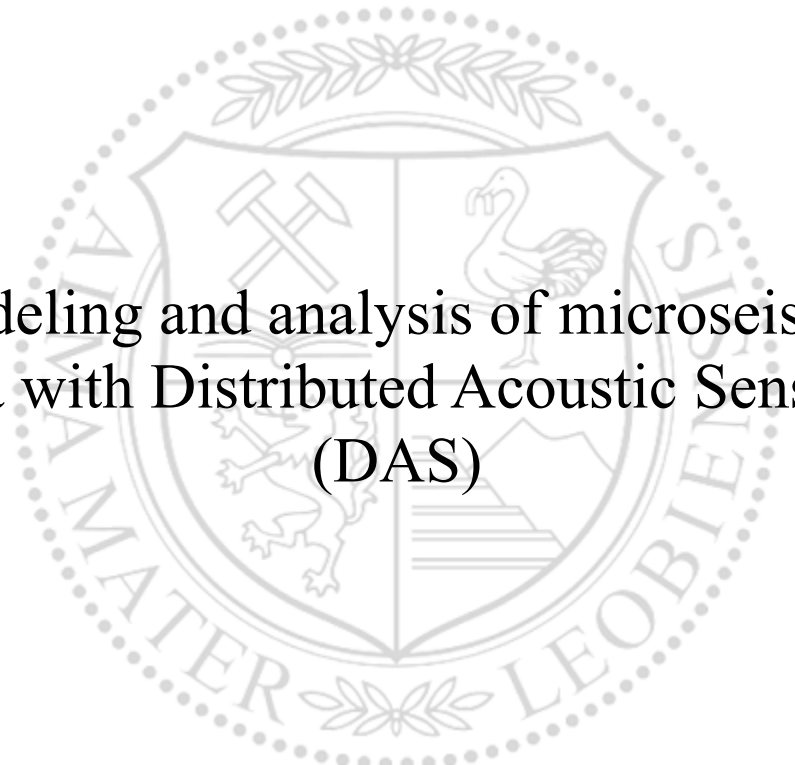




Chair of Applied Geophysics

Master's Thesis



Modeling and analysis of microseismic
data with Distributed Acoustic Sensing
(DAS)

Giacomo Rapagnani

October 2023



AFFIDAVIT

I declare on oath that I wrote this thesis independently, did not use other than the specified sources and aids, and did not otherwise use any unauthorized aids.

I declare that I have read, understood, and complied with the guidelines of the senate of the Montanuniversität Leoben for "Good Scientific Practice".

Furthermore, I declare that the electronic and printed version of the submitted thesis are identical, both, formally and with regard to content.

Date 11.10.2023

Signature Author
Giacomo Rapagnani

Contents

1	Introduction	2
2	Physical Principles of DAS Systems	4
3	Traveltime Computation for DAS	13
4	Modeling of DAS Microseismic Data	29
5	Conclusion	39

1. Introduction

Distributed Acoustic Sensing (DAS) is an emergent data acquisition technology allowing to measure the dynamic strain along the axis of an optical fiber. A DAS system consists of an optical fiber and an optoelectronic device, known as an interrogator unit (IU), that emits laser pulses within an optical fiber that can be either a classical telecommunication fiber or an engineered one. The IU is able to detect subtle phase shifts present in the backscattered light along the fiber and convert it into strain (or strain rate). In this way, it is possible to record the dynamic strain caused by an external source (e.g. seismic waves) and affecting the fiber. For this reason, this technology is becoming popular in seismology since it turns the optical fiber into a seismic array allowing for high-density and large aperture experiments. In the last few years, the interest of the seismological community in DAS technology has grown exponentially, especially for microseismic monitoring operations in borehole installations, where its high spatial and temporal sampling, if compared to standard seismological acquisition technologies, provides more detailed information on the seismic wavefield. Furthermore, DAS is particularly useful for data acquisition in logistic challenging scenarios (e.g. offshore areas, borehole installations, glaciers, or volcanic environments) where the deployment of conventional seismometers can be difficult. It's thus clear that, due to their high spatial and temporal sampling, Distributed Acoustic Sensing (DAS) systems can generate massive amounts of data, especially for long data acquisition periods. For example, a sample of 1 day of data collected with a 1 km long fiber with inter-channel distances of about 1m and temporal sampling of 0.5 ms can easily reach 2 TB. This highlights the need for specific data analysis procedures that, at the same time, are fast enough and capable of exploiting the massive amount of information contained in such data. In the last decade, DAS has been increasingly used in microseismic monitoring for conventional and unconventional hydrocarbon production, geothermal energy exploitation, and CO₂ sequestration and it is also applied in near-surface geophysics for engineering and environmental studies, particularly those requiring long-term monitoring (Lindsey, E. R. Martin, et al. 2017,). Though strain measurements have been conducted

since the early 1900s (Benioff 1935), the development of distributed optical fiber strain sensing began in the 1970s, and they were mainly used for point measurements (Bucaro, Dardy, and Carome 1977, Cole, Johnson, and Bhuta 1977). In the 1980s the use of an optical fiber light path for measurements of long-baseline strain greatly simplified the construction of interferometers (M. Zumberge, F. Wyatt, and Dong 1988). Today, DAS achieves remarkable sensitivity even compared with geophones (Piana Agostinetti, Villa, and Saccorotti 2022), enabling the detection of tidal strain, tremors, and slow slip earthquakes (Blum, Nooner, and M. A. Zumberge 2008, DeWolf et al. 2015, Hatfield, F. K. Wyatt, and M. A. Zumberge 2018). Despite the popularity of DAS systems, there is a lack of standard modeling and analysis tools that can be used within routine procedures.

In this thesis, I have developed a versatile workflow for synthetic DAS data generation based on the convolutional model. For this purpose, I have developed a travel-time calculator that solves the Eikonal equation, which is capable of managing various data acquisition geometries, including scenarios where an optical fiber is deployed in deep boreholes, either vertical or deviated ones. Synthetic DAS seismograms are then generated by using the computed travel-times both for P and S phases and the convolutional model. Although DAS synthetics calculated with the convolutional model are less realistic than the ones calculated with other methods, such as the reflectivity or the spectral element method, their computation is much faster and efficient. This aspect is particularly important in the generation of large DAS synthetic datasets. Our synthetic generation workflow can be used to 1) test newly developed seismic event detection and location methods for DAS data and 2) train machine learning models. The first part of the thesis covers the basic physical and engineering concepts about the instrument, that are needed to understand how Distributed Acoustic Sensing works. The second part of the thesis focuses on the description of the synthetic DAS data generator workflow I developed. I first describe the implementation of the Eikonal solver for travel-time calculation, followed by a detailed description of the synthetic DAS data generation process, based on the convolutional model. In the last part of this thesis, I show a comparison of the synthetics obtained using our workflow with the ones obtained using the spectral element method.

2. Physical Principles of DAS Systems

A Distributed Acoustic Sensing (DAS) system consists of two parts: the Interrogator Unit (IU) and the fiber-optic cable, illustrated in Figure (2.1). The IU serves a dual role, as it not only generates the light pulse transmitted through the fiber but also records the back-scattered signal. In this sense the IU works both a source and receiver, simultaneously sending the laser pulse and receiving the backscattered one. Meanwhile, the fiber-optic cable transmits the pulse with minimal signal attenuation. This unique configuration allows the deployment of the fiber-optic even in complex geometries and takes up minimal space.

The number of channels and the spatial sampling can be changed by the user depending on the target and the computational cost, the dx can be as small as $\sim 10cm$ or extend for $\sim 10m$ (or more). The Gauge length is the space interval over which the strain is integrated and could be referred to as a moving average filter along the fiber (Hartog 2017) and can be selected accordingly.

There are a multitude of IU and fibers but the governing principle it's the same: a laser pulse is sent in the fiber and due to small imperfections inside the core part of the pulse is back-scattered and recorded by the IU. The interesting fact about fibers used for telecommunication purposes is that back-scattering is considered noise because it's caused by absorption. For this reason, the manufacturers try to construct the fiber with as less imperfections as possible. On the contrary, for DAS systems the heterogeneity of the fiber it's necessary in order to measure the strain-rate. Finally, the scattered signal is compared to a reference one, for example, in the Double-Pulse method the IU sends two pulses with a fixed time delay and measures the phase difference. This measure is then used to calculate the strain $\varepsilon = \frac{\Delta L}{L}$ (or strain-rate $\dot{\varepsilon} = \frac{d\varepsilon}{dt}$) by exploiting the correlation between the phase shift and the elongation of the fiber itself. Within a DAS system for a given gauge length L_G the strain is calculated as (Dean et al. 2016):

$$\varepsilon_{DAS} = \int_{x-\frac{L_G}{2}}^{x+\frac{L_G}{2}} \frac{\varepsilon}{L_G} dx' \quad (2.1)$$

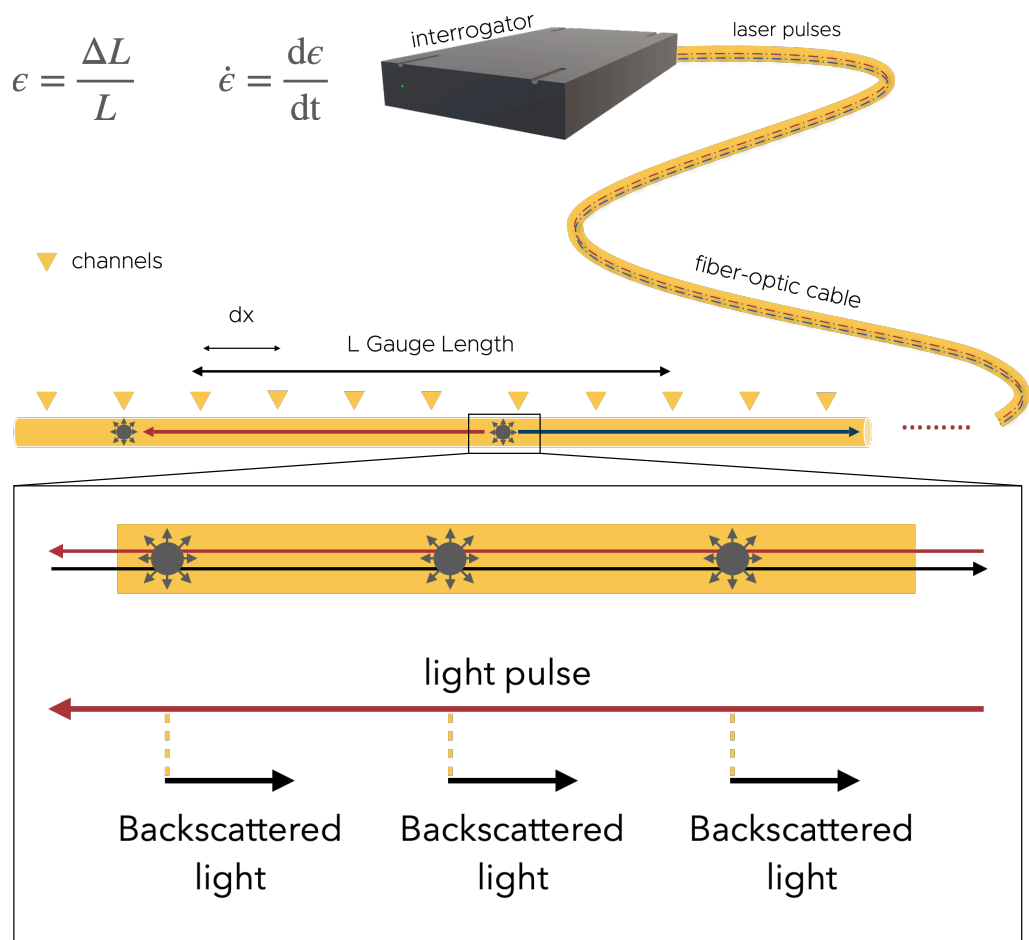


Figure 2.1: Scheme of DAS System

In this chapter I will describe the physical principles behind DAS, starting from the characteristics of the optical fibers, and continuing with the optical phenomena occurring when laser pulses propagate within the fibers. Finally, in order to understand how the final strain-rate measure is calculated, I will focus on one of the most used and common DAS systems, which uses the phase-sensitive OTDR (Optical Time Domain Reflectometry) that utilizes the Rayleigh scattering to compute the strain rate.

Fiber-optic is a dielectric material, usually made of silica or plastic, that acts as a waveguide in which light can travel with very low coherence loss and absorption (Senior and Jamro 2009).

A fiber is generally composed of three concentric layers, namely, starting from the inner one: the Core, the Cladding, and the Buffer as shown in Figure (2.2). The Buffer

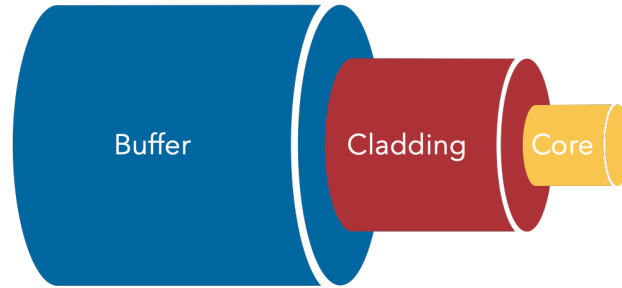


Figure 2.2: Internal Fiber Structure showing the three components: Core, Cladding, and Buffer.

acts as a protective layer insulating from external electromagnetic waves (in some fibers is not present), the Core and the Cladding layers are responsible of the transmission of the signal (Soga and Luo 2018). The core has a higher refractive index with respect to the cladding:

$$n_1 > n_2 \quad (2.2)$$

and so the ray is subjected to a critical reflection for an angle of incidence equal to or greater than θ_c which is called critical angle:

$$\theta_c = \sin^{-1} \left(\frac{n_2}{n_1} \right) \quad (2.3)$$

where n_1 is the refractive index of the Core and n_2 of the Cladding. Under this condition there is virtually no light transmission outside the Core and the signal doesn't lose energy (Figure (2.3)). In reality, there are many factors that contribute to the absorption inside the fiber (such as Rayleigh scattering), still, in a DAS System the signal can travel for tens, or in some cases even hundreds, of kilometers.

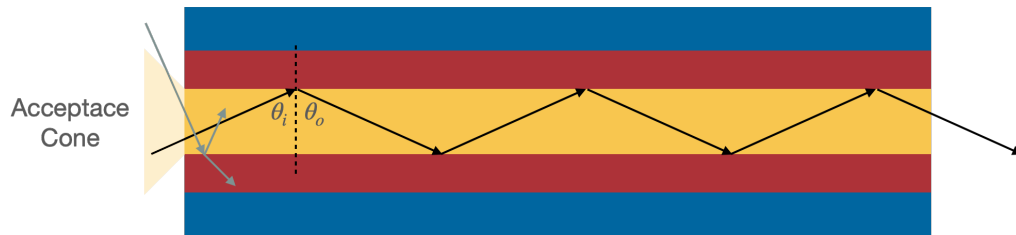


Figure 2.3: Ray Transmission in the fiber

Another characteristic that allows the fiber to be deployed for long distances is the absence of current flow, due to the fact that it is a dielectric material, and so no voltage instability problems are present.

A Fiber behavior that instead mines the quality of the phase difference measure is the dispersion of the wave pulse. Dispersion in the fiber is present in different forms, namely: material dispersion, multi-mode dispersion, polarization dispersion, and waveguide dispersion (Okamoto 2021). When a dispersion phenomenon occurs the pulse's group velocity is different from the phase velocities, causing different frequencies to travel at different speeds. The material dispersion is due to the fact that the refractive index in the fiber changes in accordance with the wavelength. Instead, the multi-mode dispersion is an arrival time difference between frequencies caused by the different reflection angles. Depending on the angle each frequency will travel more or less distance in the fiber resulting in a broadening of the pulse (Figure (2.4) shown graphically this behavior in the multi-mode fiber). The polarization dispersion is the difference in velocity of the orthogonally polarized waves caused by the non-cylindrical symmetry of the fiber-optic cable and lastly, the waveguide dispersion is an intrinsic dispersion due to the light confinement in a waveguide. It's important to point out that for DAS application a single-mode fiber is generally used, which is susceptible to chromatic dispersion only (sum of material and waveguide dispersion), and by selecting the right parameter, such as the core diameter and the fiber-optic refractive indices, a zero-dispersion wavelength can be used. So in the phase difference measure a non-dispersive pulse is considered.

As mentioned above the fiber works as a waveguide for the light to be transmitted, and the parameter that determines the frequency range of the electromagnetic wave able to pass is mainly the core diameter a as shown in the equation below (Hartog 2017):

$$V = \frac{2\pi}{\lambda} a \cdot NA \quad (2.4)$$

where V is the normalized frequency, λ is the wavelength and the numerical aperture NA is defined as

$$NA \equiv \sqrt{n_1^2 - n_2^2} = \sin \theta \quad (2.5)$$

that determines also the acceptance angle θ which is the maximum angle possible for a light ray to enter the fiber.

Two main types of fiber exist: the single-mode fibers and the multi-mode fibers (Figure 2.4)

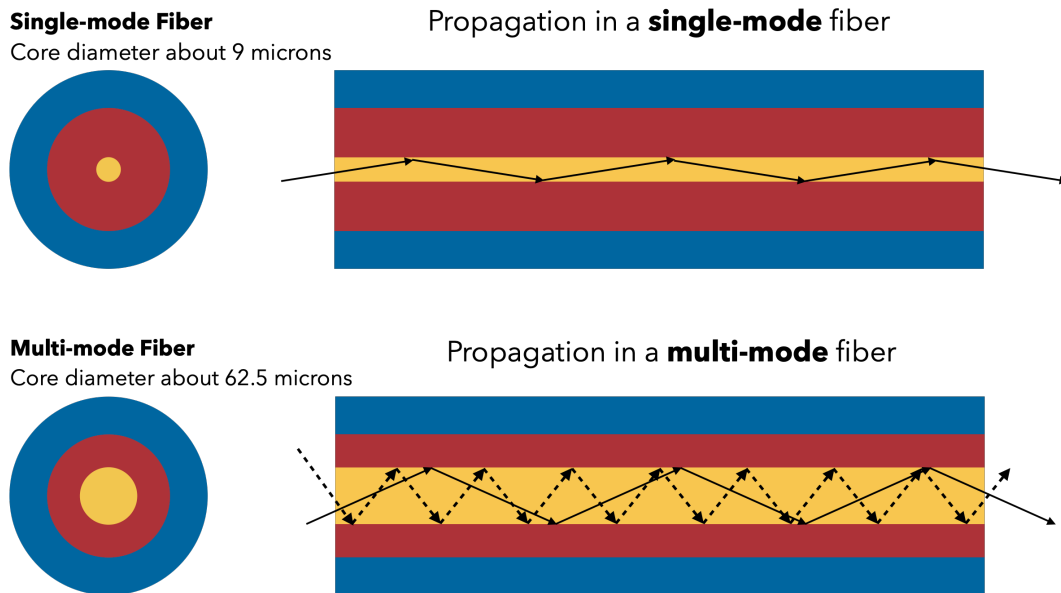


Figure 2.4: Single-mode and multi-mode fiber cross-section and ray propagation

The multi-mode have a bigger core diameter ($\sim 62.5\mu m$) and they are usually used for communication purposes because they are able to accept more frequency and consequently

carry more information. The single mode fibers instead have a smaller core ($\sim 9\mu m$) and allow the transmission of just one mode, these fibers are better suited for DAS purposes because they have a very low energy loss and dispersion (Soga and Luo 2018).

Another characteristic of the fiber, being it either single-mode or multi-mode, is the possibility of having the core with a step refractive index or graded refractive index (Hartog 2017). In the step refractive one, the core has a constant index, shown in Figure (2.5), on the other hand, the graded refractive index fiber has a non-constant shaped curve index, like the ‘bell’ curve shown in (2.5).

This non-constant shape bends the light rays inwards, increasing the incidence angle and consequently decreasing the possibility of a ray getting transmitted in the cladding and escaping the fiber causing signal energy loss. Usually, manufacturers of DAS systems and Interrogator Units, in order to simplify the calculation of the strain-rate usually take into account only the step-index fiber.

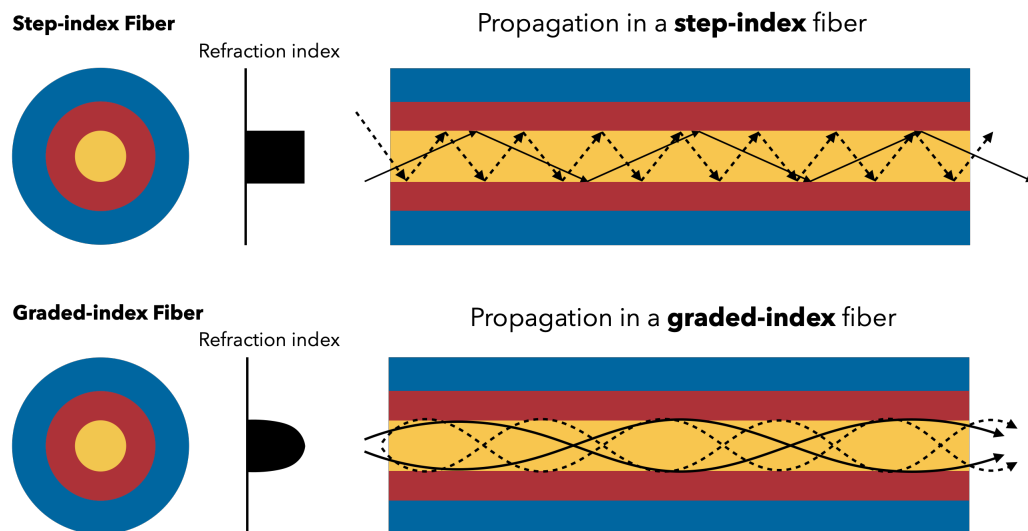


Figure 2.5: Step-index fiber and Graded-index for a multi-mode fiber cross-section, refractive index, and ray propagation.

Now that the main characteristics and structure of the fiber are explained the focus will be on the physical phenomena exploited to retrieve the signal (the light scattering) and how the strain rate is calculated. As previously mentioned, when a light pulse travels through the fiber-optic cable it encounters small imperfections, which are variations in the

refractive index, and generates scattering phenomena upon hitting them. There are three distinct types of scattering to consider in this context: Rayleigh, Raman, and Brillouin. While all three types can be used for surveys, Rayleigh scattering is typically the preferred choice for measuring strain-rates, because the other two are generally used to measure temperature changes (Fernández-Ruiz et al. 2020). Both Raman and Brillouin are inelastic scatterings caused by photon-photon interaction that releases energy. Raman is used to measure temperature changes only, instead Brillouin can be used to measure temperature or strain and it's caused by the compression of the fiber-optic cable during the traveling of the electromagnetic waves (Soga and Luo 2018).

Rayleigh scattering is generated when the wavelength is much bigger than the average radius of the defects ($\lambda \gg \text{impurities}$), and the impact occurs for low energies. With these conditions, the collision is purely elastic and therefore there are no frequency changes and the energy is conserved. A schematic representation of the three scattering phenomena is shown in Figure (2.6).

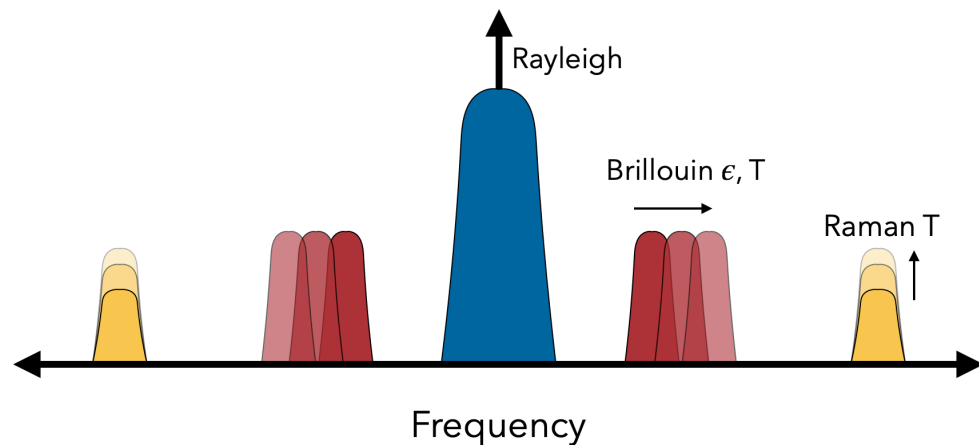


Figure 2.6: Frequency and Energy changes for light pulse scattering phenomena in the fiber.

In a DAS system, more generally called DVS (Distributed Vibration Sensor), the strain or strain-rate can be derived in different ways. We will focus on the OTDR (Optical Time Domain Reflectometry) which measures the phase difference of two signals in the time domain, like one of the most used interrogator units, the 'iDAS' by Silixa (Lindsey, Rademacher, and Ajo-Franklin 2020). The input is two time-shifted pulses, that can be

generated either by sending two pulses at a fixed time delay, the so-called dual-pulse method, or by an interferometer which takes a single pulse and shifts part of it by a fixed spacial length, the single-pulse method (Fernández-Ruiz et al. 2020).

A light pulse with a wavelength λ traveling through a fiber with a core refractive index n for a distance x undergoes a phase delay of:

$$\Theta = \frac{4\pi nx}{\lambda} \quad (2.6)$$

so two consecutive pulses have a phase difference equal to:

$$\Delta\Theta = \Theta(\tau_n + t_j, x_j) - \Theta(\tau_{n-1} + t_j, x_j) \quad (2.7)$$

where x_j is the channel position in the fiber, t_j is the time that the pulse takes to arrive at the position x_j , and τ_n is the starting time of the n pulse. The $\Delta\Theta$ in the optic-fiber is assumed to have a linear correlation with the following quantities (Lindsey, Rademacher, and Ajo-Franklin 2020):

$$\Delta\Theta(t, x_j) = \frac{4\pi n L_G}{\lambda} \left[\frac{\Delta x}{x} + \frac{\Delta n}{n} + \frac{\Delta \lambda}{\lambda} \right] \quad (2.8)$$

where L_G is the gauge length, Δx is the change in length of the fiber, Δn is the change of the refractive index, and $\Delta \lambda$ is the change in the wavelength. This equation implies that the fiber-optic is sensible to all three quantities (length, refractive index, and wavelength) and if not properly constructed, deployed, and isolated can generate very unreliable and noisy data that are dependent on multiple variables. Considering all the previous assumptions made in this chapter: single-mode, step-index, single-pulse, in time domain, and phase-based fiber, the terms $\frac{\Delta n}{n}$, mainly influenced by temperature, and $\frac{\Delta \lambda}{\lambda}$, mainly influenced by dispersion, can be considered constant and the only variable remains $\frac{\Delta n}{n}$ that can be seen equal to the strain along the fiber $\frac{\Delta n}{n} = \epsilon_{xx}$ so the equation becomes:

$$\Delta\Theta(t, x_j) = \frac{4\pi n L_G}{\lambda} \cdot \xi \cdot \epsilon_{xx}(t, x_j) \quad (2.9)$$

where ξ is a constant term depending on known material properties. Rearranging the terms the final strain equation is obtained:

$$\epsilon_{xx}(t, x_j) = \frac{\lambda}{4\pi n L_G \xi} \Delta\Theta \quad (2.10)$$

Lastly, to have a quantitative example, let's consider the 'iDAS' system with a wavelength $\lambda = 1,550nm$, a fiber refractive index $n = 1.45$, a Gauge length $L_G = 10m$ and a constant $\xi = 0.735$ provided by the manufacturer. the Equation (2.10) becomes:

$$\epsilon_{xx}(t, x_j) = \frac{1550 \cdot 10^{-9}(m)}{4\pi \cdot 1.45 \cdot 10 (m) \cdot 0.735} \Delta\Theta = 11.6 \cdot 10^{-9} \cdot \Delta\Theta \text{ (rad)} \quad (2.11)$$

3. Traveltime Computation for DAS

This chapter will describe the construction of the algorithm to compute the arrival times for an arbitrary velocity and source-receivers configurations. Most of the existing software tools for travel-time calculation are not able to manage sensors at different depths and with a very short inter-station distance (a condition required by DAS deployments in boreholes). The aim of this thesis is to develop a fast and efficient travel-time calculator based on the solution of the 2D Eikonal equation, and able to deal with high-density distribution of sensors at different depths. The approach is taken from the Podvin and Lecomte 1991, the inputs are a velocity model and a source position, and the output is an arbitrarily spaced subsurface model where the value of every point corresponds to the first arrival time.

Initially, the followed approach was the one done by Vidale 1988 based on the same principle of finite difference approximation of the eikonal equation. The main difference between the two algorithms is that in Vidale's only one wavefront is propagated from the source, so if the velocity model is complex, or a critical reflection occurs, some zones will not be illuminated. Those areas are called the "shadow zones" and the algorithm cannot calculate the travel time in them because the initial wavefront is not able to transmit further in the subsurface and gets totally reflected.

As shown in Figure (3.1) in case of a total reflection, occurring for a wave traveling from Layer 1 to Layer 2, where Layer 2 has a higher velocity with respect to Layer 1, the initial wavefront is not transmitted in Layer 2, but due to the interface condition, a head wave is generated which is able to propagate further in the subsurface. The head wave is a plane wave that travels with the velocity of Layer 2 and is constructed as the envelope of all the wavefronts generated by secondary sources present at the interface (red dots in Figure (3.1)).

The behavior of the head waves cannot be resolved with Vidale's approach, for this reason, Podvin and Lecomte 1991 proposed an eikonal solver where all the points in the model can be a secondary source of a wavefront. Podvin's algorithm is able to reconstruct the head wave by propagating multiple wavefronts when a critical reflection occurs.

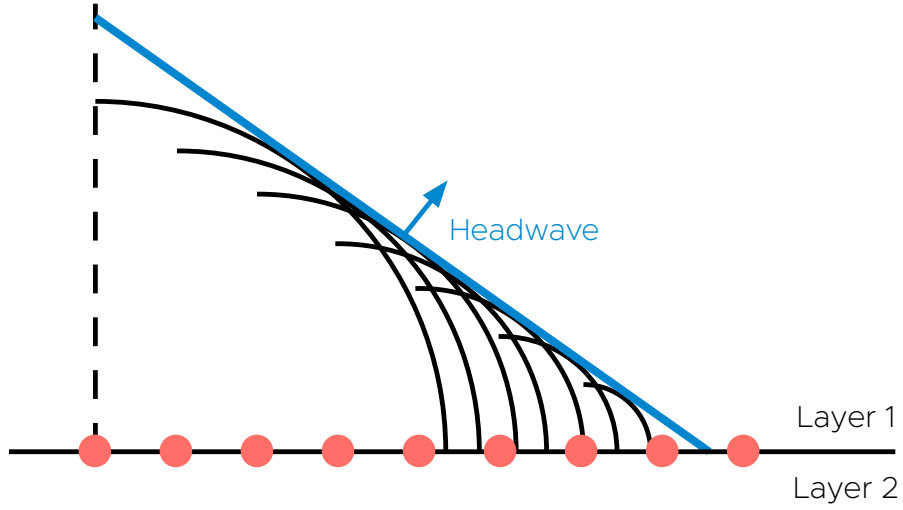


Figure 3.1: Schematic representation of a head wave propagation (Podvin and Lecomte 1991).

Before going into details on the travel time calculator, it's important to explain the derivation of the eikonal equation and why it is used (for a full, and rigorous, derivation see Ammon et al. 2020).

For a uniform, isotropic, linear elastic medium the wave equation can be written as:

$$\nabla^2 \phi = \frac{1}{\alpha^2} \ddot{\phi}, \quad (3.1)$$

where ϕ is the displacement vector's potential, and α is a constant related to the velocity of the wave in the subsurface. The displacement can be decomposed using Helmholtz's theorem into P-wave and S-wave potential. For the P-wave, it becomes:

$$u_p(\mathbf{x}, t) = \nabla \phi(\mathbf{x}, t) = \frac{\partial \phi}{\partial x_1} \hat{\mathbf{x}}_1 + \frac{\partial \phi}{\partial x_2} \hat{\mathbf{x}}_2 + \frac{\partial \phi}{\partial x_3} \hat{\mathbf{x}}_3, \quad (3.2)$$

where $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{x}}_3$ are the versors in Cartesian coordinates. Substituting Equation (3.2) into

Equation (3.1) it becomes:

$$\frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} + \frac{\partial^2 \phi}{\partial x_3^2} = \frac{1}{c^2(\mathbf{x})} \frac{\partial^2 \phi}{\partial t^2}, \quad (3.3)$$

notice that the term $1/c^2(\mathbf{x})$ represents the velocity of the wave at position \mathbf{x} . Now, if we assume a functional $\phi(\mathbf{x}, t)$ in the form:

$$\phi(\mathbf{x}, t) = A(\mathbf{x})e^{i\omega(T(\mathbf{x})-t)} \quad (3.4)$$

where $A(\mathbf{x})$ is the amplitude of the potential as a function of position, and $T(\mathbf{x})$ is also a function of position but with units of time. The contours of $T(\mathbf{x})$ represent surfaces of identical travel times, known as wavefronts. Substituting Equation (3.4) into Equation (3.3) we get:

$$\nabla^2 \left[A(\mathbf{x})e^{i\omega(T(\mathbf{x})-t)} \right] = \frac{1}{c^2(\mathbf{x})} \frac{\partial^2}{\partial t^2} \left[A(\mathbf{x})e^{i\omega(T(\mathbf{x})-t)} \right]. \quad (3.5)$$

After computing the derivatives and separating the real and imaginary parts, we can write the real part as (see full derivation on Ammon et al. 2020):

$$\left(\frac{\partial T(\mathbf{x})}{\partial x_1} \right)^2 + \left(\frac{\partial T(\mathbf{x})}{\partial x_2} \right)^2 + \left(\frac{\partial T(\mathbf{x})}{\partial x_3} \right)^2 - \frac{1}{c(\mathbf{x})^2} = \frac{1}{\omega^2 A(\mathbf{x})} \nabla^2 A(\mathbf{x}), \quad (3.6)$$

and, for high frequencies (or small wavelengths) Equation (3.6) finally becomes:

$$\left(\frac{\partial T(\mathbf{x})}{\partial x_1} \right)^2 + \left(\frac{\partial T(\mathbf{x})}{\partial x_2} \right)^2 + \left(\frac{\partial T(\mathbf{x})}{\partial x_3} \right)^2 = \frac{1}{c(\mathbf{x})^2} \quad (3.7)$$

which is called the eikonal equation.

The eikonal equation can be seen as a high-frequency approximation of the wave equation and carries information about the travel times of the wave $T(\mathbf{x})$ as a function of position. For this reason, it can be used to calculate the first arrival time of an event, without the need to compute other quantities (such as the amplitudes), and quickly provide a travel time curve that will be used to create the synthetic seismograms.

The code developed, as already mentioned, is based on Huygens' principle, the travel time of one point is calculated locally, based on all the travel times of his surrounding grid-points, which act as a secondary source of wavefronts. This code, made to generate

synthetic seismograms, simulates the wavefronts in the 2-D case. 3-D implementation can be done by following the same approach written below, but for our aim was not necessary.

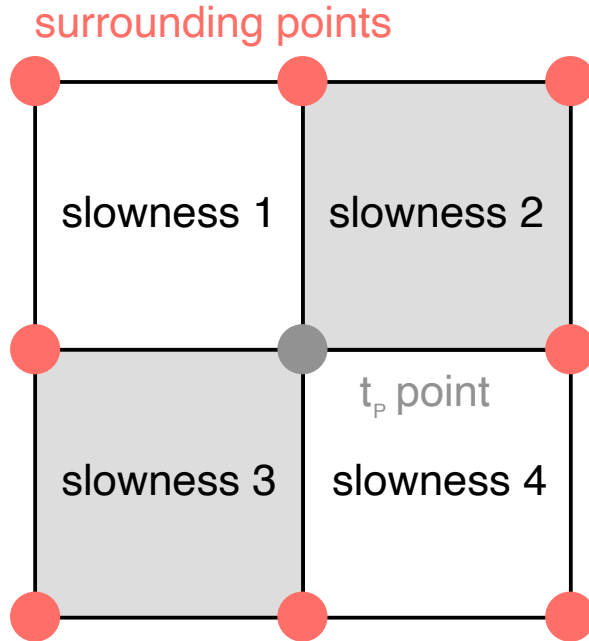


Figure 3.2: Schematic representation of the grid-points and their slowness. The travel time is calculated for the central point t_p based on the arrival time of the surrounding points (Podvin and Lecomte 1991).

Figure (3.2) shows the portion of the grid needed to compute a travel time calculation. The central point t_p in grey will have a travel time value based solely on the travel times of his surrounding points in red and their values of slowness ($s = 1/c$), regardless of the source position, the location of the t_p in the grid, and the fact that the surrounding point where reached by a wavefront or not (if not the default travel time will be infinite). Considering the fact that in the mesh squares the slowness is constant and the interpolation between the grid-points is linear, the waves propagating in a single square can be seen as local plane wavefronts. So, locally, Huygens' principle is not used as an explanation tool for the computation of the travel times, instead, plane wavefront (or rays) can describe the propagation.

A very important parameter that must be set before computing the eikonal equation

is the slowness model mesh dimension, as shown in Figure (3.2). The slowness model is divided into a number of equal size squares and each square is assigned a constant value of slowness. The vertices of the square are the travel time grid points. The square side and the time point spacing have the same length, and as a general rule should be one order of magnitude smaller than the slowness anomalies that we want to identify.

Every point that acts as a secondary source can generate three different types of waves: Transmitted waves, Head waves, and Diffracted waves. The 2-D algorithm will have for a point P 8 neighbors (Figure (3.2)), creating: 8 Transmitted waves, from the outside interfaces, 4 Head waves, from the inner interfaces separating the four slowness, and 4 Diffracted waves, from the four corners of the square formed by the eight neighbors, giving a total of 16 wavefronts reaching the point t_P . The times will be confronted and the smallest one will become the travel time for the point t_P .

Under this conditions, the eikonal equation (3.7) for a 2-D model becomes:

$$\left(\frac{\partial t}{\partial x}\right)^2 + \left(\frac{\partial t}{\partial y}\right)^2 = s^2 \quad (3.8)$$

where s is the slowness value of a single square. Now, if we discretize the partial derivatives with respect to x and y , considering a spacing h between two points, we get, respectively :

$$\frac{\partial t}{\partial x} = \frac{t_N - t_M}{h} \quad (3.9)$$

$$\frac{\partial t}{\partial y} = \frac{t_P - t_N}{h} \quad (3.10)$$

where t_P, t_M, t_N are the arrival times calculated for the points P, M, N . A graphical representation is shown in Figure (3.3)

After substituting Equation (3.9) and Equation (3.10) into Equation (3.8), the eikonal equation becomes:

$$\left(\frac{t_N - t_M}{h}\right)^2 + \left(\frac{t_P - t_N}{h}\right)^2 = s^2. \quad (3.11)$$

Then, after some steps, we are able to isolate the t_P which is the travel time to be calculated:

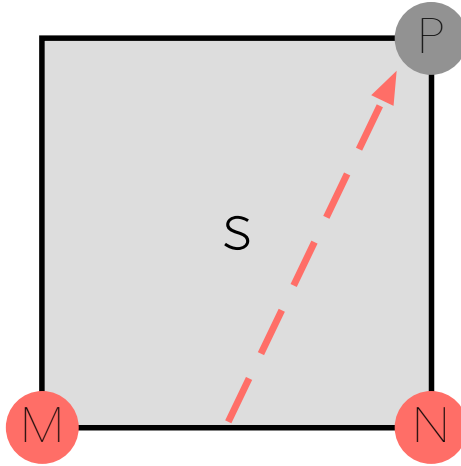


Figure 3.3: Schematic representation of a Transmitted wave from the interface MN to the point P . P, M, N are the time points and S is the slowness.

$$(t_P - t_N)^2 = (hs)^2 - (t_N - t_M)^2 \quad (3.12)$$

$$\sqrt{(t_P - t_N)^2} = \pm \sqrt{(hs)^2 - (t_N - t_M)^2} \quad (3.13)$$

$$t_P = t_N \pm \sqrt{(hs)^2 - (t_N - t_M)^2}. \quad (3.14)$$

The final result in Equation (3.14) must satisfy two conditions, called "illumination conditions":

$$0 \leq (t_N - t_M) \leq hs/\sqrt{2}, \quad (3.15)$$

to avoid negative square roots arguments, and:

$$\frac{\partial t}{\partial y} \geq 0, \quad (3.16)$$

to have only the positive value of the square root, which means that the wavefront is propagating outward and not inward. Actually, the condition in Equation (3.16) is always satisfied as a consequence of Fermat's principle (Podvin and Lecomte 1991). A negative gradient $\frac{\partial t}{\partial y} < 0$ means that the arrival time is estimated along a path that violates the eikonal equation, and, in such case, the time will always be greater than the other first arrivals. Taking into account the condition in Equation (3.15) and the just mentioned positive gradient (Equation (3.16), Equation (3.14) can be written as:

$$t_P \equiv t_T = t_N + \sqrt{(hs)^2 - (t_N - t_M)^2} \quad (3.17)$$

where T_T is the Transmitted wave arrival time from the interface MN to the point P .

The second kind of wave that can propagate from a point source is a Head wave. The Head wave is generated between two interfaces of different slowness and travels with the highest velocity of the two.

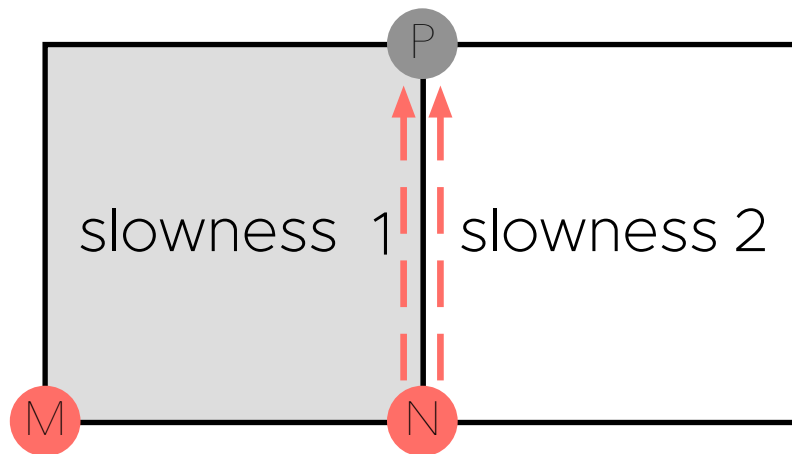


Figure 3.4: Schematic representation of a Head wave from the point N to the point P . P, M, N are the time points and $slowness1$ and $slowness2$ are the slowness.

In Figure (3.4) we can see an example of a possible Head wave propagation: let's

consider a Transmitted wave arrival time t_T from the interface MN to the point P , as seen before. t_T is not influenced by the adjacent $slowness_2$ but if his value is low enough the Transmitted wave will not be the first to arrive at point P :

$$t_T - t_N > hs_2 , \quad (3.18)$$

this means that the Head wave will be faster than the Transmitted Wave in reaching point P . The relation in Equation (3.18) means that the incidence angle of the Transmitted wave generated by the interface MN hits the interface NP exceeding the critical angle, consequently a Head wave from point N to point P must be computed. The travel time formula for a Head wave, considering the already mentioned geometry is:

$$t_P \equiv t_H = t_N + h \min(s_1, s_2) \quad (3.19)$$

Where t_H is the travel time of a Head wave from point N to point P , and s_1, s_2 are respectively $slowness_1, slowness_2$.

The Third, and last, wave considered is the Diffracted wave. The Diffracted wave is generated by one of the corner points that act as a diffractor point from which the wavefront is propagated.

In Figure (3.5) an example is shown: if the point P cannot be reached by the transmitted waves inside the grid a shadow zone will form. In that case, a Diffracted wave from the point M will illuminate the point P traveling at the constant velocity of the cell. The travel time formula for a Head wave is the following:

$$t_P \equiv t_D = t_M + hs\sqrt{2} \quad (3.20)$$

where t_D is the arrival time of a head wave from the point M to the point P .

Now that the wavefront arrivals at a single point P are described we can calculate the travel time for all the points in the grid. Knowing the velocity model and the source position we can create a mesh with arbitrary spacing for the travel time grid. the grid is created with all times set to infinite and then the source position is set to zero $t_{source} = 0$ as the starting point of the first wavefront. Then the algorithm proceeds outward calculating

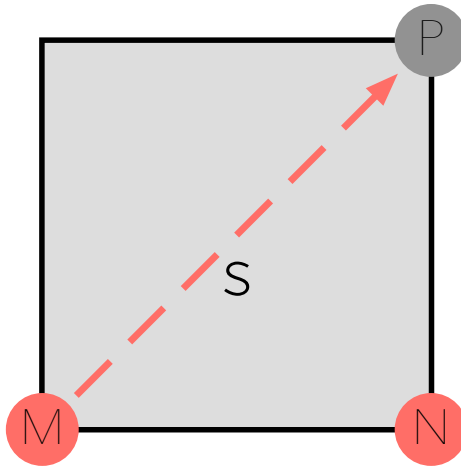


Figure 3.5: Schematic representation of a Diffracted wave from the point M to the point P . P, M, N are the time points and S is the slowness.

the travel times for points laying on squares of increasing sizes until all the sides of the squares hit the boundaries of the grid. The propagation scheme and the order of calculation for the points on all the sides are shown in Figure (3.6). Let's take as an example the upper side of the square, marked as side "1" in Figure (3.6), the points will be calculated starting from the earliest arrival of the previous row and going towards the latest arrival time, also from the previous row. Then, the newly calculated times will dictate the computation order for the next, more outward, row. This is done to mimic the propagation behavior of the wavefronts and to ensure that the first arrivals are the first to propagate in the next row, so that possible head or diffracted waves can be correctly simulated.

To test the reliability of the code we tested it with different velocity models of increasing complexity. firstly, we calculated the travel time for a homogeneous velocity model of $v = 2000 \frac{m}{s}$ and a source position at $1500m$ depth and $1500m$ distance, shown in Figure (3.7).

The right picture corresponds to the velocity model, with the white star corresponding

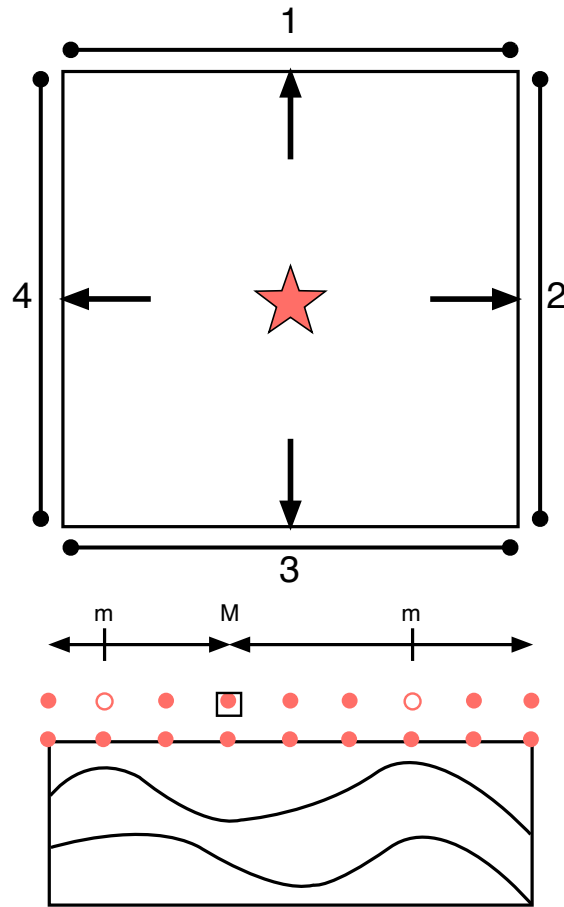


Figure 3.6: Scheme of the algorithm propagation from the source (red star) and calculation order of arrival time for a single row (Podvin and Lecomte 1991).

to the source position, and the left one is a graphical visualization of the wavefronts. All the wavefront plots in this thesis are meant to give the reader a visual idea of how the travel time matrix values change with respect to the velocity model, this is because the results of the travel time matrix are difficult to interpret by just looking at the color gradient. From the travel time matrix we extract the maximum value of arrival time, and then the

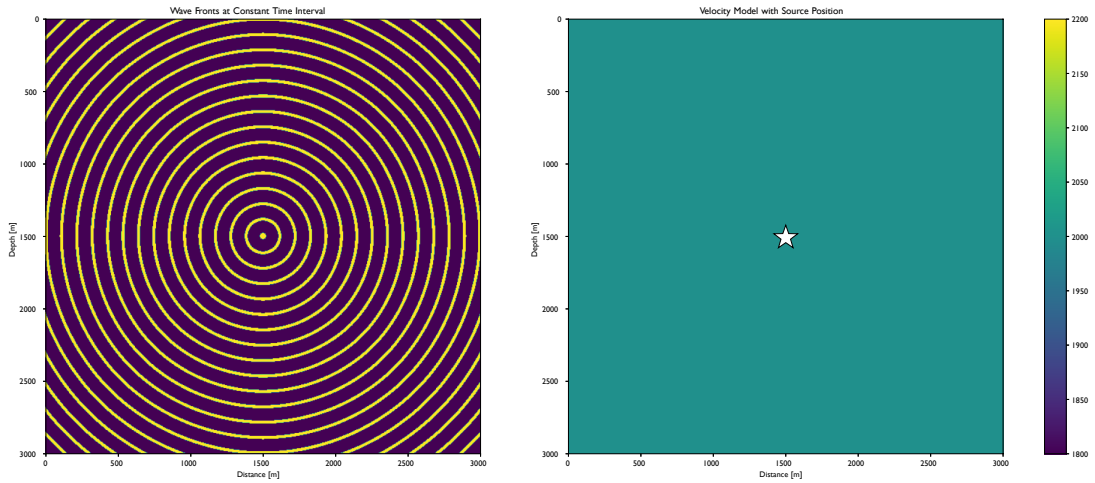


Figure 3.7: Wavefront and velocity model for a homogeneous subsurface model.

wavefronts plot is generated by creating 20 equally time-spaced wavefronts with a time duration of a hundredth of the maximum $\frac{1}{100} \cdot \text{MAX}(\text{traveltime})$.

The first test compares the travel time grid of a homogeneous velocity model with the analytical solution. The analytical solution is computed by calculating the travel time from the source to the points in the mesh:

$$t_P = \sqrt{(z_s - z_r)^2 + (x_s - x_r)^2} \cdot \frac{h}{v}, \quad (3.21)$$

where z_s, x_s are the coordinates of the source and z_r, x_r are the coordinates of the receivers, h is the mesh spacing, and v is the layer velocity (constant). Taking the result from the simulation done in Figure(3.7) and the analytical solution we get the results in Figure (3.8). The result clearly shows that the bigger error is present in the vicinity of the source, this is because the numerical approximation and the algorithm construction fail to reproduce the wave propagation. This issue was mentioned also in Podvin and Lecomte 1991 and Vidale 1988 and there are quite a few ways to achieve better results, one of the easiest approaches is to calculate the analytical solution for the biggest square with a homogeneous velocity around the source. Regardless, for our aim those solutions are not needed because we don't normally have events so close to the receivers and they don't improve the accuracy of the far-field travel times. For a grid length of 3km by 3km and a spacing of 5m , the

errors drop to below 1% when we are more than 100m away from the source.

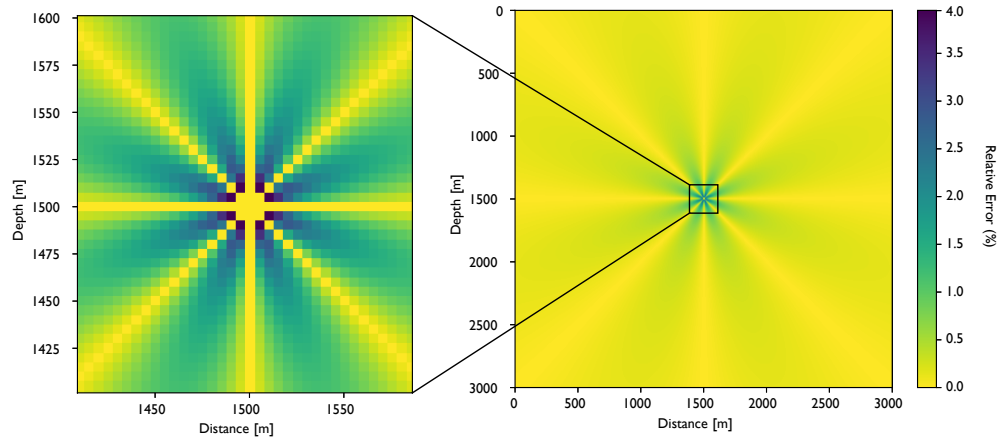


Figure 3.8: Relative difference, expressed in percentage, between the analytically calculated travel time and the travel time from the eikonal solver.

Having verified that the eikonal solver is able to calculate with a good level of precision the travel times, it must be tested with some velocity model with increasingly complex geometries and higher velocity contrasts. Nevertheless, as pointed out by Podvin and Lecomte 1991 there is one last issue to be fixed regarding the critical reflections. When the wavefront travels from a layer with a slower velocity to a layer of faster velocity a critical reflection occurs and a head wave is generated. Let's consider the upper side of the propagation square, as shown in Figure (3.9).

If that side encounters a boundary between a slow and fast medium there might be a point at which the head wave is marked as the first arrival, if that is so, the head wave must be propagated to the side's extremity to correctly time it. Then, the normal procedure of travel time order calculation will be applied for that side, as explained above. The last procedure to be implemented is the so-called back-propagation. After calculating the head wave along the side of the square where the critical reflection occurred it must be back-propagated to the source, as shown in Figure (3.9). The back-propagation stops when no arrival time is lowered on the entire row, after that, the normal outward propagation resumes from the row where the critical reflection first occurred. This procedure mimics the behavior of the wavefronts and it's able to reconstruct the head wave propa-

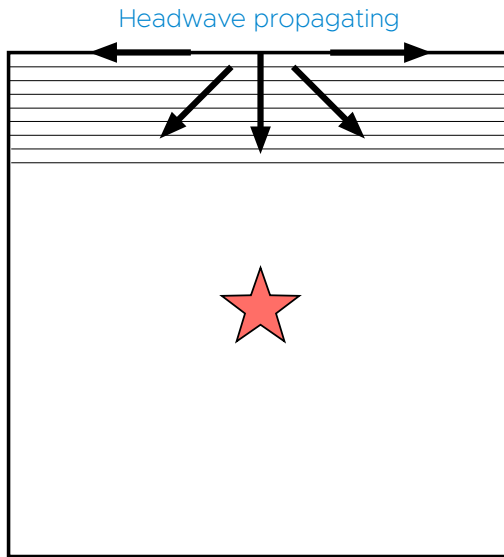


Figure 3.9: Schematic representation of the back-propagation implementation after the occurrence of a critical reflection (Podvin and Lecomte 1991).

gation. An example of the implementation of the back-propagation can be seen in Figure (3.10). The velocity model has 3 horizontal layers with increasing velocities and a strong velocity contrast to emphasize the head wave propagation: the second layer has a velocity of $3000 \frac{m}{s}$ and the third layer has a velocity of $10000 \frac{m}{s}$. The source is positioned at $0m$ depth and $0m$ distance. In Figure (3.10 b) the result without back-propagation is shown: in the red rectangle we can see some artifacts and the discontinuity of the wavefronts where the head wave firstly occurs, which are nonphysical results, and in the green rectangle we can see how these wrong calculations generate scattered wavefronts in the first layer. In Figure (3.10 c) instead the back-propagation is implemented: it's able to reconstruct correctly the head wave generation in the red square, eliminate the blur in the green square, and calculate the correct travel times for all the grid.

Now that the algorithm is complete it is able to handle very complex geometries and strong velocity contrast. The first model tested is a layered model similar to the one seen in Figure (3.10 a) with increasing velocities, in this case, the source is positioned at $500m$

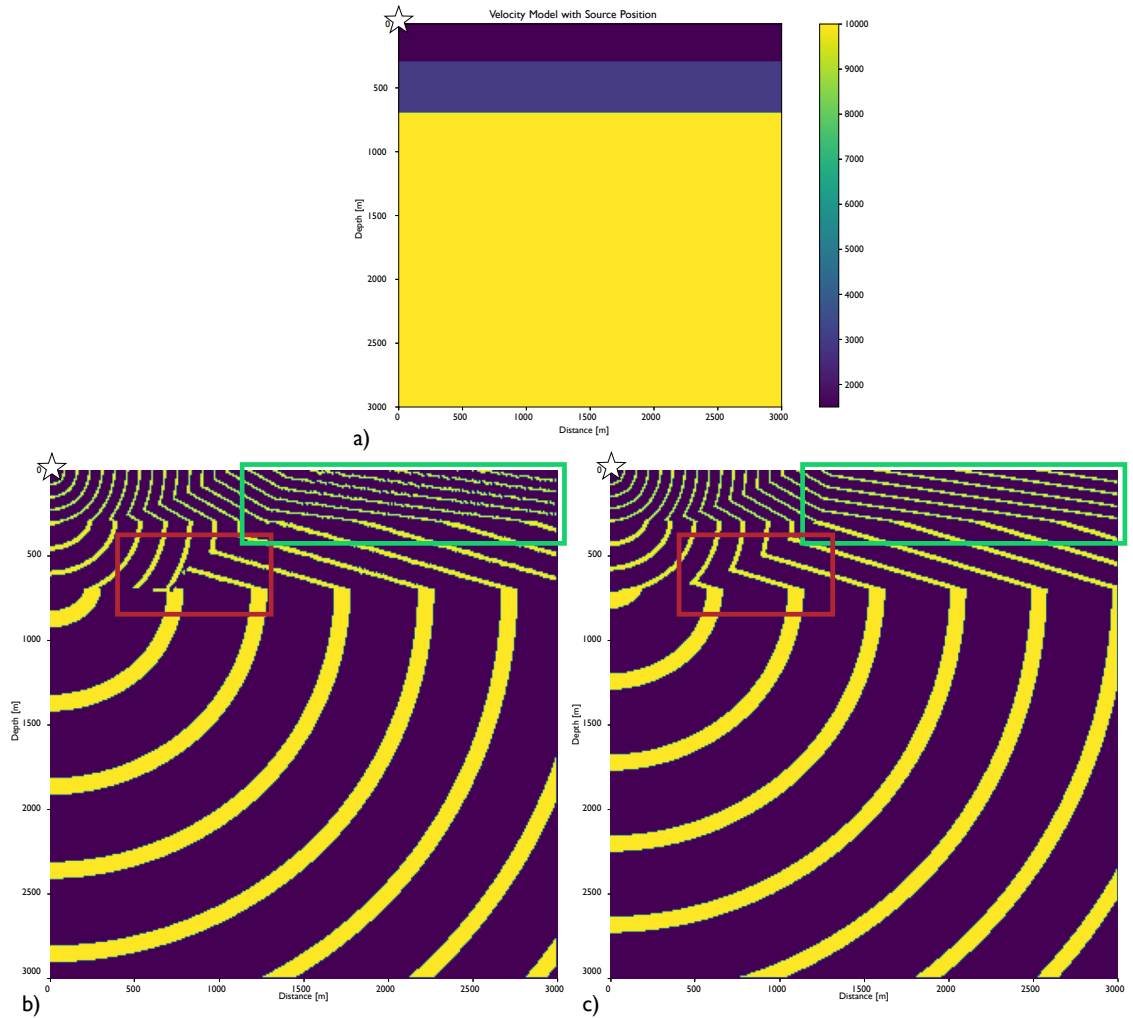


Figure 3.10: a) Velocity model used for the travel time calculation. b) Travel time matrix generated without the back-propagation algorithm, note the nonphysical results in the red and green rectangles. c) Travel time matrix generated with the back-propagation algorithm, note how the nonphysical results in the rectangle are resolved.

depth and 500m distance. The resulting wavefronts are shown in Figure (3.11) and we can see the head waves generated at every interface and the wavefronts that are getting thicker by traveling in higher velocity mediums.

The next "benchmark" model is a layered model with a dipping interface. In Figure (3.12) we can see that the velocity of the first and third layers is $2000 \frac{m}{s}$ and for the second is $6000 \frac{m}{s}$. the source is positioned in the origin at 0m depth and 0m distance. The algorithm correctly reproduces the wavefronts and handles the head wave generation even

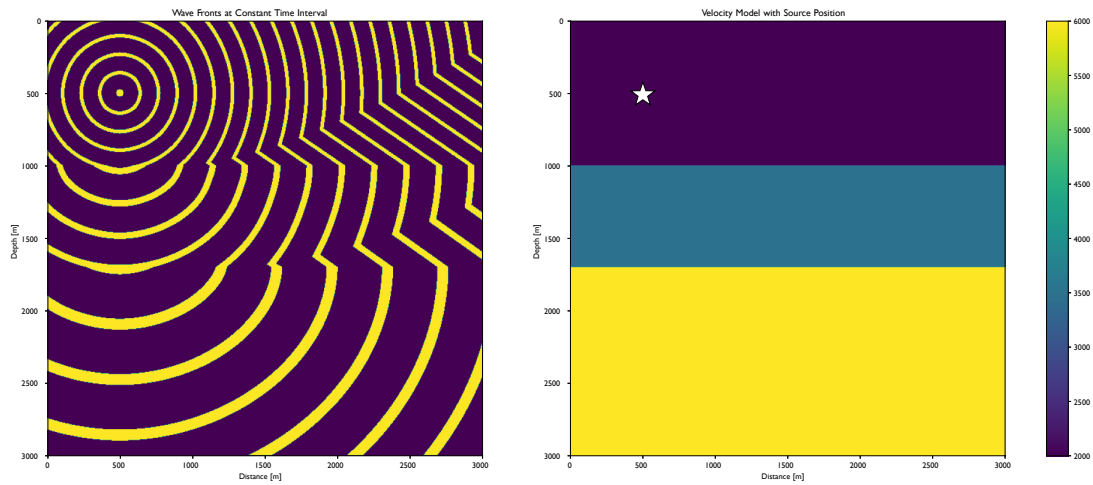


Figure 3.11: Wavefront and velocity model for a layered subsurface model.

on a dipping interface.

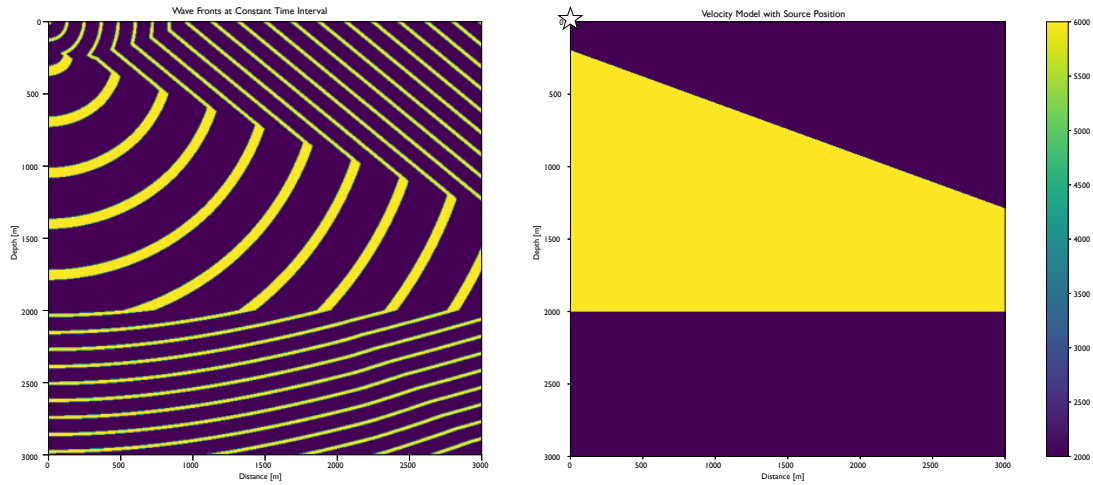


Figure 3.12: Wavefront and velocity model for a dipping layered subsurface model.

The last test was done also by Podvin and Lecomte 1991 and consists of a homogeneous fast model with a squared heterogeneity of slow velocity, which can be seen in Figure (3.13). The resulting wavefronts show how the algorithm is able to reconstruct the wavefront superposition inside and outside the "box".

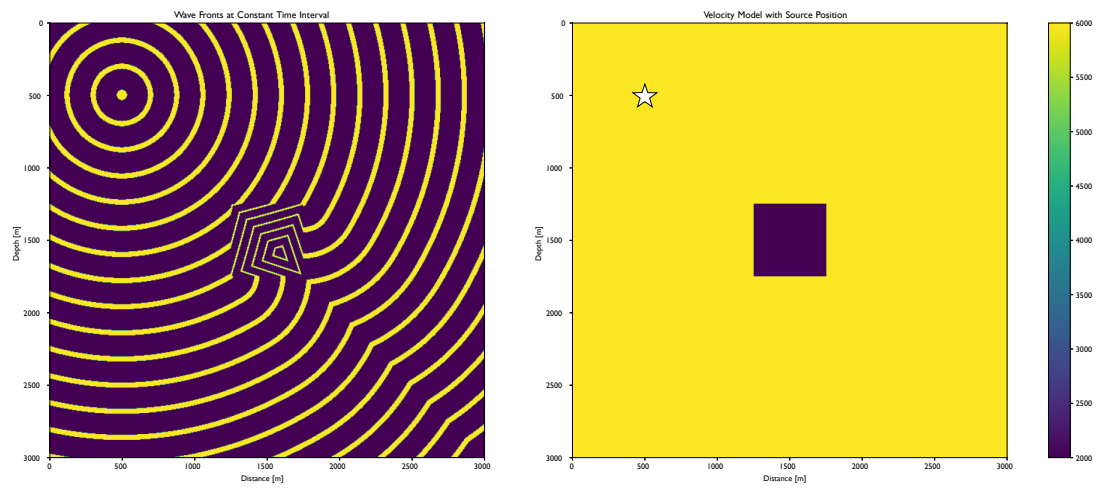


Figure 3.13: Wavefront and velocity model for homogeneous subsurface with a square heterogeneity

4. Modeling of DAS Microseismic Data

This chapter will describe the workflow we implemented to generate synthetic DAS seismograms. Although DAS is becoming very popular in seismological applications, there is still a scarcity of publicly available data, thus the testing of new data analysis methods can be only done using synthetics. However, the algorithms used to generate synthetic DAS waveforms are still few and computationally heavy due to the high spatial and temporal sampling of DAS data acquisition. Our aim is to create a workflow for the generation of synthetic DAS waveforms suitable for an arbitrary fiber geometry, source location, and source mechanism. To achieve this, firstly we use the Python library for the calculation of the first P and S onset times in a 2D medium (Podvin and Lecomte 1991) that we developed and described in the previous section. Then, for a chosen fiber geometry, we generate DAS synthetic recordings using the convolutional model with an arbitrary wavelet, the source radiation pattern, and the directivity of the fiber. Despite conceptually simpler than many other algorithms used to generate synthetic seismograms, the convolutional model is fast and provides first-order approximation (using only kinematic information) synthetics that may be useful in many applications including the testing new data analysis methods or the evaluation of the location performances of a particular DAS acquisition geometry. Furthermore, data intensive applications requiring large DAS microseismic datasets (e.g. the training of Machine Learning methods for the detection and location of microseismicity) may benefit from this workflow. In this case we can generate synthetics that can be used for the training process of Machine Learning Models. While a single simulation with our workflow requires a few seconds in a laptop computer (2 cores and 8 GB of RAM), the spectral element simulation requires a few minutes in a multi-core computer (with 18 cores and 64 GB of RAM).

It is important to mention that the model is only kinematic since the energy partition at the interfaces is not taken into account. Although this may seem a strong approximation, it is important to remind that DAS systems record strain or strain-rate and not displacements or velocities, furthermore absolute amplitudes of DAS waveform are not used because

the instrument response of the interrogator is generally not known. For this reason, it is often useful to work with normalized waveforms. In this context, the workflow we implemented to generate synthetics, even if not physically accurate, is certainly useful for the applications above described.

We model DAS P and S wave seismograms $S(\mathbf{x}_s, \mathbf{x}_r, t)$ for a given source location x_s , receiver location x_r and time t (setting the origin time equal to 0) using the following convolution equation:

$$S(\mathbf{x}_s, \mathbf{x}_r, t)^k = \int W(\theta) A^k(\mathbf{x}_s, \mathbf{x}_r) \delta(t - \tau^k(\mathbf{x}_s, \mathbf{x}_r) - \theta) d\theta \quad k \in \{P, S\} \quad (4.1)$$

where δ is the Dirac's delta, $\tau^k(\mathbf{x}_s, \mathbf{x}_r)$ is the computed P or S traveltime at a fiber's channel, and it is based on the eikonal equation computed with the method illustrated in Chapter (3), as a function of the channels \mathbf{x}_r and source \mathbf{x}_s positions. W is the source wavelet, in this case we tested two different types of wavelets: an exponentially decaying sinusoidal wavelet (Figure(4.1)) and the derivative of a Ricker wavelet (Figure(4.2)) to compare how they achieve to reconstruct a the seismograms.

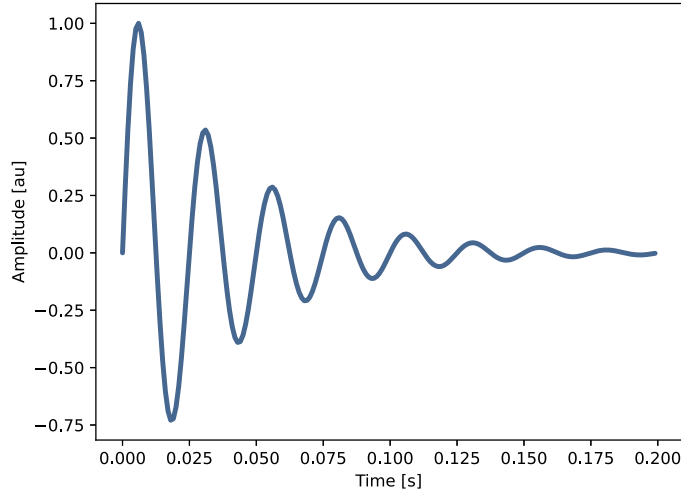


Figure 4.1: Exponentially decaying sinusoidal wavelet.

$A^k(\mathbf{x}_s, \mathbf{x}_r)$ is an amplitude correction term that takes into account the source radiation pattern and the directivity of the fiber, and it is described by the following equation:

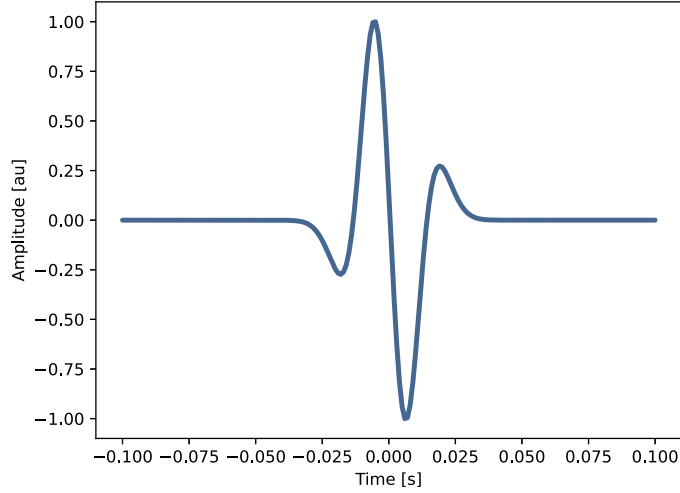


Figure 4.2: Derivative of Ricker wavelet.

$$A^k(\mathbf{x}_s, \mathbf{x}_r) = \frac{A}{r} D^k(i_h) \mathcal{F}^k(\phi_f, \delta, \lambda, i_h) \quad k \in \{P, S\} \quad (4.2)$$

Where A is an arbitrary constant, $D(i_h)$ is the Directivity of the fiber as a function of the incidence angle i_h , r is the source-channel distance, while $\mathcal{F}(\phi_f, \delta, \lambda, i_h)$ is the source radiation pattern both for P and S waves of the source as a function of the strike ϕ_f , the dip δ , the rake λ , and the incidence angle i_h . The Directivity $D(i_h)$ depends on the angle of incidence between the source and the channels, this is because the fiber elongates only along its main axis and therefore can register a signal only if the waves are polarized in the same direction. For P-waves and S-waves respectively the Directivity is (Mateeva et al. 2014, Näsholm et al. 2022):

$$D_P(i_h) = \cos^2 i_h \quad , \quad (4.3)$$

$$D_S(i_h) = \sin 2i_h \quad . \quad (4.4)$$

This relationship is better explained graphically in Figure(4.3) where also the comparison with a standard seismometer is shown. For a horizontal fiber axis, the directivity of the P-wave has a maximum for $i_h = 0$ in both instruments and goes to 0 at $i_h = \frac{\pi}{2}$. Instead,

for an S-wave the behavior differs between the two instruments: for the DAS we observe four lobes with different polarities and four zeros at $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$.

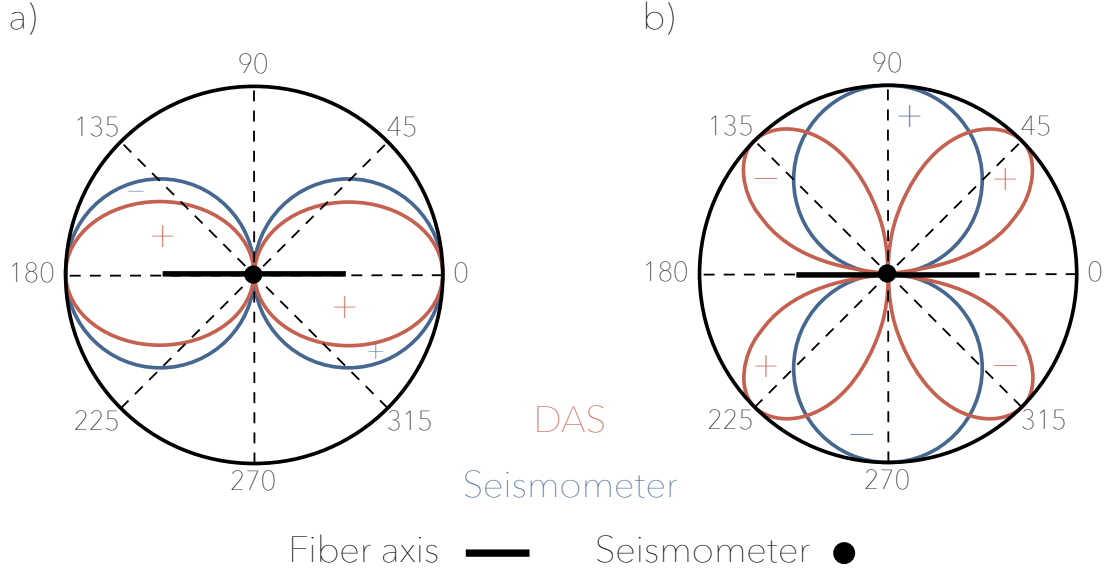


Figure 4.3: Directivity comparison between a fiber (red curve) deployed horizontally and a seismometer (blue curve) for a) P-waves, b) S-waves.

The term $\mathcal{F}(\phi_f, \delta, \lambda, i_h)$ is the source Radiation Pattern that describes how the forces generated by the source influence the displacement and velocity of the subsurface. for a double-couple source and a far-field approximation, the Radiation Pattern of P, SV, and SH waves is (Ammon et al. 2020)(Aki and Richards 2002):

$$\begin{aligned} \mathcal{F}^P = & [\cos \lambda \sin \delta \sin(-2\phi_f) - \sin \lambda \sin(2\delta) \sin^2(-\phi_f)] \sin^2 i_h \\ & + [\sin \lambda \cos(2\delta) \sin(-\phi_f) - \cos \lambda \cos \delta \cos(-\phi_f)] \sin(2i_h) \\ & + \sin \lambda \sin(2\delta) \cos^2 i_h \end{aligned} \quad (4.5)$$

$$\begin{aligned} \mathcal{F}^{SV} = & [\sin \lambda \cos(2\delta) \sin(-\phi_f) - \cos \lambda \cos \delta \cos(-\phi_f)] \cos(2i_h) \\ & + \frac{1}{2} \cos \lambda \sin \delta \sin(-2\phi_f) \sin(2i_h) \\ & - \frac{1}{2} \sin \lambda \sin(2\delta) \sin(2i_h) [1 + \sin^2(-\phi_f)] \end{aligned} \quad (4.6)$$

$$\begin{aligned}
\mathcal{F}^{SH} = & [\cos \lambda \cos \delta \sin(-\phi_f) + \sin \lambda \cos(2\delta) \cos(-\phi_f)] \cos i_h \\
& + [\cos \lambda \sin \delta \cos(-2\phi_f) \\
& - \frac{1}{2} \sin \lambda \sin(2\delta) \sin(-2\phi_f)] \sin i_h
\end{aligned} \tag{4.7}$$

where ϕ_f is the strike, δ the dip, λ the rake, and i_h the incidence angle. Finally the simulated DAS waveforms $S(\mathbf{x}_s, \mathbf{x}_r, t)$ can be written as:

$$S(\mathbf{x}_s, \mathbf{x}_r, t) = S(\mathbf{x}_s, \mathbf{x}_r, t)^P + S(\mathbf{x}_s, \mathbf{x}_r, t)^S \tag{4.8}$$

It is important to note that, while real DAS data consists of strain or strain-rate measures, our simulations contain information regarding the arrival times extracted by solving the eikonal equation, which does not depend on strain or strain-rate. Nevertheless, for our aim those parameters are not needed, and we can enrich the seismograms with properties such as directivity or radiation pattern. Our model takes into account only the kinematics of the first P and S wave arrivals, and although some amplitude correction term is included, it does not compute the energy partition at the interfaces. For this reason, our model applications do not need seismic amplitude information. In fact, we developed this workflow to test the detection and location algorithms (also ML-based) that use DAS data that we are going to develop within the Seismology group at the University of Pisa.

We are going to show in the following part various applications and comparisons with both synthetic and real data. The database used for computing the synthetic seismograms is taken from the FORGE experiment based in Utah, USA (T. Martin and Nash 2019) where they deployed the optic fiber in 3 wells along with seismometers to measure both signals. In this work, we will only use the data recorded by the optic fiber. The first step is to get the velocity model (courtesy of Lellouch et al. 2021) measured directly in the well until approximately 1000m depth and extend it to a depth of 3000m using geological references. The model extracted is a 1D velocity model with a layer spacing of 50m for v-P and v-S, which for seismological scales is a very detailed model (right images in Figure(4.4) and Figure(4.5)).

Then, for the first simulation we considered an event with a source at 2437m depth and 1000m distance from the source and we computed the wavefronts for both P and S waves.

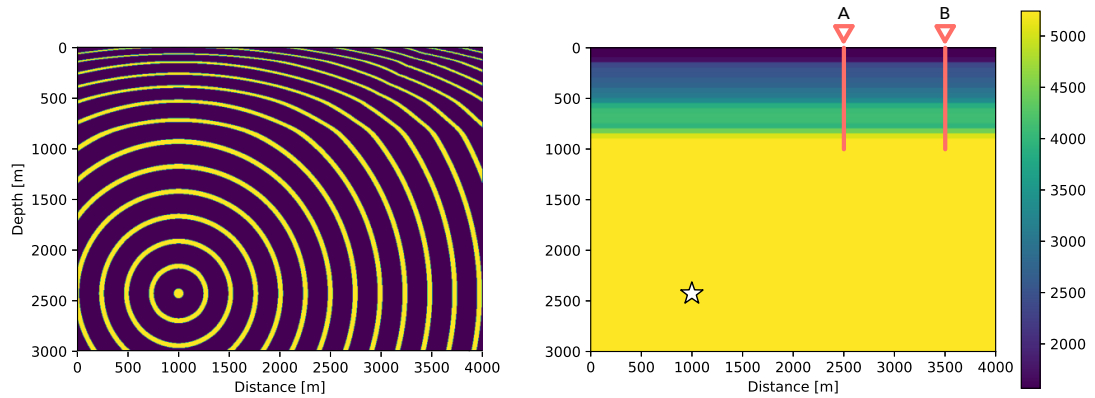


Figure 4.4: Wavefront and V-p velocity model for the FORGE experiment (Utah, USA).

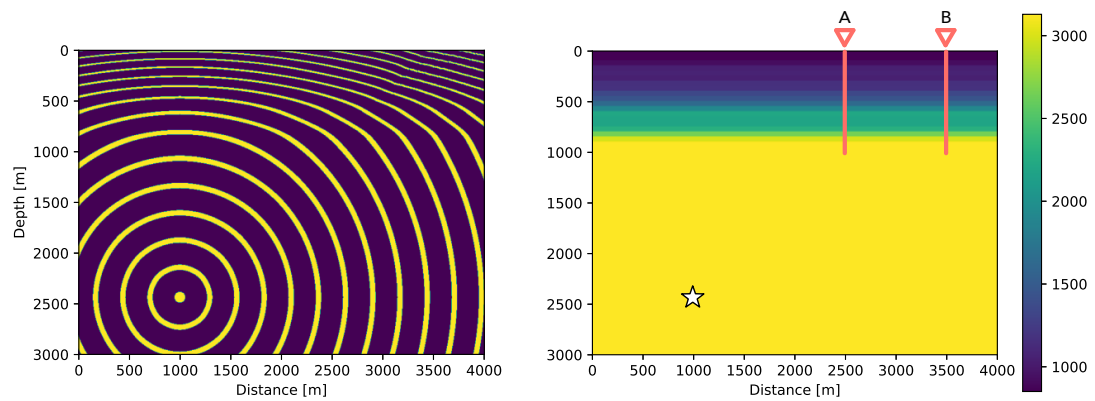


Figure 4.5: Wavefront and V-s velocity model for the FORGE experiment (Utah, USA).

The results are shown in the left images of Figure(4.4) and Figure(4.5). The wavefronts look similar due to the similar velocity model geometries, but the first arrivals recorded are clearly different as we will see below. The Figures also highlight the position of Well A and Well B which are respectively 1500m and 2500m away from the epicenter of the source.

The FORGE experiment area and Well position (red triangles) are shown in the map in Figure(4.6), along with the source epicenter (blue star). All 3 synthetic seismograms for the 3 Wells have been computed, only the ones from Well A will be shown, without losing the generality of the solution. To test if the generated synthetic seismograms are

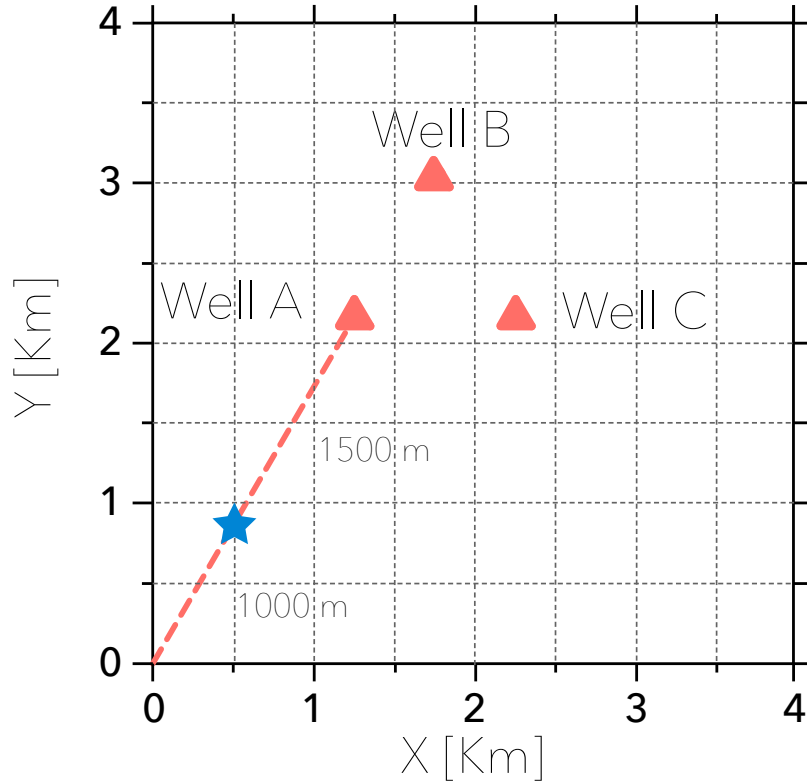


Figure 4.6: stations geometry of Well A, Well B, Well C and source position of the FORGE experiment in Utah (USA).

correct we compared them with the ones simulated by SALVUS (Afanasiev et al. 2019). Initially, we did not include the radiation pattern but only the directivity, to mainly test the correct calculation of the arrival time and the different wavelet convolution. The results are shown in Figure(4.7) and Figure(4.8) and it is clear how the arrival times calculated with the method we developed match the ones computed by SALVUS, considering also the fact that the two wavelets are different. Furthermore, the difference between the P arrivals and the S arrivals $t_p - t_s$ strictly matches in both Figures. the main difference to highlight is how the derivative of the Ricker wavelet better resembles the ones from SALVUS and provides sharp boundaries to more easily pick the first arrivals. The Directivity effect is not clearly visible in this synthetic because the source is far away from the fiber and so the range of the angle of incidence i_h is very narrow.

The last synthetic generated is the one in Figure(4.9) where the source's radiation

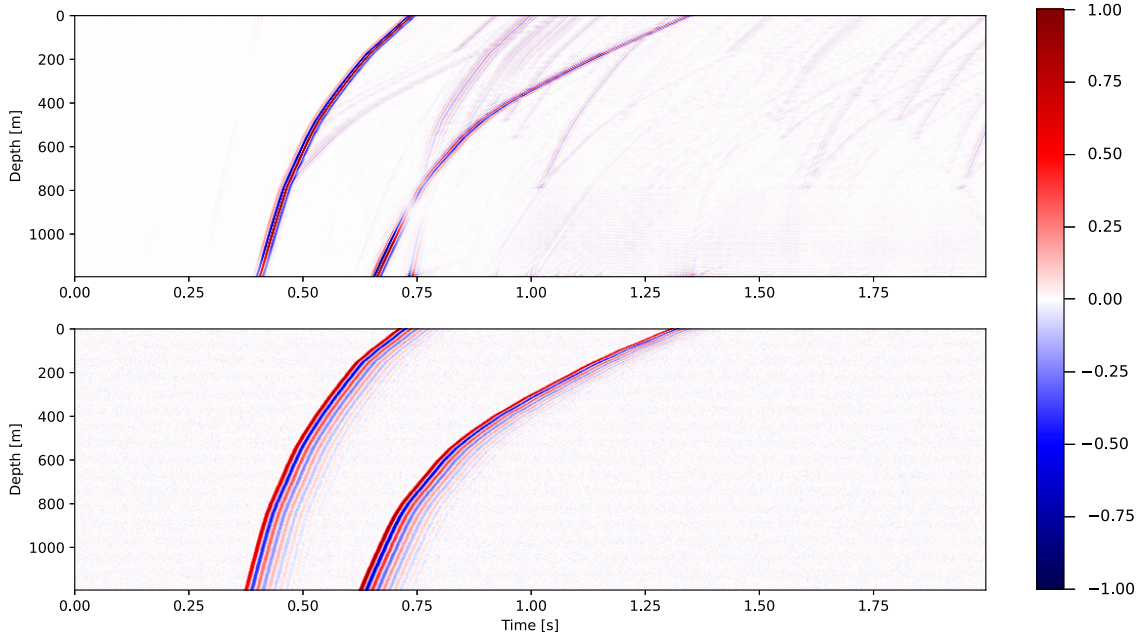


Figure 4.7: Comparison between the simulated seismogram with the exponentially decaying sinusoidal wavelet, and the one computed with SALVUS, for an event at distance $1500m$ and depth $2437m$ from the Well A.

pattern effect is included. The simulated fault has a Strike = 12° , a Dip = 45° , and a Rake = 0° . By substituting those values in the Equations(4.5),(4.6),(4.7) and calculating \mathcal{F}^P for the P waves and $\mathcal{F}^{SH}, \mathcal{F}^{SH}$ for the S waves, we generate a seismogram with both directivity and radiation pattern. The arrival times are again correct, and this time we can see the effect of the radiation pattern on P and S waves. The P wave has a higher amplitude with respect to the S wave, furthermore, the S wave has a polarity shift at around $850m$ depth that is clearly visible also in the simulation computed using SALVUS. The results show a good level of similarity between the two plots.

As a last, more qualitative example, we show a real event recorded in the FORGE experiment. The initial uncut and unfiltered recording has a time window length is $15s$ and several noise is present in the upper part of the seismogram, as shown in Figure (4.10). the event of interest is present in the first $2s$ of the recording. After computing a box-filter in the F-K spectrum to eliminate the noise present at wavenumber = 0 (Figure(4.11)) the resulting seismogram is shown in Figure(4.10). Still, some noise is present but the first P and S arrival are clearly visible and their shapes resemble greatly the ones from the

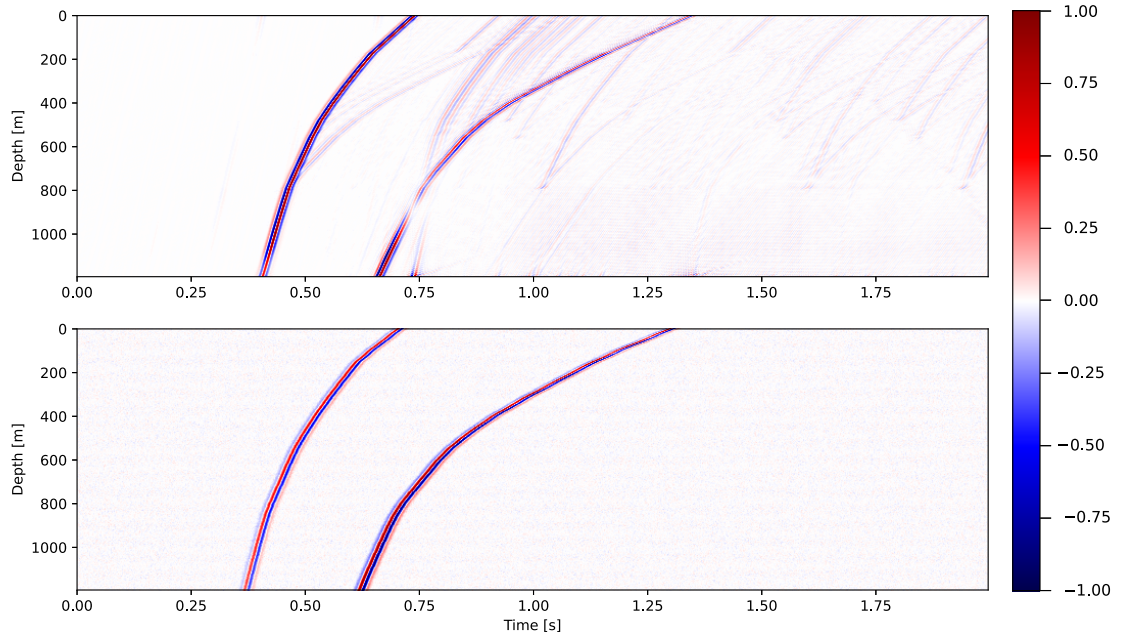


Figure 4.8: Comparison between the simulated seismogram with the derivative of the Ricker wavelet, and the one computed with SALVUS, for an event at distance $1500m$ and depth $2437m$ from the Well A.

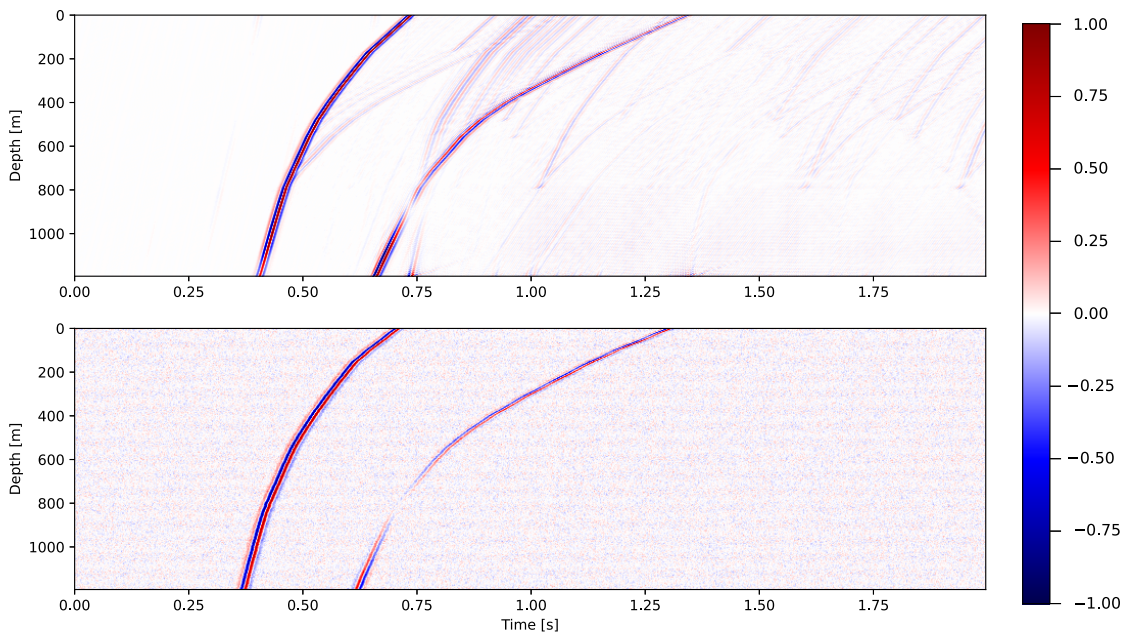


Figure 4.9: Comparison between the simulated seismogram with radiation pattern, and the one computed with SALVUS, for an event at distance $1500m$ and depth $2437m$ from the Well A.

synthetic seismograms.

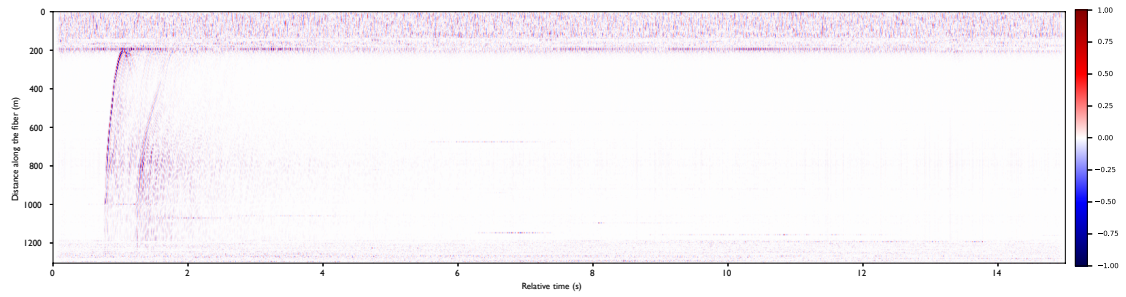


Figure 4.10: Real seismogram recorded in the Forge experiment (Utah, USA). Event present in first 2s.

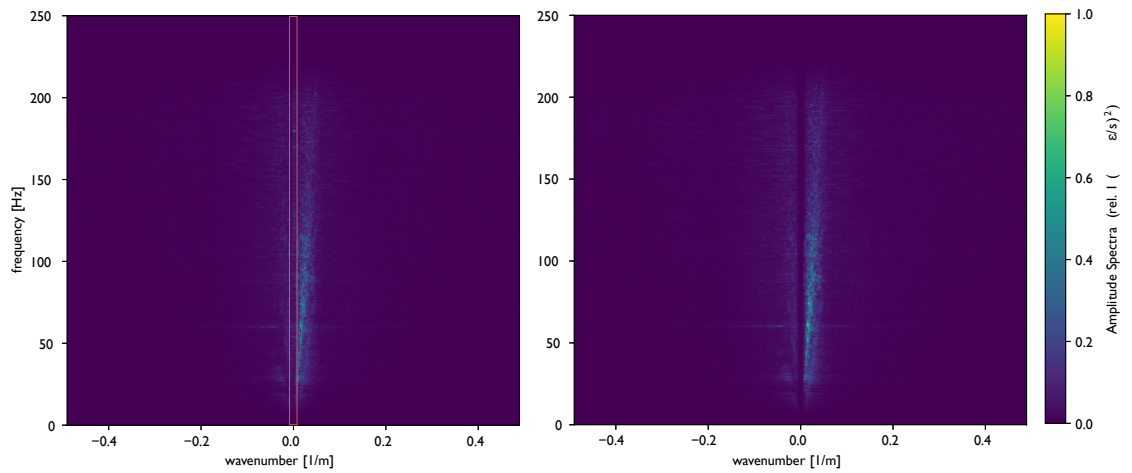


Figure 4.11: F-K spectrum before and after filtering. The box-filter is centered on the 0 value for the wavenumber and comprehends all the frequencies.

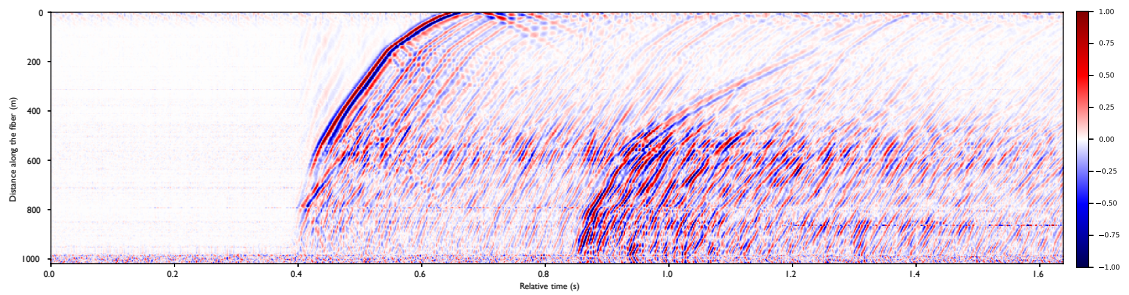


Figure 4.12: Real event recorded in the Forge experiment (Utah, USA)

5. Conclusion

In this study, we developed a tool for the calculation of traveltimes by solving the Eikonal equation in 2D media (or media with cylindrical symmetry, e.g. 1D). The final aim is to implement a fast and efficient workflow for the simulation of DAS microseismic recordings by using the convolutional model. Despite its widespread popularity, DAS still lacks specialized data analysis techniques, and there is a scarcity of publicly available datasets for developing and testing new data processing methods. Furthermore, DAS datasets are exceptionally large, reaching approximately 1 Terabyte of data for just one day of recording over a 1-kilometer length of fiber. To address this challenge, we propose the development of a workflow aimed at constructing first order approximation synthetic datasets that can be used to test new methods or train machine learning models for DAS data analysis.

In the initial section of this thesis, we described the fundamental physical concepts behind Distributed Acoustic Sensing (DAS) acquisition systems. Our focus then shifts to the pivotal contribution of this thesis – the development of a robust Python library for the calculation of seismic travel times in 2D media able to manage arbitrary source-receiver geometries and receivers spacings. This achievement is particularly significant considering the inherent challenges posed by the density of channels and the complex acquisition geometries unique to DAS installations. These challenges are further compounded when the fiber optic cables are deployed within vertical boreholes, a scenario not adequately addressed by existing tools and methodologies. We successfully tested the codes we developed by comparing our results with the "exact" ones calculated analytically for an homogeneous medium. Then we further validated the traveltime calculator we implemented with more complex velocity models.

The final part of this thesis focuses on the implementation of a workflow for synthetic DAS generation. Here, synthetic datasets are generated using the convolutional model including the directivity effect of the DAS and the source radiation pattern and attenuation due to the geometrical spreading for more realistic simulations. We then compared our results with those obtained using the spectral element modeling software Salvus, devel-

oped by Mondaic. Our comparison shows that our workflow allow to correctly reproduce the main kinematic patterns of the seismic wavefield propagating in a 1D velocity model. While a single simulation with our workflow require few seconds in a laptop computer (2 cores and 8 GB of RAM), the spectral element simulation requires few minutes in a multi-core computer (with 18 cores and 64 GB of RAM).

Our Python library for travel time generation is currently designed to operate in serial mode. In the future, our focus will be on implementing "technical" enhancements to increase the computational speed of the Eikonal Solver. This will involve rewriting part of the codes in a low-level language like C and parallelizing it with GPUs, possibly utilizing frameworks like OpenACC. By adopting these optimizations, we aim to significantly enhance computational efficiency, but keeping the input-otuput interface in python, that makes the libary easy to use. This improved efficiency will enable our DAS synthetic simulation workflow for the extensive generation of synthetic datasets that, despite simple, could allow to bridge the gap in available data and enable comprehensive testing and development of new DAS microseismic data analysis methods.

Appendix

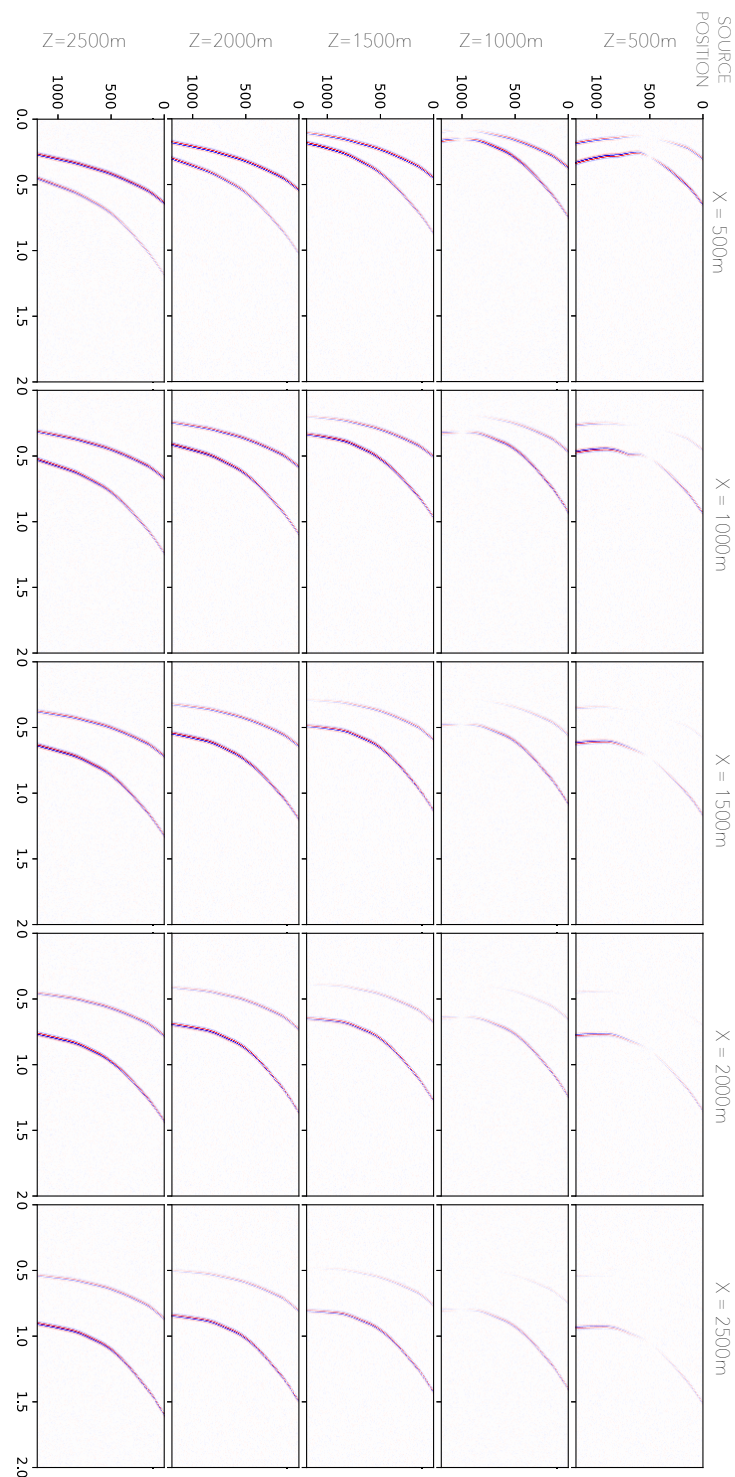


Figure 5.1: Seismograms simulated with Directivity effect, for a fiber deployed vertically and a source positioned between 500m and 2500m (steps of 500m) in depth and between 500m and 2500m in distance (steps of 500m).

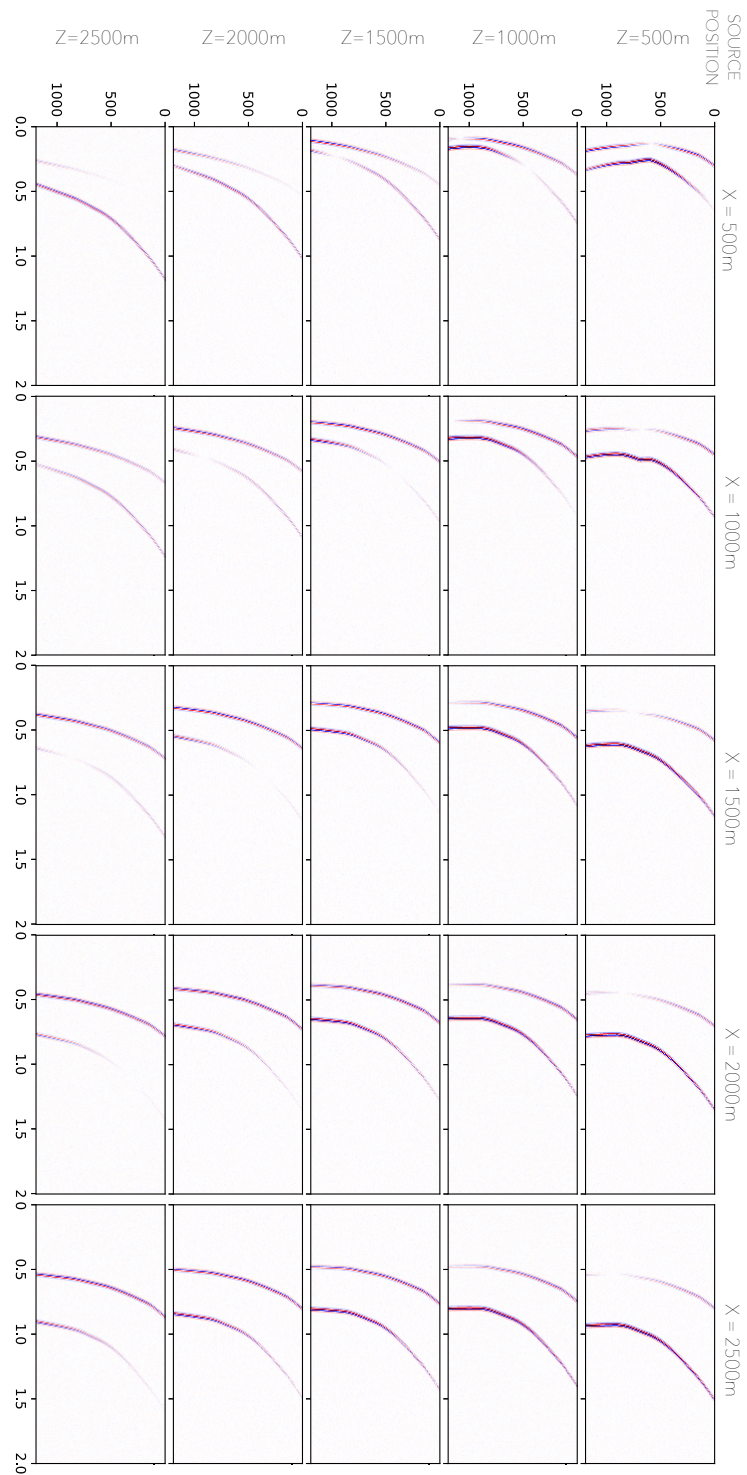


Figure 5.2: Seismograms simulated with Radiation Pattern effect, for a fiber deployed vertically and a source positioned between 500m and 2500m (steps of 500m) in depth and between 500m and 2500m in distance (steps of 500m).

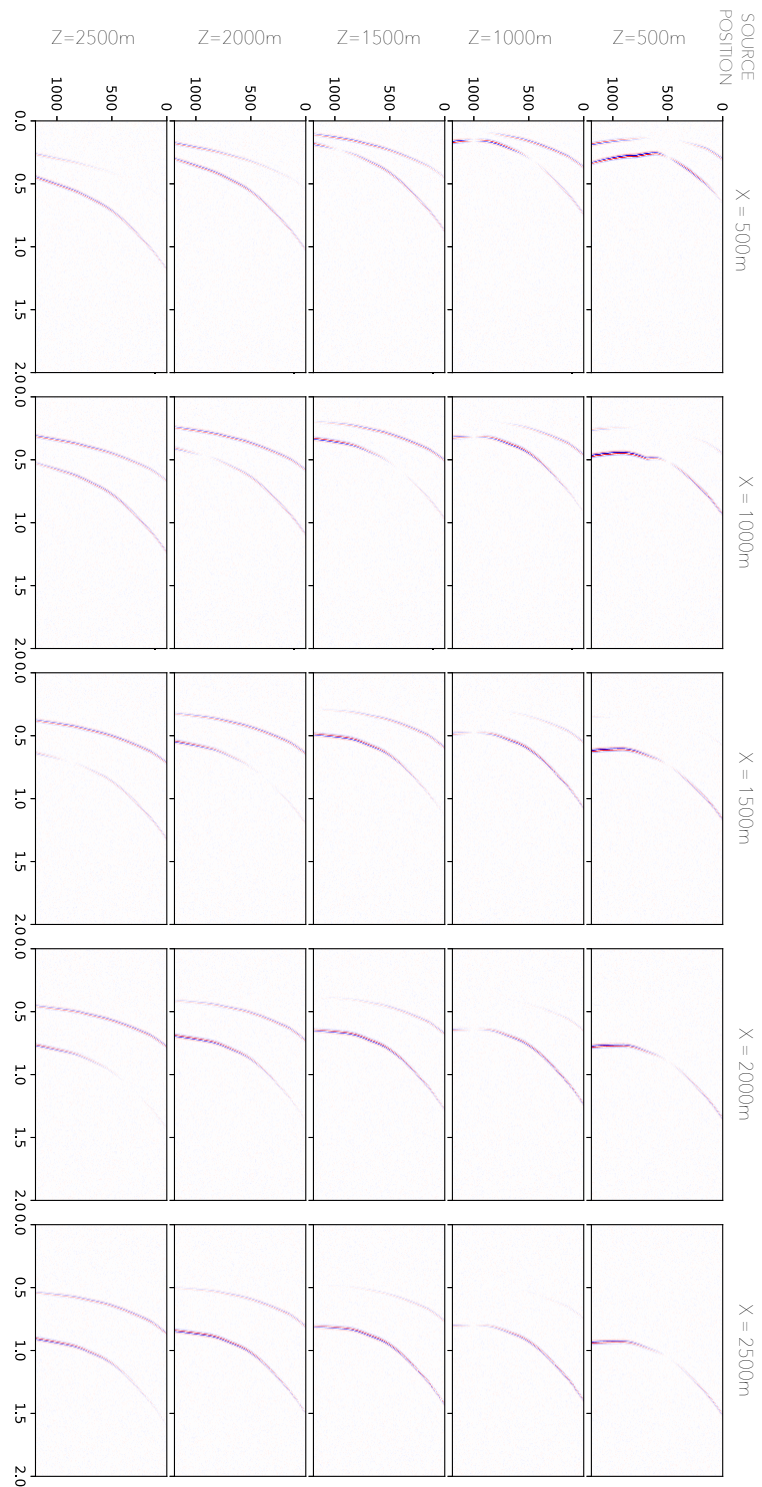


Figure 5.3: Seismograms simulated with Directivity and Radiation Pattern effects, for a fiber deployed vertically and a source positioned between 500m and 2500m (steps of 500m) in depth and between 500m and 2500m in distance (steps of 500m).

Bibliography

- Afanasiev, Michael et al. (2019). “Modular and flexible spectral-element waveform modelling in two and three dimensions”. In: *Geophysical Journal International* 216.3, pp. 1675–1692. DOI: 10.1093/gji/ggy469.
- Aki, Keiiti and Paul G. Richards (2002). *Quantitative seismology*. 2nd ed. Sausalito, Calif: University Science Books. ISBN: 9780935702965.
- Ammon, Charles J et al. (2020). *Foundations of modern global seismology*. Academic Press.
- Benioff, Victor Hugo (1935). “A Linear Strain Seismograph. The Physical Evaluation of Seismic Destructiveness. A Method for the Instrumental Determination of the Extent of Faulting”. PhD thesis. California Institute of Technology.
- Blum, John A, Scott L Nooner, and Mark A Zumberge (2008). “Recording earth strain with optical fibers”. In: *IEEE Sensors Journal* 8.7, pp. 1152–1160.
- Bucaro, JA, HD Dardy, and EF Carome (1977). “Fiber-optic hydrophone”. In: *The Journal of the Acoustical Society of America* 62.5, pp. 1302–1304.
- Cole, JH, RL Johnson, and PG Bhuta (1977). “Fiber-optic detection of sound”. In: *The Journal of the Acoustical Society of America* 62.5, pp. 1136–1138.
- Dean, Tim et al. (2016). “Distributed vibration sensing for seismic acquisition”. In: *The Leading Edge* 35.7, pp. 600–604.
- DeWolf, Scott et al. (2015). “Improved vertical optical fiber borehole strainmeter design for measuring Earth strain”. In: *Review of Scientific Instruments* 86.11.
- Fernández-Ruiz, María R. et al. (Mar. 2020). “Distributed acoustic sensing for seismic activity monitoring”. In: *APL Photonics* 5.3, p. 030901. ISSN: 2378-0967. DOI: 10.1063/1.5139602. eprint: https://pubs.aip.org/aip/app/article-pdf/doi/10.1063/1.5139602/14569690/030901_1_online.pdf. URL: <https://doi.org/10.1063/1.5139602>.
- Hartog, Arthur H (2017). *An introduction to distributed optical fibre sensors*. CRC press.

- Hatfield, William, Frank K Wyatt, and Mark A Zumberge (2018). “Optical Fiber Strainmeters and the Potential for the Detection of Slow Slip Events”. In: *AGU Fall Meeting Abstracts*. Vol. 2018, T51J–0320.
- Lellouch, A. et al. (2021). “Low-Magnitude Seismicity With a Downhole Distributed Acoustic Sensing Array—Examples From the FORGE Geothermal Experiment”. In: *Journal of Geophysical Research: Solid Earth* 126.1. e2020JB020462
2020JB020462, e2020JB020462. DOI:
<https://doi.org/10.1029/2020JB020462>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JB020462>.
URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JB020462>.
- Lindsey, Nathaniel J., Eileen R. Martin, et al. (2017). “Fiber-Optic Network Observations of Earthquake Wavefields”. In: *Geophysical Research Letters* 44.23, pp. 11, 792–11, 799. DOI: <https://doi.org/10.1002/2017GL075722>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017GL075722>.
URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL075722>.
- Lindsey, Nathaniel J., Horst Rademacher, and Jonathan B. Ajo-Franklin (2020). “On the Broadband Instrument Response of Fiber-Optic DAS Arrays”. In: *Journal of Geophysical Research: Solid Earth* 125.2. e2019JB018145 10.1029/2019JB018145, e2019JB018145. DOI: <https://doi.org/10.1029/2019JB018145>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019JB018145>.
URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JB018145>.
- Martin, Taylor and Greg Nash (2019). “Utah FORGE: High-Resolution DAS Microseismic Data from Well 78-32”. In: DOI: 10.15121/1603679. URL: <https://www.osti.gov/biblio/1603679>.
- Mateeva, Albena et al. (2014). “Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling”. In: *Geophysical Prospecting* 62.4-Vertical Seismic Profiling and Microseismicity Frontiers, pp. 679–692.

- Näsholm, Sven Peter et al. (2022). “Array signal processing on distributed acoustic sensing data: Directivity effects in slowness space”. In: *Journal of Geophysical Research: Solid Earth* 127.2, e2021JB023587.
- Okamoto, Katsunari (2021). *Fundamentals of optical waveguides*. Elsevier.
- Piana Agostinetti, N., A. Villa, and G. Saccorotti (2022). “Distributed acoustic sensing as a tool for subsurface mapping and seismic event monitoring: a proof of concept”. In: *Solid Earth* 13.2, pp. 449–468. DOI: 10.5194/se-13-449-2022. URL: <https://se.copernicus.org/articles/13/449/2022/>.
- Podvin, Pascal and Isabelle Lecomte (Apr. 1991). “Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools”. en. In: *Geophysical Journal International* 105.1, pp. 271–284. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.1991.tb03461.x. URL: <https://academic.oup.com/gji/article-lookup/doi/10.1111/j.1365-246X.1991.tb03461.x> (visited on 08/01/2022).
- Senior, John M and M Yousif Jamro (2009). *Optical fiber communications: principles and practice*. Pearson Education.
- Soga, Kenichi and Linqing Luo (Jan. 2018). “Distributed fiber optics sensors for civil engineering infrastructure sensing”. In: *Journal of Structural Integrity and Maintenance* 3, pp. 1–21. DOI: 10.1080/24705314.2018.1426138.
- Vidale, John (Dec. 1988). “Finite-difference calculation of travel times”. In: *Bulletin of the Seismological Society of America* 78.6, pp. 2062–2076. ISSN: 0037-1106. DOI: 10.1785/BSSA0780062062. eprint: <https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/78/6/2062/5334306/bssa0780062062.pdf>. URL: <https://doi.org/10.1785/BSSA0780062062>.
- Zumberge, Mark, Frank Wyatt, and X Yu Dong (1988). “An optical fiber strainmeter and test of its stability”. In: *Optical Fiber Sensors*. Optica Publishing Group, ThCC9.