generalized datasets like ImageNet, are utilized purely as automated feature extractors. By removing the final dense classification layers, the network acts as a deterministic mathematical function: it processes an input image and outputs a highly descriptive, high-dimensional numerical vector, that is the rappresentation that the CNN architecture is building upon that input image.

- Input dimensionality and channel mapping: Pre-trained models rigorously expect inputs in a specific format, typically 3-channel RGB images with fixed spatial dimensions. Conversely, the calculated B-Scans are inherently 2D single-channel arrays representing acoustic energy or wave amplitude. To satisfy the network's architectural constraints without losing physical meaning, the B-Scans must be systematically cropped using original algorithms. Furthermore, the single-channel data must be converted into a 3-channel format: this was done triplicating the grayscale matrix across the three RGB channels.

- Data normalization: To align the ultrasonic data with the statistical distribution of the original ImageNet training set, the input matrices are subjected to normalization and standard scaling, ensuring that the feature extraction process is not biased by absolute amplitude disparities between different experimental acquisitions.

- Label association: A critical operational step is ensuring robust data tracking. Each generated image must be meticulously associated with its experimental metadata, the damage type, and the PD configuration within the DoE. This precise association ensures that once the feature vectors are extracted, they are unambiguous.

Once the preprocessing pipeline is set up, every B-Scan image is fed into the chosen pre-trained CNNs. The output for each image is a dense, high-dimensional vector. Although these vectors offer an exceptionally detailed topological and spatial description of the Lamb modes wavefronts, their high dimensionality and abstract nature pose significant challenges in terms of direct visualization and interpretation. Furthermore, it is worth considering whether the objective of constructing abstract features from the B-scan images is to identify the "best match"

[4] A process that would require millions of annotated images and enormous computational resources.

that can serve as a surrogate for the BVID upon the DoE, specifically the one that possesses the identical feature vector as the mean BVID[5].

To address this, Principal Component Analysis is reintroduced into the methodological framework. Just as PCA was used in the previous chapter to project the manually extracted energy features into a lower-dimensional subspace, it is now applied to the dense CNN-extracted latent vectors. This mathematical projection compresses the thousands of deep features into a limited set of Principal Components that capture the maximum variance within the dataset. By plotting the first few principal components, it becomes possible to visually assess the clustering quality: evaluating whether the deep neural network has successfully discovered intrinsic, abstract representations that mathematically separate the 200 pseudo-damaged configurations and the 3 barely visible impact damages. Building from that visual assessment, a benchmark is established to systematically compare the feature extraction capabilities of the four pre-trained architectures selected as possible feature extractors. Since the proposed methodology relies on unsupervised representation learning rather than traditional supervised classification, standard evaluation metrics such as accuracy or cross-entropy loss are inapplicable. Instead, the performance of each network is measured by evaluating the informational density and compactness of its resulting PCA subspace.
The optimal architecture is selected among a pool of 4 different CNN architectures (VGG16, DenseNet121, ResNet50 and EfficientNetB0) as the one capable of maximizing the cumulative explained variance retained within exactly three principal components. These three components correspond to the three leading eigenvectors of the latent space's covariance matrix. By concentrating the highest percentage of the dataset's variance into three orthogonal dimensions, the winning network filters out noise and distills thousands of deep features into a compact, meaningful geometric space. Maximizing the variance captured by these top three PCs enables comprehensive 3D visualization of clusters and provides a stable foundation for identifying the best candidate among 200 pseudo-damages by calculating the minimum variance-weighted Euclidean distance[6] to real BVID targets in this reduced 3D subspace.

Table 7.1: Explained Variance Ratio captured by the first three Principal Components across the benchmarked pre-trained CNN architectures.

Principal Component	**VGG16**	**ResNet50**	**DenseNet121**	**EfficientNetB0**
PC1 Variance	16.9%	16.1%	26.0%	26.6%
PC2 Variance	6.4%	8.1%	11.0%	10.2%
PC3 Variance	3.5%	4.9%	6.4%	5.9%
3D total	**26.8%**	**29.1%**	**43.3%**	**42.7%**

[5]To assess this, one would need to compare the differences of all the thousands of rows of the two vectors, for each one of the PD configuration across the DoE.)

[6]In the principal component space, each principal component carries a percentage of the information: the weightening of the euclidean distance takes into account the percentage of information as weights.

7.4 CNN-based similarity assessment

As detailed in Table 7.1, DenseNet121 rises as the best of all the pre-trained architectures in distilling the structural complexity of ultrasonic B-Scans: it captures the highest cumulative variance within the leading principal components, effectively compressing the intricate topological features of the acoustic wavefields into an informative latent space. This feature-extraction capability is probably due to its architectural design, which employs dense connectivity patterns to ensure maximum information flow and feature reuse across its layers.For this reason, DenseNet121 was selected as the definitive feature extractor for the proposed methodology.

Having established the optimal deep learning architecture, the thousands-dimensional feature vectors extracted from the network were projected onto the reduced principal component subspace to visually and mathematically evaluate the correlation between the pseudo-damaged configurations and the real impact damages.

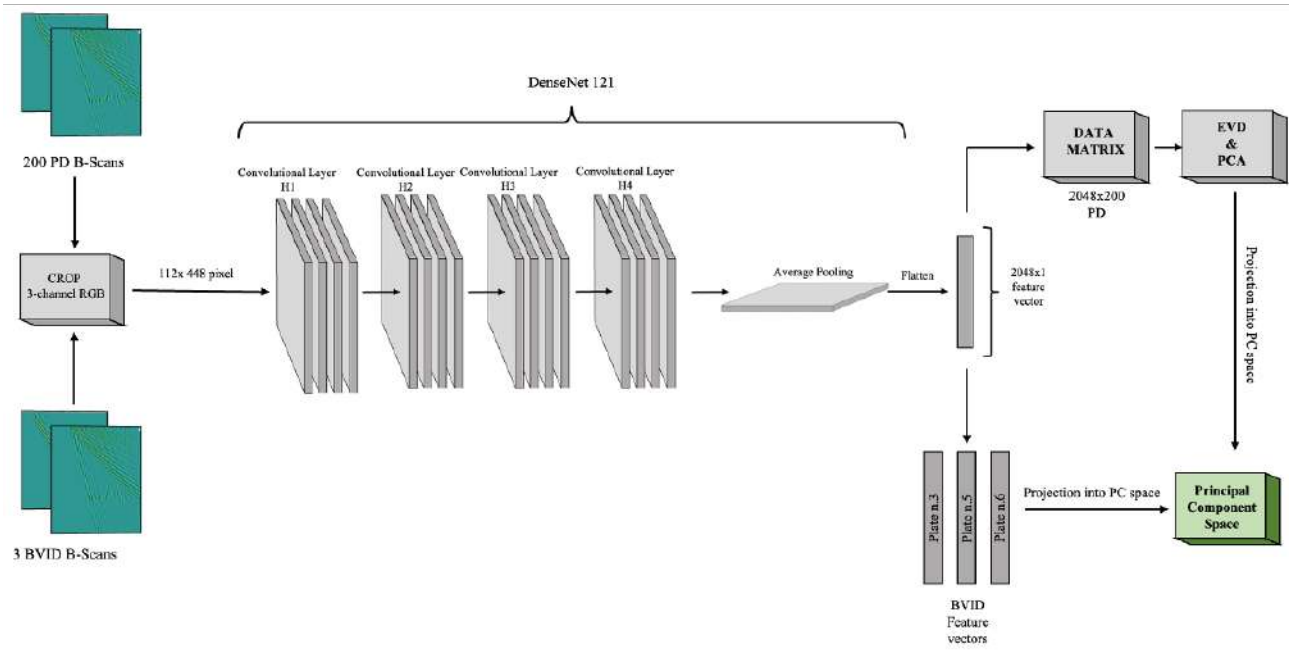


Figure 7.2: Block diagram for the CNN based similarity assessment automatic pipeline.

The resulting two-dimensional PCA space provides a comprehensive topological mapping of the Lamb modes interactions with BVID and PD configurations. Within this subspace, the spatial coordinates of each plotted point represent the fundamental representation of the corresponding ultrasonic B-scan. The 200 PDs distribute across differents regions dictated by their defining physical parameters, demonstrating the network's ability to cluster the data based on the pseudo-damaging conditions. By projecting the 3 BVIDs into this exact same metric space, their relative positioning can be objectively assessed: the geometric proximity between a target BVID point and a specific PD cluster directly translates to a measure of morphological similarity of the interaction[7]. This spatial mapping serves as the rigorous foundation for the

[7]That is granted by the CNN itself which, untrained, extract features that are representative of the input image.

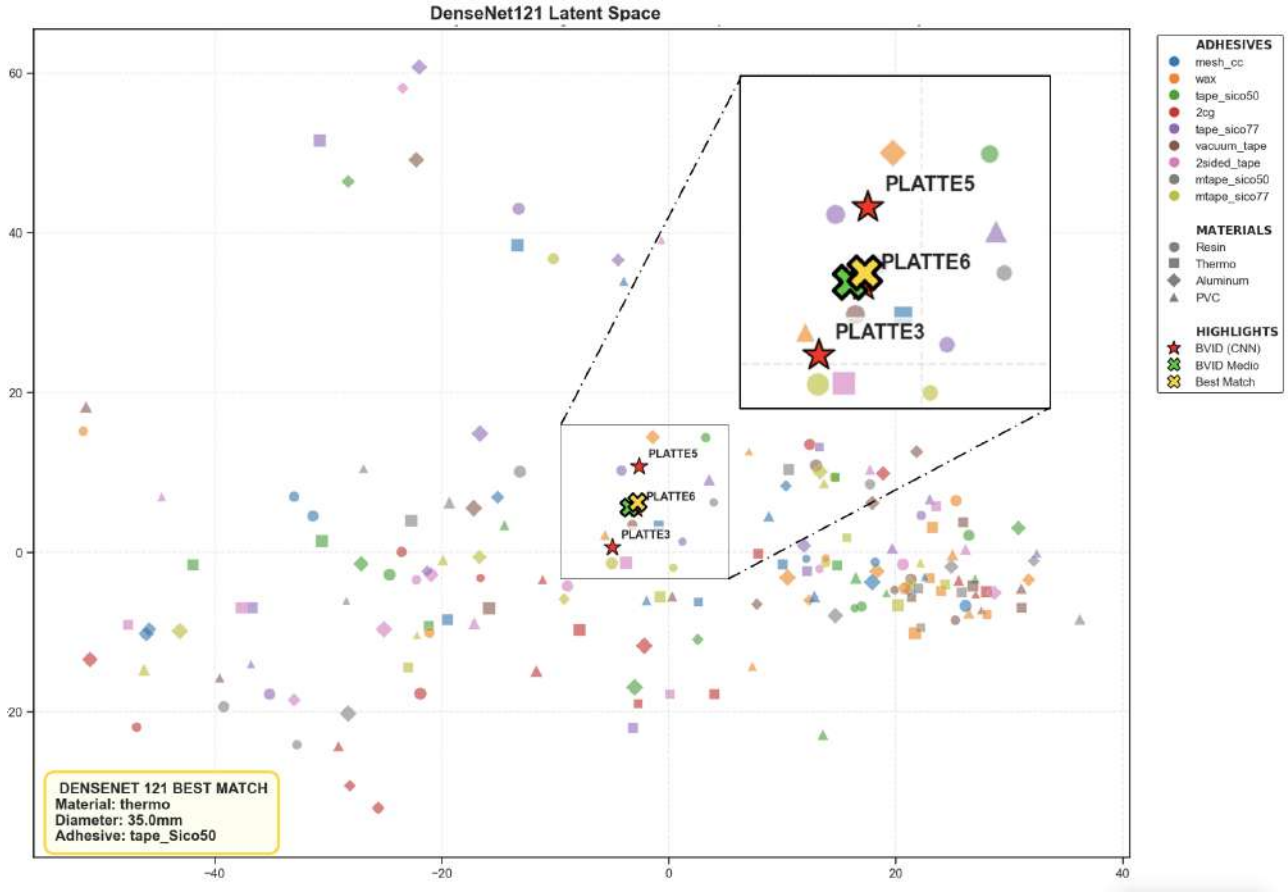


Figure 7.3: 2D CNN-based principal component space.

final diagnostic step: executing the nearest-neighbor search via the variance-weighted Euclidean distance to isolate the "CNN-based best in class".

The representation of the 2D DenseNet121 latent space, as depicted in the 7.4, reveals crucial insights. When compared to the physical-based PCA developed in the previous chapter, the CNN-generated space exhibits a significant advantage. Previously, the three real impact damages were greatly separated, risking to appair as outliers, since phisical-based energy features struggled to fully capture the multiple scattering complexities of the defects. Conversely, the deep latent representations seamlessly embed the BVID instances within the pseudo-damage clusters. This topological integration enables an exceptionally precise nearest-neighbor identification, pinpointing a specific Best Match candidate that tightly overlaps with the mean BVID centroid.

However, this mathematically superior clustering introduces a critical trade-off inherently tied to deep learning methodologies: the severe loss of explicit physical interpretability. Unlike the physical-based PCA, where the principal components were directly correlable with measurable acoustic phenomena, such as A_0 mode reflection or S_0 mode transmission energies, the CNN's latent space is constructed upon thousands of abstract, non-linear convolutional filters[8]. Consequently, the direct analytical link to the prescribed DoE is mathematically obscured. The CNN in the pipeline effectively acts as a "black box": while it excels at assessing overall

[8]Which are based upon different, general-purpose classification tasks.

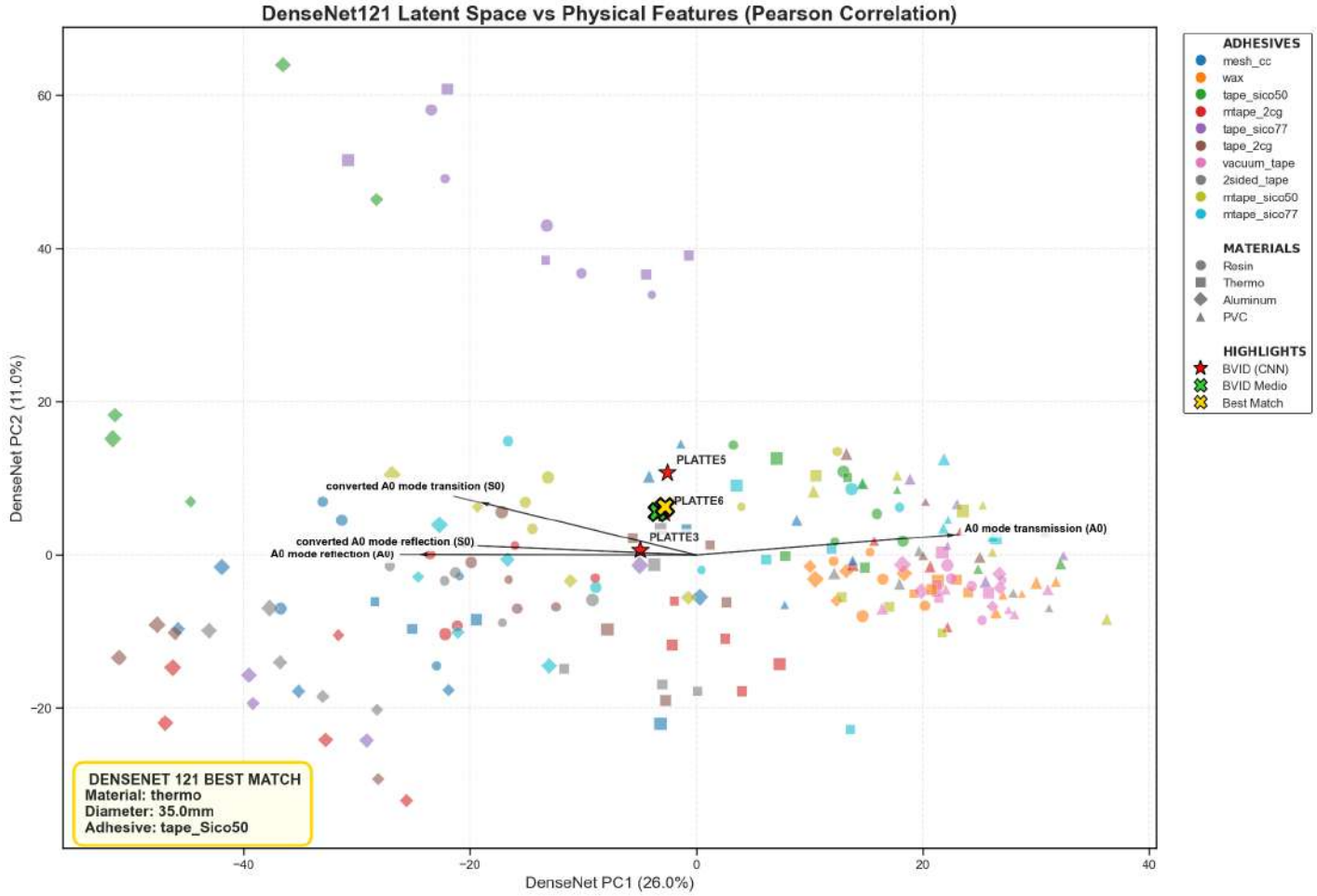


Figure 7.4: the four top features vectors inside the CNN-based principal component space.

morphological and acoustic similarity, it does not explicitly disclose which specific physical parameters, wave modes, or material properties are driving the clustering process.
To address the inherent lack of interpretability of deep convolutional filters and bridge the conceptual gap between automated feature extraction and physical wave mechanics, the manually extracted acoustic energy features were mapped onto the DenseNet121-generated PCA subspace utilizing Pearson correlation coefficients.

The resulting two-dimensional projection, represented in 7.4, unveils a highly structured and physically coherent latent representation. Several critical observations can be drawn from this representation:

- Precision of the identification: The zoomed section of the PC space illustrates the exceptional accuracy of the deep learning approach. The mathematically computed Best Match, indicated by the yellow cross, almost perfectly intersects the mean BVID centroid, the green cross, and is positioned adjacent to Plate 6. The network definitively identifies the configuration comprising a HDPE material insert, a 35.0 mm diameter, bonded with *tape_Sico50* as the optimal acoustic surrogate. This specific combination evidently replicates the highly complex spatial attenuation and scattering profile of a real composite delamination better than any other DoE candidate.

- Physical features vectors via Pearson Vectors: The superimposition of the Pearson correlation vectors unequivocally proves that the DenseNet121 architecture, despite being purely data-driven, has independently grasped the fundamental physics of ultrasonic guided waves. The primary axis of variation PC1 acts as an energy barrier descriptor. Moving towards the right side of the plot strongly correlates with high A0 mode transmission energies, whereas moving towards the left side strongly correlates with high A0 mode reflection and converted A0 mode reflection (S0).

- The Acoustic Profile of the BVID: Guided by these vectors, the exact positioning of the BVIDs and the Best Match on the negative side of PC1 acquires profound physical meaning. The deep network clustered two of the three BVIDs with the same PC1 coordinate, while the outlier one is slightly moved along th converted S0 to Ao loading vector: this is not a surprise, as it has been already proved via C-scans that the plate 3 BVID has a different conversion pattern. What is interesting here is that the CNN is looking just at the conversion pattern, without knowing nothing about it and still is able to understand the differences of the three of them.

- Differences surrogate-mean BVID: as already noted, in this PC space the differences in terms of best surrogate to mean BVID look smaller[9]: noting the direction of the loading, the main differences are in terms of the energy transmitted in the A0 interaction .

This representation validates the entire methodological framework. It demonstrates that advanced CNNs distill latent representations that are strictly isomorphic to the underlying acoustic dynamics. By fusing the DenseNet121 automatic extraction with Pearson correlation vectors, the proposed approach successfully achieves both another, non-physical, similarity assessment while granting also a robust physical interpretability.

The transition to the energy-wise metric introduced in Chapter 6 as a way to assess the similarity between the BVIDs, their mean and the identified "best in class" can be reintroduced here. Evaluating the CNN-based "best in class" under the same energy-wise metric reveals a profound trade-off, highlighting the network's prioritization of complex scattering morphologies. A significant achievement of the CNN-based candidate is observed in the S0 mode. The deep learning architecture successfully resolved the critical flaw of the physical-based model: the "acoustic barrier" effect. As depicted in the energy heatmaps, the deep-learned surrogate generates a highly diffused and spatially distributed energy profile that closely mimics the multiple scattering signature of the actual micro-delaminations. By accurately replicating this complex spatial diffusion, the localized residual error drops drastically from 48.1% for the previous candidate to an impressive 26.4%.

This superior high-frequency performance, however, comes at the explicit cost of low-frequency fidelity. When evaluating the A0 mode, the maximum residual error for the CNN candidate increases significantly to 43.4%. This represents a substantial performance degradation compared to the Physical-Based Surrogate, which achieved an exceptional minimum error of only 11.6% under the exact same energy-wise evaluation.

[9]This fact itself does not imply anything, because of the abstract nature of the projection of the CNN features onto its PC space.

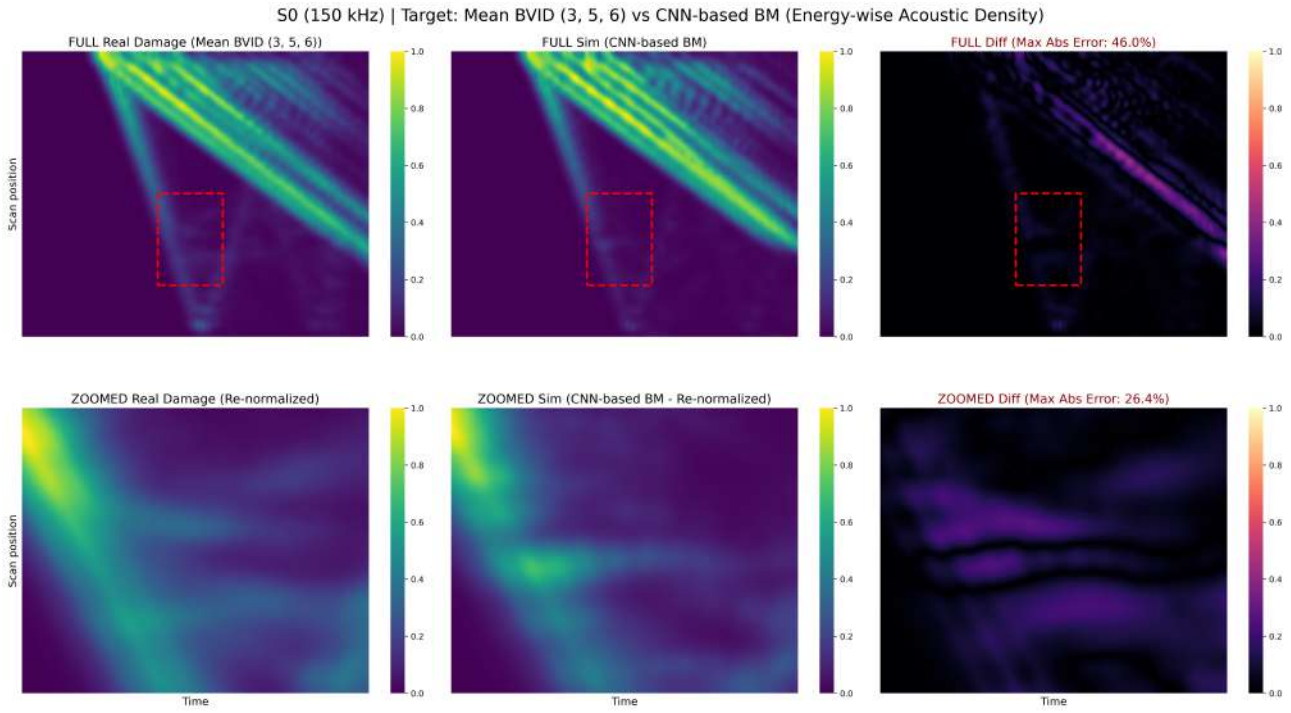


Figure 7.5: energy-based similarity assessment, case S0 - mean of the BVIDs.

The physical justification for this numerical disparity lies in the fundamental scattering mechanics of the flexural Lamb wave. At 90 kHz, the A0 mode is characterized by a relatively long wavelength, making the propagating wavefield primarily sensitive to global, macroscopic reductions in bulk stiffness and mass. The physical-based surrogate, the PVC insert bonded with wax, perfectly embodied this specific acoustic behavior: it acted as a uniform, homogeneous obstacle that successfully produced a smooth, continuous transmission shadow, closely matching the global energy attenuation of the real BVID. Conversely, the Deep-Learned Surrogate introduces an excessive degree of acoustic complexity at this scale. Visual inspection of the energy heatmaps reveals that the network's chosen configuration induces severe, localized wavefront fragmentation and irregular energy dispersion. While this highly fragmented scattering profile was the exact key to successfully replicating the high-frequency multiple scattering of the S0 mode, it proves detrimental in the low-frequency regime. Essentially, the CNN surrogate "over-complicates" the wave-defect interaction. By generating spurious local reflections and secondary trailing waves, it struggles to match the smoother, continuous macroscopic attenuation envelope exhibited by the mean BVID, directly resulting in the elevated 43.4% residual penalty.

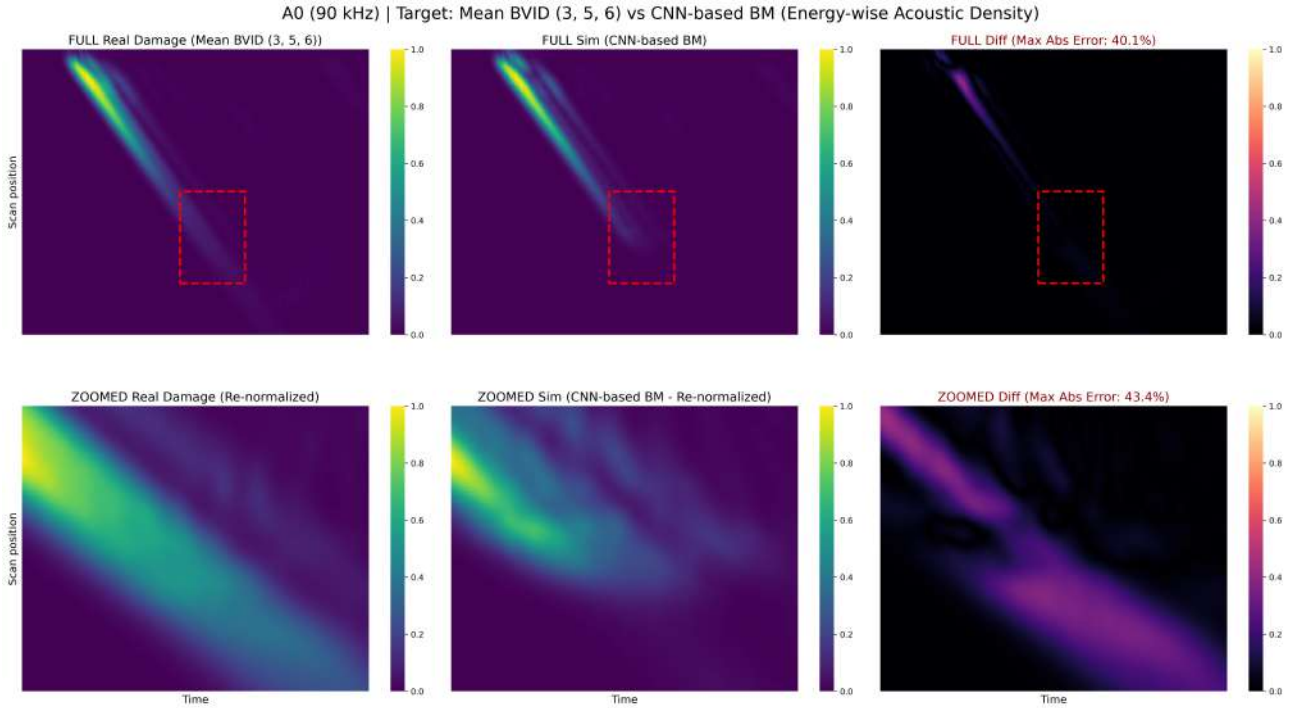


Figure 7.6: energy-based similarity assessment, case A0 - mean of the BVIDs.

Ultimately, this comparative analysis demonstrates that the DenseNet121 latent space prioritizes the emulation of complex, high-frequency internal scattering over simple, low-frequency bulk attenuation. While the physical-based surrogate remains optimal for macroscopic wave blockage, the deep-learned surrogate proves vastly superior in replicating the authentic morphological signature of a shattered composite interface.

Chapter 8

Conclusions & future perspectives

Non-Destructive Testing (NDT) and Structural Health Monitoring (SHM) techniques are of primary relevance within a damage tolerance design approach for airframe structural components. Data-driven SHM algorithms, however, require massive datasets of wave-defect interactions, which are inherently expensive and destructive to acquire. This work presented a preliminary but comprehensive framework in which the concept of simulating impact damages was rigorously explored. If proven reliably accurate, this concept could serve as a highly cost-effective and representative strategy for developing, training, and validating future SHM algorithms.

To achieve this, the core methodology relied on a DoE-based experimental campaign coupled with a physics-based feature definition and extraction, which culminated in a quadratic regression analysis. Crucially, the systematic categorization of the surrogate configurations within the DoE was made possible through a preliminary experimental evaluation of the shear stiffness values associated with the employed materials. This comprehensive framework was pivotal not only to systematically generate the surrogate Pseudo-Damages, but also to qualitatively identify the physical dependencies of the fundamental Lamb wave modes on the DoE independent variables. The regression models successfully proved representative of the defined physical features and highlighted that the flexural A0 mode is primarily sensitive to the inertial properties of the PD, whereas the extensional S0 mode is highly dependent on localized stiffness variations. Furthermore, this regression analysis established the theoretical foundation required to interpret the physical feature distributions within the two constructed Principal Component spaces: the physics-based feature space and the CNN-based latent space. Ultimately, this robust characterization provided the essential physical baseline necessary to evaluate how effectively these reversible tools can mimic genuine impact events while also providing information for further experiments.

Following the wavefield characterization, the study focused on assessing the physical similarity between the actual impact damages and the engineered PDs by first perform impact testing and then by deploying the two defined feature extraction strategies. The similarity assessment strongly relied upon the collected B-scans and a critical finding of this research was the inadequacy of raw pixel-wise difference metrics, which severely penalized the surrogates due to microscopic, unpreventable phase shifts. By transitioning to a robust energy-wise acoustic density evaluation, the genuine physical behaviors were decoupled from mathematical artifacts.

The similarity assessment revealed a fundamental diagnostic trade-off when comparing the physical-based surrogate, optimized via Mahalanobis distance, and the CNN-based Surrogate extracted from the DenseNet121 latent space. The physics-based candidate proved optimal in the low-frequency regime of the A0 mode, successfully replicating the macroscopic wave blockage and global energy attenuation of the mean BVID. Conversely, the CNN-based candidate proved vastly superior in the high-frequency regime of the S0 mode. By prioritizing complex morphological features, the deep-learned surrogate successfully reproduced the intricate, spatially diffused multiple scattering signature typical of a shattered composite interface, effectively resolving the "acoustic barrier" limitation of the physical model. Furthermore, exploring the principal component space provided profound insights into the topological clustering between the engineered PDs and the experimental BVIDs, a grouping behavior clearly evident in both evaluated cases. Crucially, the implementation of supplementary variable projection served to definitively bridge the gap between the interpretable physical-based feature analysis and the abstract, high-dimensional CNN representations. This projection technique successfully reconfirmed the underlying physical links initially identified in the physical-based PCA, demonstrating that the convolutional latent space implicitly encodes the fundamental elastodynamic dependencies necessary to accurately characterize and correlate the Pseudo-Damages with the real Impact Damages.

While the proposed framework demonstrated the high potential of engineered pseudo-damages, the investigation highlighted two primary challenges that heavily complicate the task of finding a universally feasible, cost-effective, and quasi-reversible BVID simulation strategy. The first major weakness stems from the highly diverse and stochastic morphology of real impact damages: as explicitly demonstrated through the complementary C-scan analyses, the internal network of delaminations and matrix cracks is highly irregular and fundamentally unique to each impact. This profound structural variability directly accounts for the significant fluctuations observed in the physical features extracted across the different BVIDs. The second critical challenge lies in the inherent elastodynamic complexity of the wave-defect interactions. Specifically,while the B-scans provides visual informations about Lamb modes propagation, there is a relative difficulty in fully comprehending and isolating the intricate refraction and multiple scattering patterns of the Lamb wave modes as they propagate through both the stochastic BVIDs and the engineered boundary conditions of the PDs. Consequently, it is imperative to conduct further tests and expand the prescribed experimental campaign. Increasing the statistical pool of BVIDs, combined with advanced wavefield investigations, will provide deeper insights into these complex refraction phenomena and the inherent variance of Lamb mode interactions, ultimately establishing a more robust "ground truth" for future similarity assessments.

Furthermore, the intersection of these acoustic phenomena with advanced Machine Learning techniques offers highly promising future perspectives:

- Explainable AI (XAI) and Grad-CAM Integration: Future works should thoroughly analyze the response of DenseNet121 and similar Convolutional Neural Networks to both PD and BVID B-scans by employing Gradient-weighted Class Activation Mapping (Grad-CAM). Visualizing the network's activation maps will explicitly reveal which spatial and temporal topological features drive the similarity metrics within the latent space. By extracting these heatmaps, researchers can transition from a "black-box" similarity assess-

ment to a physically interpretable framework. This visualization would serve to validate whether the CNN successfully focuses on genuine elastodynamic phenomena such as the transmission shadow, primary reflections, or mode conversion trails, rather than being biased by experimental artifacts or boundary reflections. Moreover, directly comparing the activation maps of a real BVID with its Best Match PD would precisely highlight micro-structural discrepancies, revealing exactly where the physical surrogate fails to emulate reality. Ultimately, understanding the topological features prioritized by the deep learning model would create a direct feedback loop, guiding a new iteration of the DoE to physically tailor the geometric and material properties of future pseudo-damages.

- Classical Machine Learning Deployments and Feature-Based Baselines: Alongside the deployment of deep learning architectures, implementing established classical Machine Learning techniques, such as Random Forest and Support Vector Machines (SVM), represents a critical future objective. Because these algorithms rely on the explicitly defined, physics-based feature, rather than abstract latent representations, they offer a highly interpretable analytical baseline. For instance, algorithms like Random Forest intrinsically provide quantitative feature importance metrics, which would explicitly rank which physical parameters are statistically most decisive in defining the similarity or discrepancy between a Pseudo-Damage and a genuine BVID. Concurrently, Support Vector Machines excel at defining clear, margin-based decision boundaries within a multi-dimensional feature space. Applying SVMs to the previously established Principal Component space would allow researchers to mathematically quantify the exact regions of overlap, mapping definitive similarity zones between the engineered surrogates and stochastic impact damages. Establishing this robust classical ML baseline would rigorously justify the computational expense of deep learning, definitively answering whether complex morphological feature extraction is strictly necessary or if a well-crafted, physically interpretable feature set is sufficient for surrogate validation.

- Cross-Domain Classification Tasks: Finally, if the convolutional architecture is proven robust and reliable through XAI validation, a definitive way to assess the ultimate similarity between PDs and BVIDs is to frame the problem as a generalized, cross-domain classification task. A network could be trained extensively over a significantly larger, parametrically controlled PD dataset, learning the complete spectrum of engineered wave-defect interactions. Subsequently, this trained model would be asked to classify entirely unseen BVID B-scans. Achieving a high classification accuracy in this challenging cross-domain scenario would be the ultimate diagnostic validation. It would definitively prove that the engineered surrogate defects possess the full physical and morphological representativeness required to successfully replace real, destructive impact events in the training pipelines of future SHM algorithms.

Bibliography

[1] François Chollet. *Deep Learning with Python*. 2nd. Shelter Island, NY: Manning Publications, 2021.

[2] Armin Huber. *Dispersion Calculator*. Deutch Zentrum fur Luft und Raundfarth (DLR e.V.) (2018-2025), 2018.

[3] Wierach Peter Tim Behrens Moix-Bonet Maria Schmidt Daniel. "Charakterisierung von Lamb-Wellen-Schadensinteraktionen für die Entwicklung von Pseudo-Schäden". In: *e-Journal of Nondestructive Testing* (2023).

[4] Peter Wierach. *Lamb-wave based structural health monitoring in polymer composites*. Springer, 2018. ISBN: 978-3-319-49714-3. DOI: 10.1007/978-3-319-49715-0.

[5] Ye Lu Zhongqing Su Lin Ye. "Guided Lamb waves for identification of damage in composite structures: A review". In: *Journal of Sound and Vibration 295* (2006).