Development of new methods for non destructive testing for railway infrastructure : Use of statistical signal processing

Quentin Mayolle

CENTRALE LILLE

THÈSE

présentée en vue d'obtenir le grade de

DOCTEUR

Spécialité : Automatique, Génie informatique, Traitement du Signal et des Images

par

Quentin Mayolle

Doctorat délivré par Centrale Lille

Titre de la thèse :

Development of new methods for Non Destructive Testing for Railway Infrastructure. Use of Statistical Signal Processing.

Développement de Nouvelles Méthodes pour l'Évaluation Non-Destructive des Infrastructures Ferroviaires. Utilisation de Traitements Statistiques des Signaux.

Soutenue le 13 juillet 2021 devant le jury d'examen :

Président	Emmanuel Duflos	Professeur	Centrale Lille
Rapporteur	Sriram Narasimhan	Professeur	University of California, Los Angeles
Rapporteur	Christophe Pouet	Professeur	Centrale Marseille
Membre	Giovanni Cascante	Professeur	University of Waterloo
Membre	Audrey Giremus	Maître de conférence	Université de Bordeaux
Membre	Hannes Gräbe	Associate Professor	University of Pretoria
Membre	Sophie Tison	Professeur	Université de Lille
Directeur de thèse	Philippe Vanheeghe	Professeur	Centrale Lille
Invité	Pierre-Marie Rogier		RAILENIUM
Invité	Francis Sabatier		Eurotunnel

Thèse préparée dans le Laboratoire UMR 9189 -CRIStAL
École Doctorale SPI 072

Ce projet est cofinancé par l'Union Européenne au travers du Fonds Européen de Développement Régional (FEDER)

The owl of Minerva spreads its wings only with the coming of the dusk.

Hegel

Acknowledgements

It is a particularly hard task to pay tribute to the numerous persons who shared these three wonderful years. If I take the chronological axis to help me in this task, my first thanks are obviously addressed to the man who initiated this great adventure: my supervisor Philippe Vanheeghe, Professor at Centrale Lille. He not only gave me his trust from the beginning, but acted wisely with patience for me during all these years. I hope sincerely that we will have the chance to continue our investigations for a while. My thanks go next evidently to Giovanni Cascante, Professor at the University of Waterloo (ON, Canada), who transmitted his curiosity for the mechanical fields, invited me to its university, and spend assiduously time with us on atypical subject during the thesis.

I would like to thanks the rapporteurs, Mr Sriram Narasimhan, Professor at the University of California, Los Angeles, and Mr Christophe Pouet, Professor at Centrale Marseille, for spending time to analyse my work. I am grateful to the jury members, Mr Emmanuel Duflos, Professor at Centrale Lille, Audrey Giremus, Maître de conférence at the Université de Bordeaux, Hannes Gräbe, Professor at the University of Pretoria, and Sophie Tison, Professor at the Université de Lille, who accepted to examine and assess my work.

I now have to thank my colleague Nouha Jaoua, for her constant help during these years. She valiantly accepted to answer my relentless questions, and was a great help on every subject. I would like to thank Francis Sabatier, who invited us to Eurotunnel and shared his tremendous knowledge about the railway world. I will surely keep for a very long time the memories of our night operations in the impressive tunnel.

I thanks Pierre-Marie Rogier, who guided me in RAILENIUM, and has been the driving force which conducted our various experiments in the north of France. I do not forget Denovan Lampin, whose good mood has been a great help for our long hours spend on puzzling problems. I have a deep gratitude for Carole Magniez and Moïse Vouters, for making this thesis possible, and for their encouragements. I thanks Alain Rivero for his inspiration about modern inspection techniques, and for sharing its railways stories.

I express my gratefulness for the Sigma team, which has been a fertile ground for exchanges, debates, support and motivation. A small wink to my room-mates Yoan and Markus, and to my precious Solène, who brightened numerous days.

During this thesis I was welcomed 2 months at the non-destructive testing laboratory at the department of Civil and Environmental Engineering at the University of Waterloo (ON, Canada). This visit was funding by the laboratory and RAILENIUM. I thank Mr and Mrs Saccomanno for their hospitality and their kindness. I obviously cannot speak about these months without a sweet thought for the two persons who welcomed me with open arms, Taher and Cristobal, who made me enjoy every day spent in the university.

I salute my parents and my grandmothers for their cherished support during these valuable but sometimes rough years. A special mention goes to my two brothers, Thibaut and François, without whom life would have been definitively less funny. A delicate thought to my grandfathers, who still inspire us from their celestial resting places.

Thanks to my friends, who endured bravely me: Joris, Nelson, Etienne, Rémy, François, Pierre-Yoann and my two Camille(s). May our adventures continue for one more century.

Lastly, I would like to thanks the Hauts-de-France Region, RAILENIUM and Eurotunnel for making this three years thesis possible.

Ce projet est cofinancé par l'Union Européenne au travers du Fonds Européen de Développement Régional (FEDER).

Contents

List of Figures

List of Tables

Notations

Operations

- $\langle f,g \rangle = \int f(t) g^*(t)\, dt$ the inner product of functions f and g.
- $\|f\|^2 = \langle f,f \rangle$ the square norm of function f.
- $(f * g)(t) = \int f(\tau) g(t-\tau)\, d\tau = \int f(t-\tau) g(\tau)\, d\tau$ the convolution between f and g at point τ.
- X^T : denote the transpose of matrix X

Signal Processing notations

- c^* : complex conjugate of a variable c
- f^* : function which is the complex conjugate of a function f
- ν : constant frequency
- $\nu(t)$: instantaneous frequency
- Ψ : mother wavelet function
- $\mathcal{W}_\Psi(s,u)$: wavelet coefficient at scale s and time u with mother wavelet Ψ
- $\mathcal{S}_\Psi(s,u) = |\mathcal{W}_\Psi(s,u)|^2$: scalogram, or energy of the wavelet coefficient at scale s and time u
- $\mathfrak{R}(s(u),u)$: ridge, or set of local maxima of the scalogram, defined for different values of time u, with their corresponding scales $s(u)$

Probabilistic notations

- $\mathbb{N}$: set integers
- $\mathbb{N}^*$: set strictly positive integers
- $\mathbb{R}$: set of real numbers
- $\mathbb{R}^+$: set of positive real numbers
- $\mathbb{R}^{+*}$: set of strictly positive real numbers
- $\mathbb{C}$: set of complex numbers
- $\hat{c}$: estimator of a variable c
- x : state vector
- x_k : state vector at step k

- $x_{1:k}$: concatenation of state vectors $x_1, x_2, .., x_k$
- z : measurement vector
- z_k : measurement vector at step k
- $(z_k)_{k=1}^{N}$: sequence of measurement vectors from step 1 to N
- $z_{1:k}$: concatenation of measurement vectors $z_1, z_2, .., z_k$
- $\mathcal{L}(\theta)$ or $\mathcal{L}_{\mathcal{D}}(\theta)$: likelihood of data $\mathcal{D}$ for parameter θ
- $Q\mathcal{L}(\theta)$ or $Q\mathcal{L}_{\mathcal{D}}(\theta)$: quasi-likelihood of data $\mathcal{D}$ for parameter θ
- $|X|$: cardinality of a RFS X, or number of vectors contains in the set X
- $[\![p, q]\!]$ with $p, q \in \mathbb{N}$ and $p \leq q$: discrete interval containing all the integer between p and q, p and q included

Acronyms

- AC : Alternative current
- BFGS : Broyden–Fletcher–Goldfarb–Shanno
- BO : Bayesian Optimization
- COI : Cone of Influence
- CWT : Continuous Wavelet Transform
- δ-GLMB : delta Generalized Labelled Multi-Bernoulli
- EC : Eddy Current
- EI : Expected Improvement
- EMAT : Electro-Magnetic Acoustic Transducers
- EM : Expectation Maximization
- EMD: Empirical Mode Decomposition
- EWT : Empirical Wavelet Transform
- FISST : Finite Set Statistic Theory
- GM-PHD (filter) : Gaussian Mixture Probability Hypothesis Density (filter)
- GP : Gaussian Processes
- HHT : Hilbert-Huang Transform
- IMF : Intrinsic Mode Function
- LMB : Labelled Multi-Bernoulli
- MAP : Maximum A posteriori

- MCMC : Monte Carlo Markov Chain
- MCWM : Monte Carlo Within Metropolis
- MH : Metropolis-Hastings
- MLE : Maximum Likelihood Estimator
- MP : Matching Pursuit
- MRAE : Mean Relative Absolute Error
- MSE : Mean Square Error
- NDT : Non-Destructive Testing
- NMSE : Normalized Mean Square Error
- NPMMH : Noisy Particle Marginal Metropolis-Hastings
- OSPA : Optimal SubPattern Assignment
- PF : Particle Filter
- PHD : Probability Hypothesis Density
- PMMH : Particle Marginal Metropolis-Hastings
- PRF : Pulse Repetition Frequency
- QML : Quasi-Maximum Likelihood
- RBPF : Rao-Blackwellized Particle Filter
- RCF : Rolling Contact Fatigue
- RFS : Random Finite Set
- RK4 : Runge-Kutta of order 4
- SAGE : Space Alternating Generalized EM
- SDOF : Single Degree of Freedom
- SHM : Structural Health Monitoring
- SNR : Signal to Noise Ratio
- STFT : Short Term Fourier Transform
- ToF : Time of Flight
- US : Ultrasonic
- WVD : Wigner-Ville distribution

Introduction 0

Motivations

Monitoring of railway infrastructures relies on the control of quality and lifespan of equipments. Many devices interact continuously to guaranty the passenger service. The train is travelling on the track, which rests on a specifically adapted ground. Train is also in contact with an electrical alimentation, provided by the catenary. Control and signalization devices are distributed on the rail network to control trains and manage circulation. This phenomenal web of objects made of wood, iron, copper, concrete,... requires a proper maintenance to prevent disturbances of the traffic.

From the previous infrastructures, rails deserve a focused attention. Any damage on rail could evolve into rail break, with dramatic consequences. External environment, train traffic or manufacturing deficiencies are sources of degradation of the rail track. Mechanical stress has a major influence on the development of rail defects. Structural Health Monitoring (SHM) has therefore to perform assessment of the rail health to record, characterize and follow defects through time. A proper knowledge of both the lifespan of material and changes of defects allows an efficient management of the maintenance operations, and a reduction of risk associated to traffic interruption.

Recent years have seen the development of specific train cars equipped with numerous sensors for inspection of rails [DKT16]. Railways infrastructure operations require a reduction of time dedicated to maintenance, and optimization of processes. Measurement trains or trolleys present an interesting perspective for fast assessment of rails, provided that the quality of inspection is preserved. However, suitable interaction of an external track with costly measurement instruments is hardly achieved. To address this major concern, a new methodology for non-contact inspection of rail has been designed by RAILENIUM [Nda16]. Ultrasonic sensors with piezoelectric transducers is a well known technique for rail inspection, which has demonstrated its capability to perform reliable and efficient inspection of the rail [PRD08], but it requires a direct contact with the rail surface. Electro-Magnetic Acoustic Transducers (EMAT) sensors are used to generate a surface wave, called a Rayleigh wave, on the rail. This wave can interact with possible defects. Measurement issues are yet still observable. Rail condition and external environment bring disturbances which deteriorate the inspection quality, interfering with detection of defects. A further step is thus necessary to interpret signals coming from sensors, extract information of interest and estimate the level of rail damage. A deliberately choice have been made in the following work: to operate in the field of statistical signal processing.

The thesis aims at proposing a methodology for sequential processing of vibratory signals. Notably, our work is based on the joint use of sparse decomposition techniques, whose purpose is to extract features from the signal, and sequential filtering techniques, which estimate hidden information from evolution of measurements across time. The full process is summarized in figure 1. An ultrasonic measurement vector z_k obtained at a step k is introduced in the decomposition process, to produce a set of characteristics Z_k. This set is then managed by a sequential filter, which performs association with data from previous measurements (for example z_{k-1}). The produced estimation set of hidden characteristics $\widehat{X}_k$ serves as a new basis for evaluation of the rail condition. An integration of the evolution of ultrasonic signals over time is an innovative approach in the Non-Destructive Testing (NDT) domain.

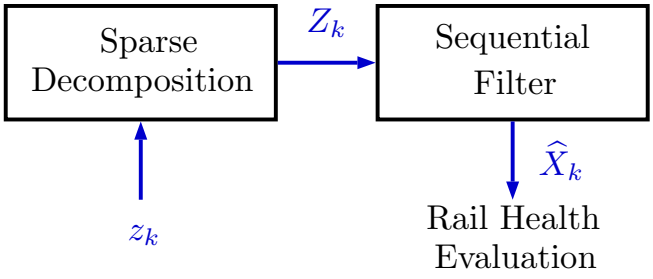

Figure 1: Proposed methodology for rail inspection with EMAT sensors.

Analysis tools from the literature do not express the variability of accessible events inside a signal. The decomposition method allows an expression of one signal as a sum of Gabor functions, which offers an efficient representation of ultrasonic measurements [DS01b]. A set of vectors is obtained from this process. Our analysis of rails are conducted with this specific representation. An important key for inclusion of changes of representations with time are the Random Finite Sets (RFS). They allow to model evolution of sets of objects from accumulation of imprecise measurements. The Finite Set Statistics (FISST) theory developed by Mahler constructed the basis for Bayesian multi-object filters, which express the uncertainty of estimations with probability density functions. The Probability Hypothesis Densitity (PHD) filter [Mah03] proposes a practical approximation of the multi-object posterior to propagate information about sets across time. The delta Generalized Labelled Multi-Bernoulli (δ-GLMB) filter [VV13] is the first Bayesian multi-target tracker, which allows a direct association between estimated states. The present thesis proposes to include this powerful estimation algorithms in the processing of ultrasonic signals.

However, transition from existing research and applications of the RFS filters and trackers is not straightforward. Our original approach is strongly different from available examples of multi-object tracking. Consequently, the model has to be constructed, and all its related parameters defined. For our specific application on EMAT measurements, we propose a new methodology to estimate the parameters of

both the state space model and the δ-GLMB filter. The originality is to combine the filtering process with a probabilistic optimization procedure, which automates the search of parameters. Our complete methodology, from acquisitions to estimations of features with time associations, has been first tested on laboratory to confirm the proposed framework. Field experimentations on active rail tracks have then been conducted to assess the rail health, attesting the validity of the presented methodology.

Organization of the thesis

This thesis is organized as follows:

Chapter 1 presents the specificities or rail damages studied in this thesis. Surface defects which occur at the top of the rail are described, as well as the technologies of the NDT literature employed to detect them. The EMAT technology, with its advantages and drawbacks is outlined. The laboratory devices and rails are next characterized to provide reference measurements for development of our techniques.

Chapter 2 introduces the main theoretical aspects to perform analyses of vibratory and ultrasonic signals. Time, frequency and time-frequency techniques are used to extract relevant information to assess the quality of signals, and compare them. In some cases, a physical interpretation can be made. Application of EMAT signals highlights the potential benefit of these analyses techniques, and difficulties induced by data association, notably with introduction of filtering techniques, based on state space representations.

Chapter 3 investigates the deeper structure of the previous tools, and presents our main inspection methodology to process ultrasonic signals. Advanced filtering techniques with probabilistic formulation allow to design the dynamic evolution of acquisitions, and estimation of physical information related to them. Specific models are introduced, and tested to assess the performances of our processes. Simulations are investigated, to measure the accuracy of our methods, and estimate the related uncertainties. This chapter presents the results of dynamic inspections on the rail tracks of Eurotunnel. Evaluation of performances with moving devices is taken into account. Real defects, coming form damaged rails due to current exploitation of the railways, are processed.

Final chapter 4 investigates a side domain, to generalize the previous presented methodologies. Purpose is to access a higher understanding of vibratory processes, based on non-linear experiments conducted on a resonant-column, a tool for dynamic characterization of soils. A modified version of the single degree of freedom is examined, with a transfer to a state space approach. Non-linearities are caused by variations of the soil characteristics, that we are trying to estimate with filters. Probabilistic methods to take into account the uncertainty related to

parameters of the models are proposed. Results could be useful for more fundamental works about the propagation of perturbations on mediums.

Lastly, a general conclusion is made about the presented work, with investigations of future researches and related topics.

Simulations and results presented in this thesis were conducted through a collaboration partnership between Eurotunnel, RAILENIUM, Centrale Lille and the department of Civil and Environmental Engineering at the University of Waterloo (ON, Canada).

Background on non-contact defects detection on rail surfaces

1

1.1 Introduction

Purpose of this chapter is to introduce the environment of Non-Destructive Testing (NDT) applied on railways infrastructures. It takes an important place in the domain of Structural Health Monitoring (SHM), as an active device to control and manage the full infrastructure. A proper knowledge of degradation of materials with time is mandatory, to extend the lifespan of materials, and improve service quality. The thesis concentrated on measurements related to rail, and a deep understanding of damages affecting this essential component of the railway structure is the basis for efficient inspection and health evaluation. Section 1.2 presents the rail structure, its constituent parts, and related surface defects analysed in this thesis. A brief overview introduces the main existing rail inspection methods. Section 1.3 defines the kind of measurements processed with our methodologies, their specificities, and how they are obtained, with focus on Electro-Magnetic Acoustic Transducers (EMAT), the core non-contact inspection technique developed by RAILENIUM. In section 1.4, the two laboratory rail samples, which serve as baselines, are described. Last section 1.5 underlines the main difficulties faced for analyses of signals coming from ground acquisitions.

1.2 Rail defects

1.2.1 *Descriptive perspective of rails*

Rails are the iron structures which support and lead the train wheels. With ageing, repetition of solicitations over years, loads or occasional shocks, the medium structure can be damaged. Propagation of cracks, if not detected and corrected, can finally cause rail break or pieces detachment. Rail accidents are the major events to be prevented. But even a full rail replacement is costly and time-consuming: it induces traffic delay, displacement and installation of rails, with solicitation of workforces. Early detection is therefore a major interest of today's rail inspection to expend rail lifespans, and optimize the global monitoring of the railway infrastructure management process.

Majority of French railway infrastructures use $UIC\,60$ rails [69]. Their morphology is represented in figure 1.1.

For purpose of clarity, an orthogonal spatial system $(O\vec{x}\,\vec{y}\,\vec{z})$ will be used for all representations of rails. The longitudinal axis, noted $(O\vec{x})$, follows the natural rail orientation. Transversal axis $(O\vec{y})$ and height axis $(O\vec{z})$ are also represented in figure 1.1.

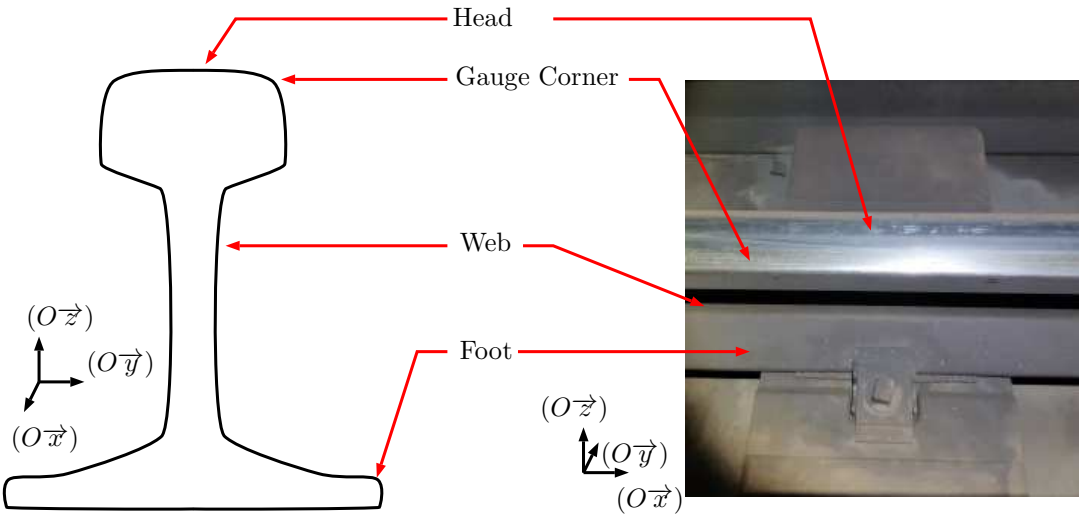

Figure 1.1: Cross sectional view of $UIC\,60$ rail.

The rail foot is fixed to railroad ties (or railway sleepers), which land on the surface, generally made of ballast in external environments. Even if wood is the traditional material for railroad ties, concrete blocks are available for covered areas, such as tunnels. Detection of defects inside the foot is a hard part for current technologies, and still an open problem. For the web part of the rail, a weak area is around the holes, which allow a linking between rails. Strains can induce cracks starting at hole's border, propagating into the web. Head of the rail receives the train wheels. Contact area between the surfaces of rail and wheels is situated between the head centre and the gauge corner, and depends on many physical parameters: height differences between rails, curvature of the tracks, shape of the wheels, distance between rails and wear intensity.

1.2.2 *Surface rail defects*

A distinction is made between internal and surface defects, although a cracks initiated at the surface can penetrate the rail and create severe internal damages. A wide range of surface defects can appear on the head of the rail [02]. Their origins are not clearly identified. They arise in the area of the contact between the wheel of the train and the head of the rail. *Rolling contact fatigue* (RCF) defects are a general class of defects due to the development of excessive shear stress at the contact interface. All the following defects, to be studied in this thesis, belong to this class. *Shelling* or *wheel burn* appear during the fast increase of temperature due to a fast slipping of the wheel. On the contrary, *head-checks* and *squats* develop gradually over months or years: micro-cracks, which initiate at the surface of the rail, grow down and propagate below the surface, and can cause rail breaks. Squats visually look like dark stains at the surface of the rail. Studies have been conducted to understand how they appear, and what are their mechanic characteristics. They are related to the apparition of

a white etching layer, a very hard part due to an adiabatic transformation of the steel, at low temperature and high strain levels, caused by the repetition of micro-slips of the wheel [KWM08]. A common solution to defect development in rail is the surface *grinding* [PRD08]. Small defects are not dangerous by themselves, but their propagation inside the rail head can induce severe cracks. The objective is then to remove a thin layer of rail to eliminate these small defects before their development. This method could also shorten the size of medium cracks, but this will only slow their propagation with time. However, this operation is not neutral for the rail: it will produce some grinding marks, not to be confused with head-checks. Visual differences can be observed on figure 1.2, with localization of grinding mark at the gauge corner, but close to area of head-checking.

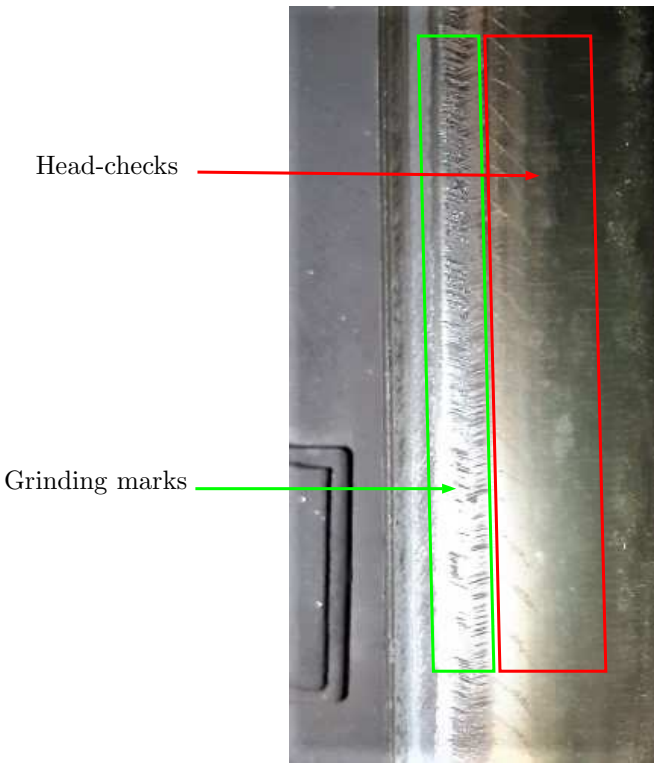

Figure 1.2: Rail sample with both grinding marks (green-left) and head-checks (red-right).

1.2.3 *Inspection techniques*

Before the cracks propagate in the rail, it is therefore necessary to identify micro-cracks at an early stage, to be able to perform maintenance actions, like grinding of the surface, to remove these cracks.

To perform detection of these defects, a wide range of measurements techniques has been proposed in research. The two main techniques are described as:

- Ultrasonic (US) inspection with piezoelectric sensors. It is the most common technique in NDT. The transducer, in contact with the rail, emits a pulse, which propagates in it. A defect will reflect this

pulse, and this reflection will be detected by receiver transducers (or the emitter itself if it works in emitter-receiver mode). The beam orientation in the rail is calculated with the Snell-Descartes law, to reach the values of 0°, 37°, 45° or 70° and detect cracks with various orientations. However, US inspection is unable to detect spot surface defects, and need a perfect contact with the surface to be analysed, most of the time using a couplant. For rail application, it has shown good performances for rail head and web inspection, but poor performances for RCF defects with size $\leq 4\,mm$ [PRD08]. The physical reachable limit is estimated around $60\,km/h$ [SNV17] for mobile inspection. The detectability is reduced when velocity increases.

- Eddy-current (EC) sensors are a non-contact inspection technique. An alternative current (AC) powered coil generates a primary magnetic field, which, by interaction with a ferromagnetic material, induces an Eddy current in it. The circulation of this induced Eddy current in the material will equally induce a secondary magnetic field, which can be measured by a sensor [PRD08]. Any change in the magnetic properties of the material can alter the secondary field, and therefore modify the measurement [DKT16]. Such changes of properties include modification of hardness, cracks or shelling. The sensor detection relies mostly on impedance changes. Amplitude and phase are extracted from the measured complex signal. The intensity of phase changes have been correlated with the crack depths. However, this inspection technique suffers from a serious drawback. The AC current tends to concentrate at the material surface, limiting the possible inspection depth. This effect is known as the *skin effect*. Current density becomes closer to the surface for higher frequencies. The physical quantity which summarizes this phenomenon is the penetration depth δ, calculated as [Son+11]:

$$\delta = \sqrt{\frac{1}{\pi f \mu \sigma}} \tag{1.2.1}$$

with σ and μ the material conductivity and permeability, and f the current frequency. Inside the material, current density is maximal at depth 0, and exponentially decays with depth. It is reduced, at depth δ, to $1/e \approx 0.37$ its value at depth 0, (and 0.05 at depth 3δ). See [GGV11] for details. Higher frequencies produce a concentration near the surface, and lower frequencies explore more volume of the material. Moreover, EC testing is extremely dependent on the distance between the sensors and the medium surface, called the *lift-off*, the material magnetic permittivity and conductivity. During inspection, even complex mechanisms could not prevent lift-off changes, due to the track geometry. Skin effect brings huge limitations, reducing the material inspected depth. Lower frequencies would be necessary to inspect deeper the material, producing less clear signal and higher sensitivity to disturbances. Lastly, as noted in [Raj+18], even for shallow cracks, EC testing tends to overestimate the depth of the cracks, and the estimation of size is not

Environment of measurement impacts hugely the performances of EC methods. In wire and tubes inspection, probes are static (encircling coil probes), the inspected piece can be moved with high precision: as a results, speed up to $150\,m/s$ are reached. We can clearly see the difference with rail inspection, where lift-off (distance between coil or sensor and the inspected surface) cannot always be constant, and mechanical disturbances occur.

reliable.

Limitations of US and EC inspection, especially for surface defect inspections, has conducted to development of new sensors dedicated to low depths inspection.

1.3 Ultrasonic signals

1.3.1 *Pressure, Shear and Rayleigh waves*

In the case of NDT analysis, the 3 principal types of waves which propagate in mediums are the *pressure* waves (P-waves), the *shear* waves (S-waves) and the Rayleigh waves. The first kind, P-wave, is also known as the *compression* wave. Indeed, in this case, the direction of displacement of particles in the medium is the same as the direction of the propagation of the wave. This wave is the fastest of the three kinds, and is thus the first recorded by the measurement devices. For the S-waves, the direction of displacement of the particles is orthogonal to the direction of propagation of the wave. The last type, the Rayleigh wave, is also known as the *surface* wave. It is the slowest wave compared with the two previous described. Its specificity is that the particles move in a cyclic motion. The intensity of this move is reduced with depth. This is the reason why this wave propagates at the surface of mediums. Figure 1.3 shows a graphical representation of waves with the related vibrations of mediums. For pressure, shear and Rayleigh waves, their respective wavelengths are noted λ_P, λ_S and λ_R.

A second type of surface wave is the *Love* wave, which produce, on the surface plane, displacements orthogonal to the direction of propagation.

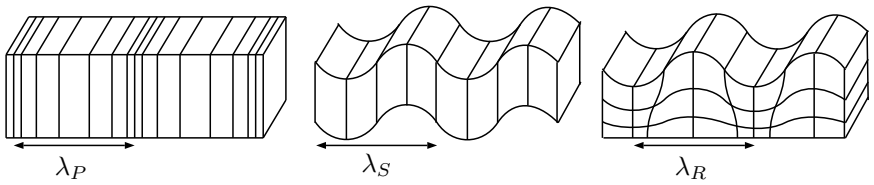

Figure 1.3: Wave propagation of Pressure (left), Shear (middle) and Rayleigh (right) waves on a medium.

The device used in this thesis, EMAT, has the particularity of generating Rayleigh waves. This is of great interest in the research of surface defects: Rayleigh waves will indeed directly interact with cracks and singularities. In a homogeneous medium, Rayleigh waves do not disperse.

1.3.2 *Electro-Magnetic Acoustic Transducers*

EMAT are a technique to generate ultrasonic waves in a ferromagnetic material, with magnetic devices. A coil induces current in the material, that interacts with a permanent magnet to produce a Lorentz force. This disturbance is transferred to the structure of the material, producing a wave that propagates in the material [IHF16]. Since the wave is directly generated in the material, this technique allows contact-less control of rail: contrary to ultrasonic piezoelectric transducers, no couplant is necessary. Furthermore, this technique can

This is the reason why EMAT are also called a *air-coupled* inspection technique.

be used to generate waves that will only propagate at the surface of medium: Rayleigh waves. Nevertheless, this technology has one huge drawback: the amplitude of the signal we get from the receiver is highly dependent on the distance between the surface tested and the transducer (also called the lift-off). This techniques has already been tried with rail application, and showed good results [Nda16].

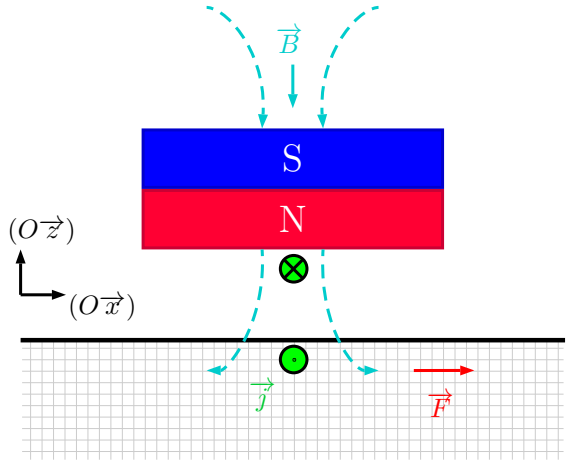

Figure 1.4: Electro-Magnetic Acoustic Transducer with single wire for force generation in a ferromagnetic medium.

Figure 1.4 represents the simplest configuration with one magnet creating an external magnetic field $\vec{B}$ and a wire inducing a current $\vec{j}$ inside the material. The interaction creates a Lorentz force $\vec{F}$, with vectorial expression defined as:

$$\vec{F} = \vec{j} \times \vec{B} \tag{1.3.1}$$

The notation $a \times b$ denotes the vector cross product between vectors a and b.

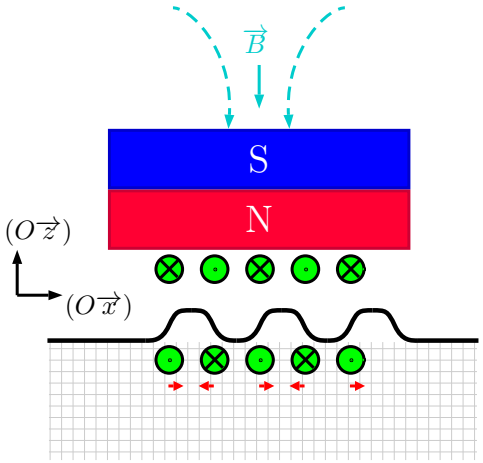

Figure 1.5: Electro-Magnetic Acoustic Transducer with meander coil for wave generation in a ferromagnetic medium.

To generate sinusoidal perturbations, a more complex configuration, using a meander coil, is described in figure 1.5. In this situation, a sequence of opposite forces are created. Depending on the coil, ultrasonic waves can be generated in the two directions (according to axis $(O\vec{x})$).

Since a Rayleigh wave is generated, defects on the propagation path will disturb the wave. The major interest of EMAT is the focus made

only on surface defects. Phenomena including absorption of the wave energy by the defects, reflections or modification of the frequency content of the pulse, provide information about the location of damages and their characteristics [Nda16].

In this thesis, the same configuration of the EMAT inspection device is kept. Sensors generate Rayleigh waves with frequency fixed at $586\,kHz$ and wavelength of $5.08\,mm$ in steel. Ultrasonic pulses are generated with impulse input signals sent by a generator. Sampling frequency f_e is set to $25\,MHz$ (sampling interval of $0.04\,\mu s$). Amplitude of signals are arbitrary units, and this variable will therefore be expressed in *u.a.* in the thesis.

1.3.3 *Sensor characterization*

Comparisons of ultrasonic sensors highlight many practical differences between piezoelectric and EMAT devices. Notably, the non-contact measurement aspect change perceptions and analyses of signals. A proper knowledge of the behaviours of sensors is critical to adapt detection tools to these specific signals.

Main physical parameters of sensors are described as:

- Lift-off issues: with an important sensitivity to variations of the distance between the sensor coil and the inspected medium surface;
- Orientation of the EMAT sensors, in particular the relative angle formed by the direction of emitter and receiver;
- Attenuation during wave propagation in the medium.

To characterize the sensors used in this thesis, laboratory experiments are conducted on a clean rail surface. Signals are averaged over at least 50 acquisitions.

Lift-off variations have a large influence on the maximum intensity of the measured signal. This phenomenon is well-known in the literature: linear variation of the lift-off causes exponential decrease of the signal intensity [Hua+09]. Therefore, detachment of an EMAT sensor from its nominal position will not cause a lack of measure, as in ultrasonic testing with piezoelectric sensors with wedges, but will induce a variation of the pulse's amplitude. Defect detection with EMAT must not rely only on amplitude changes if the lift-off is not assumed constant.

Figure 1.6 illustrates the variation of amplitude with lift-off, on a clean rail portion, with steps of $0.375\,mm$. Exponential fitting confirms the decay rate.

Angle variations between emitter and receiver has also been demonstrated as a relevant parameter. Figure 1.7 underlines the short admissible interval of angles to keep enough high amplitude signal. The directivity diagram indicates the need to maintain admissible variations of the angle of $\pm 10°$ angles to keep a maximum attenuation of 20 percent of the original signal.

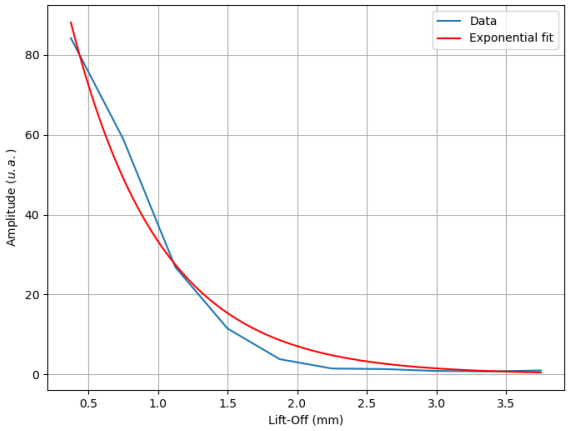

Figure 1.6: Variations of amplitude with change of lift-off.

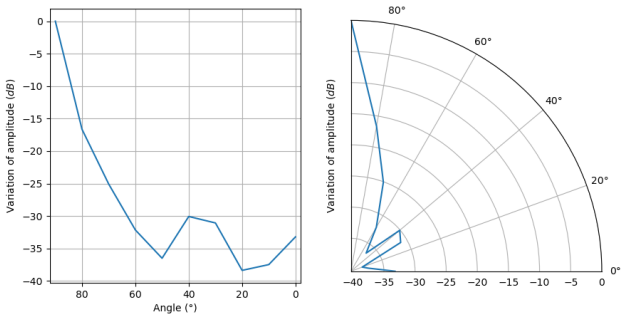

Figure 1.7: Directivity diagram: variations of the signal amplitude as a function of the angle between the receiver and the longitudinal axis: emitter and receiver are parallel at 90°, and orthogonal at 0° (attenuation in dB).

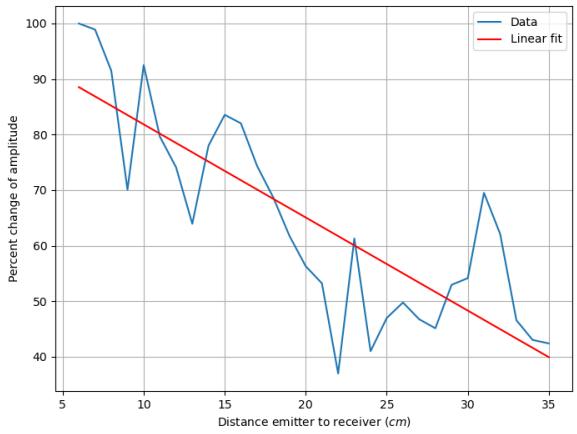

Figure 1.8: Relative variations of amplitude with change of distance between emitter and receiver (percent of change from the maximum amplitude).

The last measurement parameter, distance between the two transducers, has less importance than the previous ones. Attenuation is found to be linear, or at least exponential with slow decrease rate. Figure 1.8 highlights however local variations of the trend.

1.3.4 *Dynamic acquisition devices for ultrasonic inspection*

This section introduces a technology developed by RAILENIUM, to perform online inspection of a rail with EMAT sensors. The sensors are mounted on a structure which can move on the rail, allowing a dynamic inspection. Major goal is to achieve high speed inspection velocities.

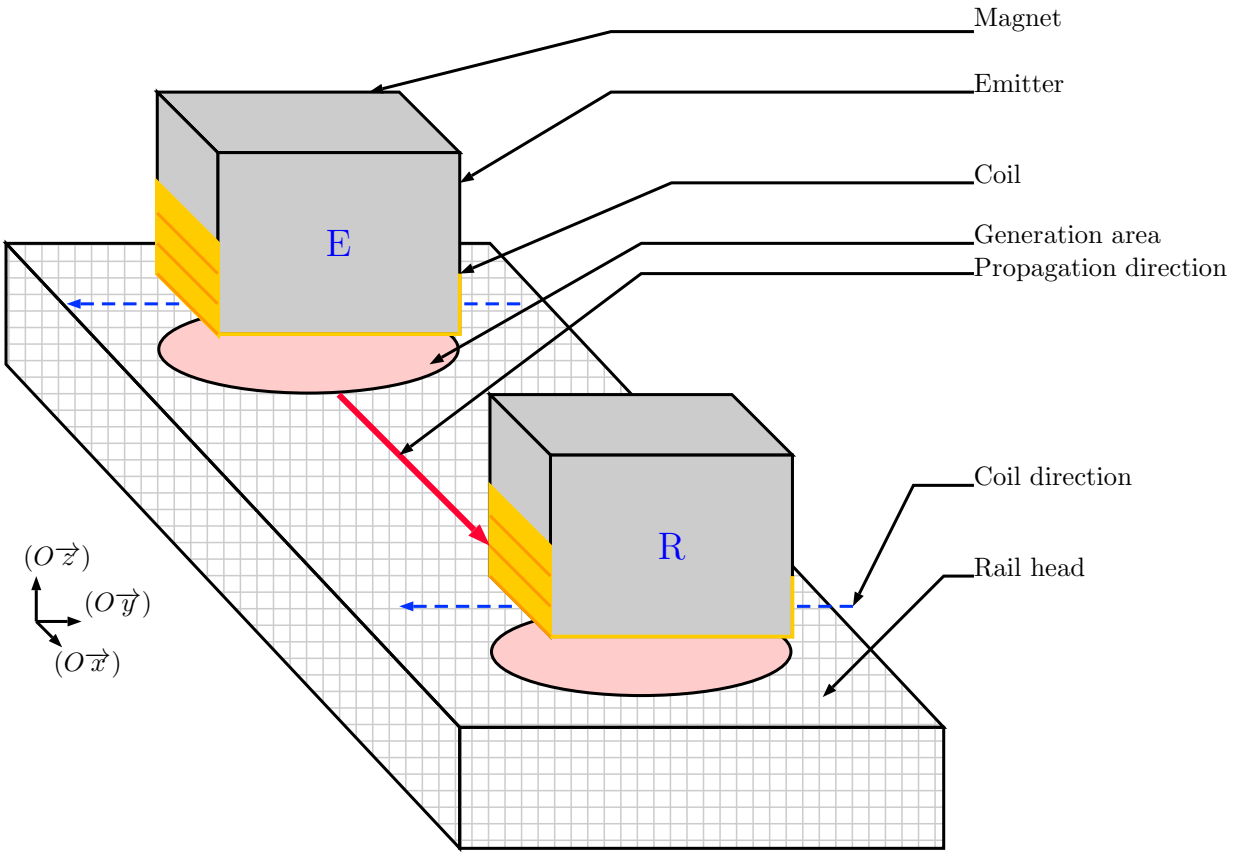

Figure 1.9: Electro-Magnetic Acoustic Transducer device for dynamic non-contact inspection on rails.

Direction of the coils, or equivalently direction of each wire which compose it, directs the direction of the current induced in the rail, and therefore the direction of propagation of the Rayleigh wave. Directions of the two coils (for emitter and receiver), are parallel to maximize the generation and reception efficiencies. In addition, since the meander coil induces several currents in the rail surface, the area of generation of the wave is of the same length order. This imply first that the generation precision is low, in comparison with classical ultrasonic sensors. Next, perturbation inside the medium is not unidirectional:

two ultrasonic pulses can be generated on the medium, propagating in opposite senses. Depending on the shape of the rail, proper reception of the signal can be disturbed.

The main inspection disposition was originally designed to propagate the wave in the longitudinal direction of the rail. A laboratory configuration is shown in figure 1.10.

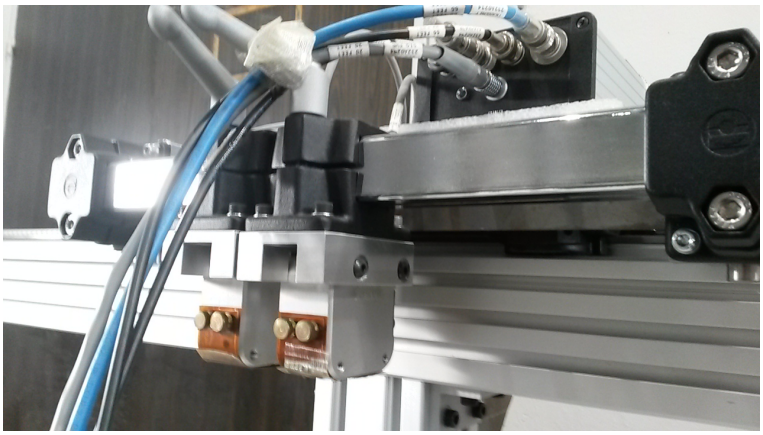

Figure 1.10: RAILENIUM laboratory configuration for dynamic inspection of rails.

The rate of acquisitions performed by the sensor is measured by the *Pulse Repetition Frequency* (PRF), which controls the number of pulses generated by the emitter. One pulse will produce one measurement signal. Range of the used PRF is between 1 and 100 (units are in acquisitions per second, but will be omitted).

1.3.5 *Ultrasonic signal representation*

A *A-scan* is the representation of an acquisition at a specific location of a medium. It traduces the reception of ultrasonic pulses by the receiver sensor [MPP03]. The y-axis measures the displacement of the medium induced by the pulse. Propagation of the pulse is represented by the variation on this axis with displacement on the x-axis, generally the time axis.

The A-scan is mostly found with variation of the medium displacement with time, but it can also have an x-axis related to a different displacement in space.

Signals obtained from EMAT sensors allow a A-scan representation with measurement of amplitude with time. Figure 1.11 shows a typical example of recorded signal. One specificity of EMAT signals are the presence, at the beginning of the A-scan, of a high amplitude event, called the *electrical* signal. This part is indeed related to the generation of the ultrasonic pulse inside the medium. The external magnetic field and the coil of the emitter interact with the receiver, producing a record. However, there is no useful information in this part, and the signal is usually translated, to have origin after a total attenuation of the electrical signal. To prevent a confusion between this content and further reception of pulses propagating in the medium, attention must be paid during the configuration of sensors, to maintain a minimum distance between emitters and receivers.

A sequence of acquisitions made at different positions forms a set of A-scans called a *B-scan*. A heat-map, or a $2D$ plot of this set allows

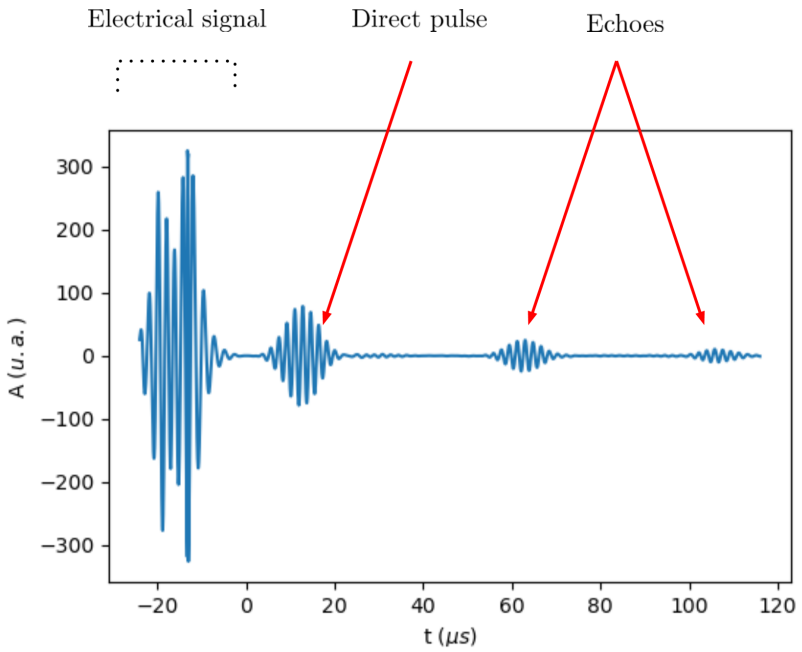

Figure 1.11: Complete A-scan representation of an EMAT signal.

visual detection of changes on A-scans, and diagnostic of the medium areas related to them.

NDT analysis includes two principal steps: detection and characterization of defects. The last step is generally related to a classification on the type of the defects, localization, or an estimation of its geometry (size, depth, severity, orientation). Numerous methods of Signal Processing have been tried, such as frequency analysis [Jia+07], multi-scale analysis [Kub+11]. Use of Gabor wavelets analysis methods showed good results for study of guided waves in steel pipes and plates [XGY09]. Detection and characterization widely use, now, machine learning techniques, for their ability to generate models from data coming from experiments or simulations: Linear Discriminant Analysis [Qiu+14] or Support Vector Machines for detection [Ben04; STF17], and Artificial Neural Networks for estimation of crack depths [Gar+18]. Civil Engineering fields still use empirical methodologies to assess the presence of defects, using several characteristics of signals. The goal is to define a *damage index*: a score, generally between 0 and 1, which indicates the severity of damages.

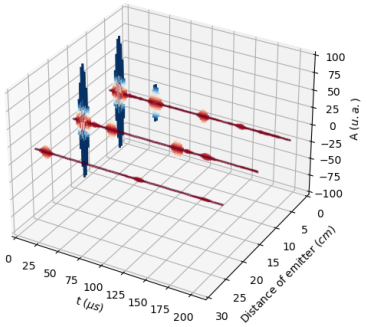

Figure 1.12: Three A-scan for different locations of sensors of the rail: red colour related to low amplitudes, and blue colour to high amplitudes.

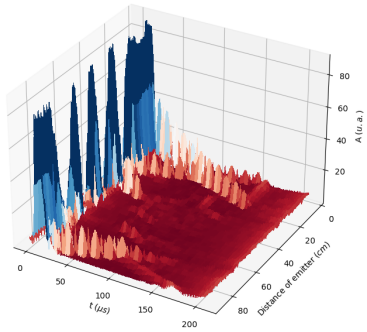

Figure 1.13: 3D envelope representation of signals recorded on a rail, with 1 acquisition per *cm*.

1.4 Simulation of defects

Laboratory experiments are conducted on two rail samples, with machined defects aiming at reproducing squats and head-checks, depending on their orientation, depth and inclination. Real defects are indeed rare, and sometimes far form each other. To provide a compact length of material to be analysed in laboratory, with known defects characteristics, controlled mechanical damages are an efficient way to test sensors and detection methods.

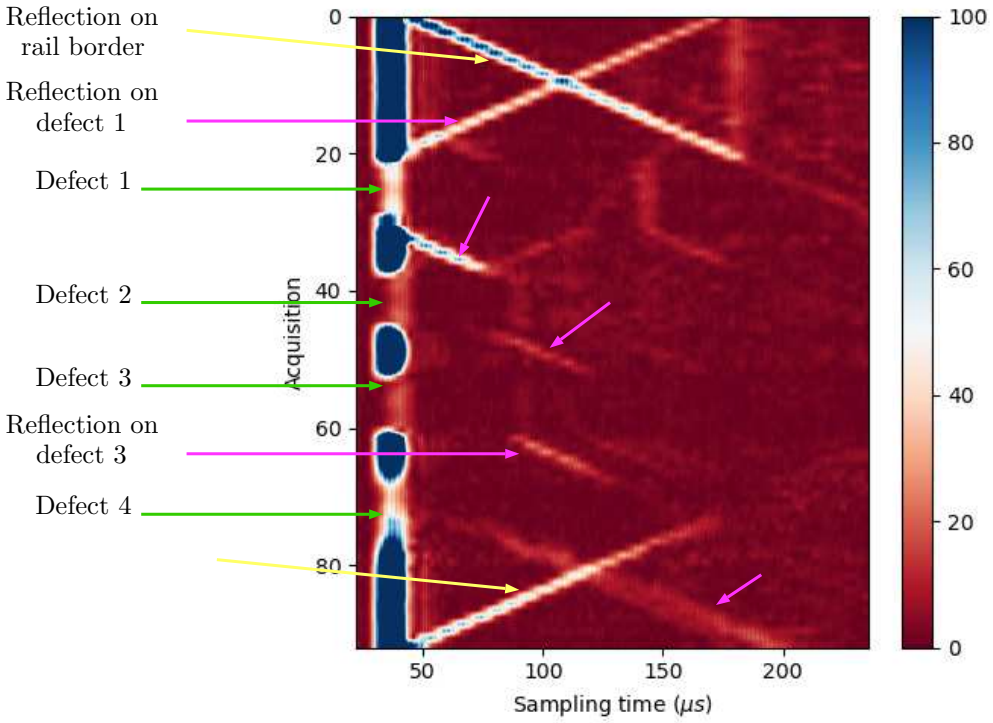

Figure 1.14: B-scan representation of EMAT signals on the $1\,m$ rail, with one measurement per cm. Green arrows target areas of attenuations because of defects. Pink arrows target reflections on defects. Yellow arrows target reflections on rail borders

The first sample, with total length of $1.03\,m$, contains damages with thickness of $1.5\,mm$. First one represents a squat, with direction in the transversal axis of the rail ($O\overrightarrow{y}$). The three other defects, simulating head-checks, have inclinations with respect to the transverse and longitudinal axes. The complete characteristics are gathered in table 1.1. For a displacement of the EMAT sensors on the longitudinal axis of the rail, the best performances of detection are expected for the first defect, and the worst for the last one with higher inclinations. This sample will be related as the *machined* rail, or the $1\,m$ rail (for simplifications).

	Machined defect			
	Defect 1	Defect 2	Defect 3	Defect 4
Depth (mm)	4	8	7	4.1
Thickness (mm)	1.5	1.5	1.5	1.5
Angle on ($O\overrightarrow{x}\,\overrightarrow{y}$)	0°	45°	45°	70°
Angle on ($O\overrightarrow{x}\,\overrightarrow{z}$)	0°	0°	45°	70°

Table 1.1: Characteristics of machined defects on the $1\,m$ rail

Figure 1.14 represents the full B-scan gathering all acquisitions. In areas of defects (when a defect is between the emitter and the receiver), the signal is clearly attenuated. In addition, the B-scan reveals phenomena of reflections on rail borders and defects, producing diagonal lines. Figure 1.15 gathers the rail top and side cross views.

The second sample (figure 1.16), with length of $70\,cm$, has thinner defects, produced with an electro-erosion device. Such damages are

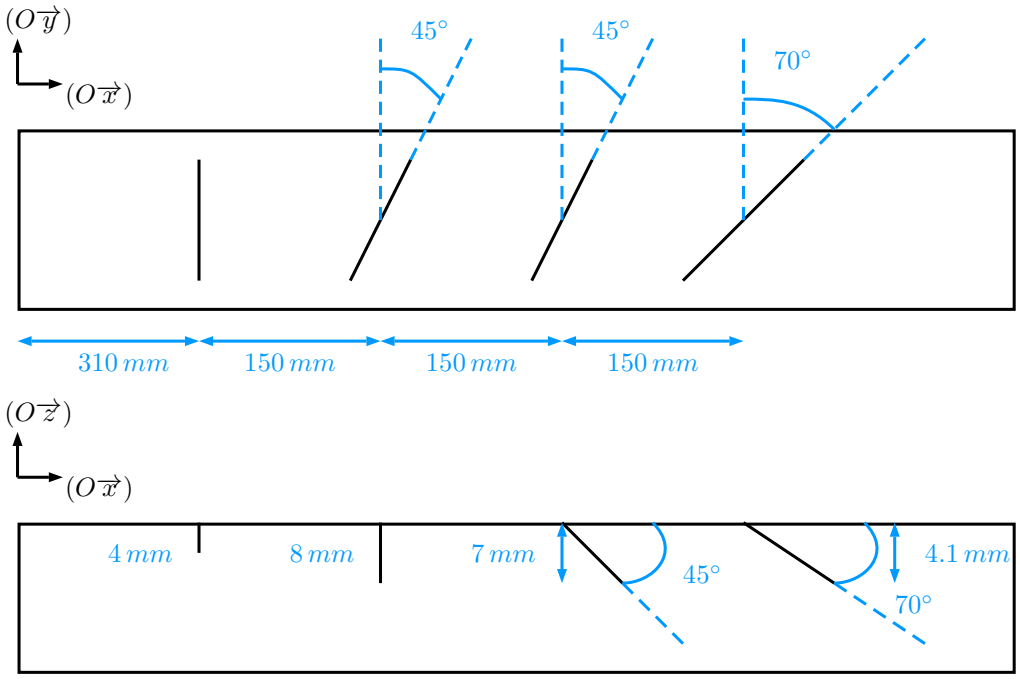

Figure 1.15: Top view and cross longitudinal views of the 1 m rail with machined defects.

closer to realistic situations. Four defects have been machined, with characteristics gathered in table 1.2. A schematic representation of defects is available on figure 1.17. This sample will be related as the *electro-erosion* rail.

Figure 1.16: The 70 cm rail with electro-erosion defects.

Online inspection requires a cheap evaluation of the rail health based on a current measurement. Previous automatic detection schemes, using thresholds or machine learning algorithms, proceed independent analyses of acquisitions. These processes are thus memoryless. On the contrary, visual inspection of signals, notably with B-scans, allows a human operator to detect several changes on acquisitions. Therefore, implicit links between measurements could bring further information for decision of the inspected rail portion.

	Electro-erosion defect			
	Defect 1	Defect 2	Defect 3	Defect 4
Depth (mm)	3	3	6	6
Thickness (mm)	0.4	0.4	0.4	0.4
Angle on $(O\vec{x}\,\vec{y})$	0°	20°	20°	0°
Angle on $(O\vec{x}\,\vec{z})$	0°	0°	0°	0°

Table 1.2: Characteristics of electro-erosion defects on a $70\,cm$ rail

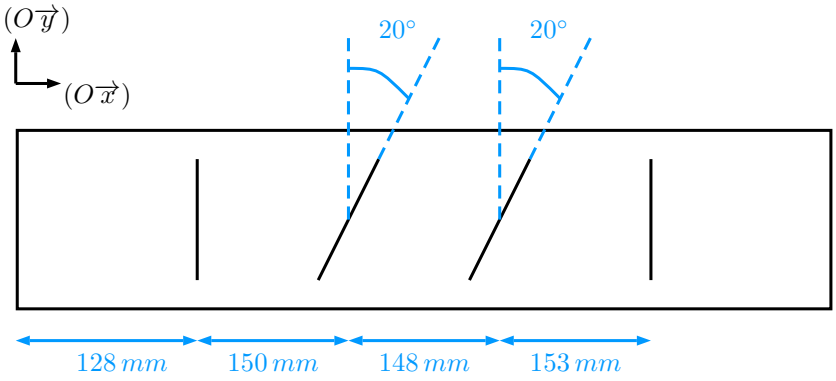

Figure 1.17: Top view of the $70\,cm$ rail with electro-erosion defects.

1.5 Practical considerations

Conducting inspection in external environments induces perturbations of the measurements. The surface condition of the rail is easily degraded by, for instance, dirt, exposition to weather events, chemical reactions or shape of the track. It is therefore necessary to include these aspects in both the measurement and analysis processes. Some important issues, such as variations of the lift-off or lateral displacements, must be addressed by mechanical solutions. Other issues can be partially covered by the signal processing which follows acquisition.

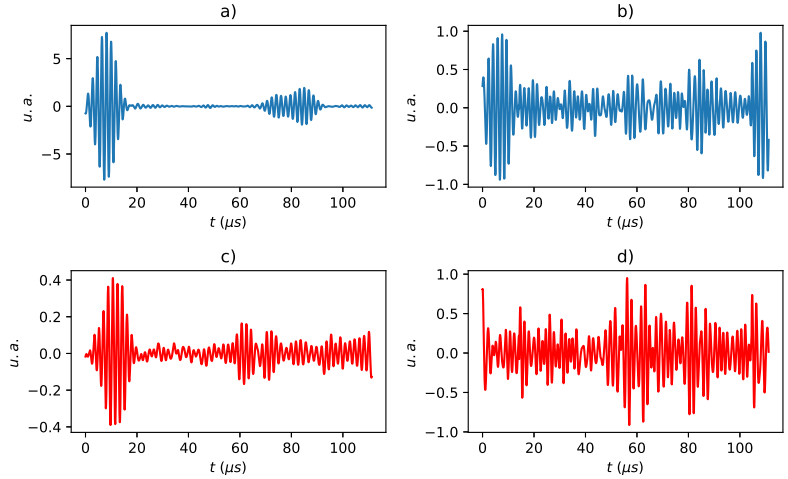

Figure 1.18: Example of noise corruption on healthy and damaged signal with dynamic acquisitions: *a*): healthy clean, *b*): defect clean, *c*): healthy noisy, *d*): defect noisy.

Several problems have been raised form early experimentations with EMAT sensors [Nda16]:

- Noise corruption, which have various sources: electromagnetic disturbances caused by other electric devices, lack of robustness of the measurement system or counter-effects due to displacement of a magnetic device over a ferromagnetic medium (here the rail). Figure 1.18 illustrates this purpose, with a comparison of static and dynamic signals, with measurements from a heathy rail area and from a damaged area. Noisy perturbations make distinction between healthy and defect signal harder. This could disturb further decision processes, which could interpret a noisy healthy measurement as a defect one. Even worse, in high noise situations (low SNR), a damage measurement could be evaluated as healthy, making the related defect invisible to the operator.

The Signal to Noise Ratio (SNR) is defined as the ratio of the power of a clean signal to the power of background noise. It is generally expressed in dB. Low SNR indicates low signal quality

- Saturation: the measurement sensors are limited to a specific amplitude windows, related to a gain. The operator has control on this measurement parameter, and can adapt it to current measurement conditions: if the signal amplitude has decreased because of poor surface condition of the rail, for example. However, the opposite case is equally troublesome: the pulse amplitude can exceed the sensor limits, and cause a phenomenon called saturation. The recorded signal is sliced, and a part of the information is lost for the operator. The real amplitude, which is a feature of interest, is therefore hidden.

Processing techniques must therefore manage these identified issues, to maintain an efficient inspection of rails.

1.6 Conclusion

This chapter introduced the framework of the thesis. Rail health monitoring focus on a specific range of damages, the surface defects which emerge on the head or edges of the rail. From the presented ultrasonic measurement systems, EMAT sensors, a non-contact inspection techniques recently adapted to rails by RAILENIUM, are the promising technology which provided the data processed in the thesis. Analysis of generated Rayleigh waves, which propagate on rail surfaces, is the key aspect for improvement of detection capabilities of measurement devices. Laboratory rails serve as a basis for understanding of vibratory phenomena which happen inside the rail. They provide known and machined defects, allowing an adjustment of inspection techniques to maximize detection performances. On the basis of these data, the main objective of this thesis is to construct and test statistical tools to combine our physical understanding of wave propagation and statistical models to enhance rail inspection.

Statistical signal processing for ultrasonic measurements analysis

2

Non-destructive evaluation of damages on a material employs sometimes complex sensors to bring information about the properties of the inspected object. Structural Health Monitoring (SHM) raised interest among diversified topics, from building inspection to car testing. The spectrum of technologies dedicated to these tasks is large enough to cover nearly all existing possible emission and reception tools. To perform the analysis of information obtained by these sensors, Signal Processing has been from the start an important topic, particularly for ultrasonic testing, in which the recorded outputs are often not directly interpretable for a human operator. Yet, much information is available, after some operations. The purpose of this chapter is first to present the methodologies of signal analysis selected during this thesis. Probabilistic data association is next investigated to complete them.

Section 2.1 highlights the most relevant techniques to extract characteristic from our ultrasonic signal. They are intended to give some answers to the issues encountered during the acquisition.

An important basis resulting from the previous tools is the ability of our signals to be decomposed into distinct elements. Section 2.2 presents how this operation is performed.

Last section 2.3, based on the two previous sections, answers the question of association between acquisitions: how the different characteristics are propagated to bring new items to assess the health of rails.

2.1 Physical features extraction

2.1.1 *Time of flight estimation*

Damage assessment in materials is classically performed using the pulse propagation velocity V_P, defined as, for a pulse-receiver system:

$$V_P = \frac{L}{T} \tag{2.1.1}$$

with L the distance between the emitter and receiver centres, and T the arrival time of the pulse, often referred as the Time of Flight (ToF). Though this definition is commonly encountered in the NDT literature, determination of the arrival time depends on the medium, and type of emitted wave. In ultrasonic testing, techniques rely on identification of a local maxima inside the signal.

Envelope detection allows identification of a maxima in a sinusoidal signal. For a real measurement vector $y(t)$, it finds a decomposition $y(t) = A(t)\cos(\phi(t))$, with $A(t)$ the *instantaneous amplitude* or *envelope*, slowly varying in comparison with the *phase* $\phi(t)$. Following condition about their first derivatives $A'(t)$ and $\phi'(t)$ is generally as-

sumed: $A'(t) \ll \phi'(t)$. Feature of interest are the maximum of the envelope, and time at which it occurs.

The analytical signal $y_a(t)$ can be constructed as the sum of its real part $y(t)$ and an imaginary part such that:

$$y_a(t) = y(t) + i\mathcal{H}_y(t) \tag{2.1.2}$$

The operator $\mathcal{H} : y \mapsto \mathcal{H}_y$ is the Hilbert transform, which acts as a reconstructor of the imaginary part of any real signal. The definition of the operator is:

$$\mathcal{H}_y = \left(y * \frac{1}{\pi t}\right)(t) \tag{2.1.3}$$

Here $*$ denote the convolution product between two functions.

This definition induces a multiplication in the frequency domain, from the Fourier transform of y noted $\hat{y}$, and the function $\omega \mapsto -i\,\mathrm{sign}(\omega)$.

Practically, from the signal y, calculation of the analytical signal y_a only requires to multiply each complex coefficients by 2 for every strictly positive frequency, by 0 for every strictly negative frequency, and performing the inverse Fourier transform.

The sign function *sign* is defined, $\forall x \in \mathbb{R}$ by :

$$sign(x) = \begin{cases} 1 & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x = 0 \\ -1 & \text{for} \quad x < 0 \end{cases}$$

From the analytical representation, the estimators $\hat{A}(t)$ and $\hat{\phi}(t)$, respectively of the *instantaneous amplitude and phase*, are computed as:

$$\hat{A}(t) = |y_a(t)| \tag{2.1.4}$$

$$\hat{\phi}(t) = \arg\left(y_a(t)\right) \tag{2.1.5}$$

For any complex $z \in \mathbb{C}$ such that $z = a + ib$ with $a, b \in \mathbb{R}$, the modulus of z is defined as

$$|z| = r = \sqrt{a^2 + b^2}$$

giving the representation $z = ze^{i\varphi}$ with $\varphi \in \mathbb{R}$ an argument in radian. Restriction of this real on the interval $]-\pi, \pi]$ gives the argument $\arg(z)$.

A additional feature to be calculated is the *instantaneous frequency* $\nu(t)$, with estimator $\hat{\nu}(t)$ defined as:

$$\hat{\nu}(t) = \frac{1}{2\pi}\frac{d}{dt}\phi(t) \tag{2.1.6}$$

Identification of the arrival time is straightforward:

$$\hat{u} = \arg\max_u A(u) \tag{2.1.7}$$

In noisy situations, signals are corrupted with random perturbations, which disturb the estimation of the arrival time. More advanced techniques must be employed, first to separate the pulse signal from the perturbations, to next apply estimation of the arrival time, and next to model both the pulse and perturbation to apply estimation methods.

2.1.2 *Wigner-Ville distribution*

Compared with the classical Short-Term-Fourier-Transform (STFT), the Wigner-Ville distribution (or WVD) does not use any window function (or time-frequency atom). It raises the highest possible time/frequency resolution, at the expense of a degradation of the visual analysis of the signal, due to interferences which can mist elements.

The definition of the WVD for a single function f is given as [Mal08]:

$$W_f(t,\omega) = \int_{-\infty}^{\infty} f(t+\frac{\tau}{2}) f^*(t-\frac{\tau}{2}) e^{-j\tau\omega}\, d\tau \tag{2.1.8}$$

ω is the angular frequency, defined as $\omega = 2\pi\nu$ with ν the frequency expressed in Hz.

or using Parseval formula, with a frequency integration:

$$W_f(t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\omega+\frac{\xi}{2}) \widehat{f}^*(\omega-\frac{\xi}{2}) e^{j\xi t}\, d\xi \tag{2.1.9}$$

The WVD of two functions f and g is defined as:

$$W_{f,g}(t,\omega) = \int_{-\infty}^{\infty} f(t+\frac{\tau}{2}) g^*(t-\frac{\tau}{2}) e^{-j\tau\omega}\, d\tau \tag{2.1.10}$$

If $f = f_1 + f_2$, then

$$W_f(t,\omega) = W_{f_1}(t,\omega) + W_{f_2}(t,\omega) + I_{f_1,f_2}(t,\omega) \tag{2.1.11}$$

with $I_{f_1,f_2}(t,\omega)$ defines the cross term (or interference term) such that

$$I_{f_1,f_2}(t,\omega) = W_{f_1,f_2}(t,\omega) + W_{f_2,f_1}(t,\omega) \tag{2.1.12}$$

Example 2.1.1. *For a Gaussian Gabor signal f:*

$$f(t) = A e^{-\frac{(t-u)^2}{s}} e^{j\zeta(t-u)} e^{j\phi} \tag{2.1.13}$$

the WVD can be formally calculated:

$$W_f(t,\omega) = \sqrt{2\pi s} A^2 e^{-\frac{s}{2}(\zeta-\omega)^2} e^{-\frac{2}{s}(u-t)^2} \tag{2.1.14}$$

Example 2.1.2. *For a Gaussian Chirplet signal f:*

$$f(t) = A e^{-\frac{(t-u)^2}{s}} e^{j2\pi[\nu_0+\beta(t-u)]} e^{j\phi} \tag{2.1.15}$$

the WVD can also be formally calculated:

$$W_f(t,\omega) = \sqrt{2\pi s} A^2 e^{-2\pi^2 s[(\frac{\omega}{2\pi}-2\beta(t-u))-f_0]^2} e^{-\frac{2}{s}(u-t)^2} \tag{2.1.16}$$

From a practical point of view, the previous formula directly shows the opposite behaviours of parameters on the temporal and frequency scales: a narrow atom in the time domain (small s) will give a wide atom on the frequency domain, and reciprocally, a spread atom in the time domain (large s) will give a narrow atom on the frequency domain.

We define the temporal and frequency centres as μ_t and μ_ω, and the temporal and frequency variances σ_t^2 and σ_ω^2 as (x is equal to t or ω, with $f(\omega)$ the Fourier transform of f):

$$\mu_x = \|f\|^{-2} \int x\, |f(x)|^2\, dx \tag{2.1.17}$$

$$\sigma_x^2 = \|f\|^{-2} \int (x-\mu_x)^2\, |f(x)|^2\, dx \tag{2.1.18}$$

Theorem 2.1.3. *[Mal08] The temporal variance and the frequency variance of $f \in L^2(\mathbb{R})$ satisfy :*

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4} \tag{2.1.19}$$

This inequality is also known as the Heisenberg uncertainty principle, in signal processing.

The formula 2.1.19 gives an equality if and only if there exist $(u,\zeta) \in \mathbb{R}^2$ and $(a,b) \in \mathbb{C}^2$ such that

$$f(t) = a \exp(i\zeta t - b(t-u)^2)$$

Such a signal is called by Mallat a *Gabor Chirp*.

The Wigner-Ville transform calculated for one Gabor pulse or one Chirplet is positive. However, for the sum of two signals, the cross-term integral can be negative at some points on the time frequency space. It produces interferences in the spectrum, which makes analysis with this tool complex with general signals.

Example 2.1.4. *For two Gabor signals f_1 and f_2 as defined in equation 2.1.13 with different parameters, the absolute value of the interference term can be formally calculated, and integrated over time and frequency domains, as:*

$$E_{f_1,f_2} = \int\int |W_{f_1,f_2}|\, d\omega\, dt = \frac{\sqrt{\pi} A_1 A_2 \sqrt{s_1 + s_2}}{2} \tag{2.1.20}$$

2.1.3 *Continuous wavelet transform*

Wavelet transform of signal x at scale s and time u, using the mother wavelet Ψ is defined as:

$$\mathcal{W}_x(u,s) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{s}} \Psi^* \left(\frac{t-u}{s} \right) dt \tag{2.1.21}$$

where Ψ^* denotes the conjugate function of Ψ.

is also found sometimes in the form $\mathcal{W}_x(u,s) = \langle x, \Psi_{u,s} \rangle$, with $\Psi_{u,s} = \frac{1}{\sqrt{s}} \Psi \left(\frac{t-u}{s} \right)$

One can also reconstruct the signal, under conditions:

$$x(t) = \frac{1}{K_\Psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \mathcal{W}_x(u,s) \frac{1}{\sqrt{s}} \Psi \left(\frac{t-u}{s} \right) du \frac{ds}{s^2} \tag{2.1.22}$$

Taking $s \in \mathbb{R}^{+,*}$ gives the Continuous Wavelet Transform (CWT), which can be used to analyse a frequency band with any precision. Taking $s = k^j, j \in \mathbb{N}$ for some $k \in \mathbb{N}^*$ gives the Discrete Wavelet Transform (DWT).

Common choice is $k = 2$

It has been demonstrated that the CWT provides an unbiased estimation of the ToF if the Morlet wavelet is used as basis. However, it is no longer possible for the central frequency, the estimation is indeed biased (see [CS05] or [CH10]), with bias which is function of the central frequency of the pulse.

However, the CWT is still a powerful tool to provide an estimation of the instantaneous amplitudes and phase of the signal. From a *scalogram*, it is possible to extract wavelet coefficients related to the local maxima points, which set is called a *ridge*, noted $\mathfrak{R}(s(u), u)$.

The scalogram at scale s and time u is defined as the energy of the wavelet coefficient at scale s and time u : $\mathcal{S}_\Psi(s,u) = |\mathcal{W}_\Psi(s,u)|^2$

The process is described as below, using the complex Morlet wavelets, from [Cur+08a]:

In the ridge, for all value of u, a specific scale $s(u)$ is attributed

1. Compute the Continuous Wavelet Transform (CWT) of y:

$$\mathcal{W}_y(s,u) \tag{2.1.23}$$

for $u = t_i,\ i = 1..N$ and the scaling factor selected s;

2. Compute the ridge of the CWT, *i.e.* the local maxima and their location $\mathfrak{R}(s(u), u)$;

3. Compute the Amplitude and the Phase, following:

$$A(u) = 2\frac{|\,\mathcal{W}_y(s(u),u)\,|}{\sqrt{s(u)}} \tag{2.1.24}$$

$$\phi(u) = \arg(\mathcal{W}_y(s,u)) \tag{2.1.25}$$

Instantaneous frequency is next easily obtained. The similarity with estimation based on the Hilbert transform is straightforward. However, selection of a specific frequency band allows to bypass the instabilities due to random noise perturbations.

2.1.4 *Hilbert-Huang transform*

The Hilbert-Huang Transform (or HHT) is an adaptative analysis tool based on the joint use of the Empirical Mode Decomposition (or EMD) scheme and the Hilbert transform. EMD has raised interest for providing an efficient method to extract, from a noisy multi-modal signals, Intrinsic Mode Functions (or IMFs), which will correspond to unimodal signals [HW08].

Definition 2.1.5. *A signal y is an IMF if:*

- *The absolute difference between the numbers of extreme and zeros crossing of y does not exceed* 1*.*
- *The mean values between the upper and lower envelope must be equal to* 0 *at any point.*

The original signal is the sum of the IMFs extracted. The previous method for amplitude and frequency estimation can then be used directly on each IMF. Moreover, the noise tends empirically to be extracted in the very first levels of decomposition, allowing fast denoising by removing the corresponding IMFs. However, this method lacks mathematical foundations [HW08], and the resulting IMFs can be very sensitive to the noise present in the original signal. More complex methods, like the Ensemble Empirical Mode Decomposition (or EEMD) [WH09], have been proposed to face this problem, by averaging the estimations on several runs of EMD on the signal corrupted with artificial noise, resulting in more stable results, at the price of an higher computational cost.

The algorithm proposed by Huang creates a sequence of IMFs, which are subtracted from the signal to produce the residual. Each IMF is constructed by removing the mean of the upper and lower envelopes from the signal, until the result fulfils the conditions to be an IMF.

After decomposition, the Hilbert transform is applied on each individual IMF. The resulting set of instantaneous amplitudes and frequencies is called the Hilbert spectrum.

This technique has been successfully used in NDT for extraction of transmitted pulses and echoes [LLY09] from an ultrasonic signal.

2.1.5 *Empirical wavelet transform*

The EMD is an adaptive method which does not rely on specific properties of the signal, in terms of frequencies, or abrupt changes. However, the decomposed IMF spectra are not necessary centred around one frequency. The Empirical Wavelet Transform (EWT) has been designed to circumvent this issue. After an adaptive partitioning of the Fourier spectrum, a wavelet filter bank is designed based on each identified support [Gil13]. The signal is next decomposed into components, each one with a compact support around a specific frequency. The only parameter to chose is the number of disjoint supports. This decomposition process has demonstrated superior performances to EMD for anomalies detection in oscillatory systems [War16]. Since the partitioning is performed in the frequency domain, this decomposition technique has poor performances if several components overlaps in the time-frequency domain. Example is the mixture of two chirps with closed centre frequencies and ToF[Gil13].

A Hilbert spectrum can be obtained using the Hilbert transform on each component, similarly to the HHT with the EMD.

However, choice of the number of supports is hard to find without knowledge about measurements. An external perturbation could easily corrupt the spectrum, and change the supports during a new measurement.

2.1.6 *Parametric estimation with the maximum likelihood estimator*

The common probabilistic model for ultrasonic pulses is the Gabor function, which is the product of a cosine, with central frequency ν_c and constant phase ϕ, and a Gaussian shape, with peak at the ToF u and spread s. An amplitude factor A is generally added. The parameters are gathered in the vector $\theta = [s, u, \nu_c, \phi, A]^T$, and the Gabor function is noted f_θ:

$$f_\theta(t) = A \exp\left(-\frac{(t-u)^2}{s}\right) \cos\left(2\pi\nu_c(t-u) + \phi\right) \tag{2.1.26}$$

Measurement perturbations are assumed to corrupt the part of the signal coming from a specific model, such as:

$$y_i = f_\theta(t_i) + w_i \tag{2.1.27}$$

with $w_i \sim \mathcal{N}(0, \sigma^2)$ for some positive scalar σ, the standard deviation of the noise.

With random Gaussian perturbation, the probability density functions of measurements given the real signal are Gaussian. If perturbations are independent, the total density function, called the likelihood of the signal, is:

$$p(y \mid \theta, \sigma) = \mathcal{L}(\theta, \sigma) \quad = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^N} \prod_{i=1}^{N} \exp\left(-\frac{(y_i - f_\theta(t_i))}{2\sigma^2}\right) \tag{2.1.28}$$

The Maximum Likelihood Estimator (or MLE) is the parameter vector which maximizes the signal likelihood, or equivalently, the signal log-likelihood, preferred for maximization, since the probability density alone can reach very small values. It is defined as:

$$\theta^{\text{MLE}} = \arg\max_{\theta} \log p(y \mid \theta, \sigma) \tag{2.1.29}$$

$$= \arg\max_{\theta} N \log\left(\sqrt{2\pi}\sigma\right)^2 - \sum_{i=1}^{N} \frac{(y_i - f_\theta(t_i))}{2\sigma^2} \tag{2.1.30}$$

The noise variance σ^2 is not known in the general case, and can be estimated by extending the parameter vector with it before the search for the MLE. However, knowledge of its value is not necessary for the estimation of only θ^{MLE}. For every variance σ^2, the previous maximization problem is equivalent to the following one:

$$\theta^{\text{MLE}} = \arg\min_{\theta} \sum_{i=1}^{N} (y_i - f_\theta(t_i))^2 \tag{2.1.31}$$

$$= \arg\min_{\theta} \|y - f_\theta(t)\|_2^2 \tag{2.1.32}$$

Therefore, finding the MLE with Gaussian noise involves a least-square minimization independent of the noise variance. Up to a constant, the optimal parameter vector of the above problem is the one which minimizes the Mean Square Error (or MSE) between the signal and the parametrized reconstruction using the Gabor function. Efficient algorithms such as the Gauss-Newton, or the Levenberg-Marquardt have largely been used in this context, taking advantage of the availability of the Jacobian matrix. Since they do not use the exact Hessian, but provide only an approximation of it, they are known to be more stable, and still converging to the global optimum. More general optimization methods exist, such as the famous Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [NW06]. However, the Gauss-Newton and Levenberg-Marquardt have, around the global minima, quadratic convergence rates. The BFGS, a quasi-Newton method, has only super-linear convergence rate. For this specific problem, using the BFGS algorithm would be counterproductive.

These properties are not preserved after calculation of the envelope. Measurements perturbations are no longer independent, and do not follow Gaussian probability functions. Indeed, [HWZ12] employ a Rice distribution to model the densities of the envelope of y, the points noted $\{A_i\}_{i=1}^N$, based on the results of [Ric45]:

$$p(A_i \mid \lambda_i) = \frac{A_i}{\sigma^2} \exp^{-\frac{A_i^2+\lambda_i^2}{2\sigma^2}} I_0\left(\frac{\lambda_i A_i}{\sigma^2}\right) \tag{2.1.33}$$

For a zeros signal, one can find the Rayleigh distribution with probability density $p(x \mid \sigma^2) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right)$ defined on $\mathbb{R}^{+,*}$.

with $\{\lambda_i\}_{i=1}^N$ the points of the enveloppe of a Gabor function, calculated from the parameter vector θ.

The Quasi-Maximum Likelihood (QML) method assumes independence between the calculated amplitudes to solve the maximization

problem the simplified version of the likelihood, called the quasi-likelihood:

$$Q\mathcal{L}(\theta, \sigma) = \prod_{i=1}^{N} p(A_i \mid \theta, \sigma) \tag{2.1.34}$$

Optimization is generally made with the more convenient logarithm of quantity $Q\mathcal{L}(\theta, \sigma)$, giving the formula for the log-quasi-likelihood $\log Q\mathcal{L}(\theta, \sigma)$, which complete expression is:

$$\log Q\mathcal{L}(\theta, \sigma) = \sum_{i=1}^{N} \log A_i - \log \sigma^2 - \sum_{i=1}^{N} \frac{A_i^2 + \lambda_i^2}{2\sigma^2} + \sum_{i=1}^{N} \log I_0 \left(\frac{\lambda_i A_i}{\sigma^2} \right) \tag{2.1.35}$$

with I_0 the modified Bessel function of order 0. Given the form of the function to minimize, the BFGS algorithm is a good choice, since the previous ones can no longer be used.

2.1.7 *Comparison of estimators*

To compare performances of the estimators presented in this section, the following signal was simulated. A Gabor signal is generated, with parameter vector $\theta = [s, u, \nu_c, \phi, A]^T = [0.25, 3, 5, 0, 1]^T$. Signal is simulated on the interval $[0, 5\,s]$ with sampling interval $\Delta t = 0.01\,s$. A random Gaussian noise is added, to reach a SNR level starting at $-10\,dB$ up to $20\,dB$, by step of $1\,dB$. For each SNR, for each estimation of ToF u and frequency ν given by the previous methods, the MSE is averaged over 1000 realizations. Methods used to estimate the parameters are: the CWT, the MLE, the maximum of the envelope computed with the Hilbert transform (called here the Hilbert method), the QML and the EMD. For estimation with the EMD, the envelope maximum of the IMF with higher correlation with the intial signal is selected (except the first IMF, related to noise). Of course, the QML method does not provide an estimation of ν. Resulting curves are gathered in figure 2.1.

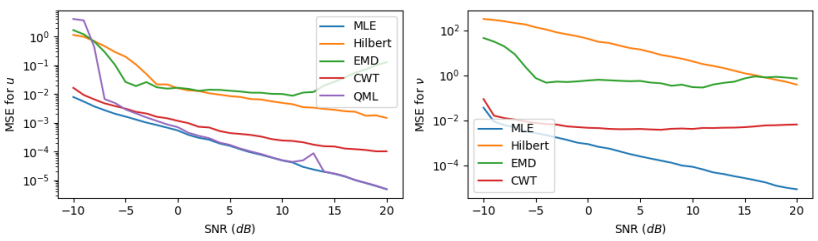

Figure 2.1: Comparison of estimation performances between CWT, MLE, Hilbert, QML and EMD method, with 1000 Monte Carlo iteration to compute the average MSE on a simulated Gabor signal.

Surprisingly, estimations with the EMD method are not reliable for very high SNR. Hilbert method is not expected to produce precise estimates, and shows only the important perturbations caused by noises. For low SNR, best methods are the CWT and the MLE. QML performs similarly, except for very low SNR. Absence of improvements for the CWT for frequency estimations is more an issue related to

the frequency range analysed. Comparing the similar performances of the MLE and the QML, using a complex model with Rice distribution seems not interesting. Moreover, the QML does not provide information about phase and frequency.

Consequently, for simple estimation on noisy signals, using the MLE or the CWT method are a practical and robust choice, under the assumption of a Gabor model.

2.2 Sparse decomposition of ultrasonic signals

2.2.1 *Gabor dictionaries*

Instead of extracting one single vector from an acquisition, it can be useful to decompose the measurement vector into a specific basis called a *dictionary*. Purpose of dictionaries is to decompose a signal y into as sum of distinct elements, called *atoms*, which belong to a set $\mathfrak{D}$, the dictionary.

An overcomplete dictionary $\mathfrak{D}$ contains a number of atoms much superior to the dimension of y, such that:

$$y \approx \mathfrak{D}x \text{ with } \begin{cases} y \in \mathbb{C}^n \\ \mathfrak{D} \in \mathbb{C}^{n \times m} \\ x \in \mathbb{C}^m \end{cases} \text{ and } m \gg n \tag{2.2.1}$$

The interest of such basis is to express the vector y with very few atoms, such that x is *sparse*.

Definition 2.2.1. *A matrix is said to be sparse if most of its elements are zeros.*

The minimization problem to be solved has the form:

$$x^\star = \arg\min_x \|y - \mathfrak{D}x\|_2^2 + \xi\|x\|_0 \tag{2.2.2}$$

With $\|x\|_0$ the norm 0 of x, defined as the number of non-zero elements of x.

for some positive scalar ξ. Higher values of ξ will promote sparser vectors. However, this problem is hard to solve numerically, and different approaches must be employed to approximate the problem.

For ultrasonic signals, a basis composed of Gabor atoms has demonstrated great performances [OF98; DS01b]. The Gabor dictionary is defined as $\mathfrak{D} = \{g_\gamma\}_{\gamma \in \Gamma}$ with Γ our set of parameters. Each element g_γ an atom, is the Gabor function parametrized by the vector $\gamma = [s, u, \nu_c, \phi]^T$, and defined as:

$$g_\gamma(t) = \exp\left(-\frac{(t-u)^2}{s}\right)\cos\left(2\pi\nu_c(t-u)+\phi\right)C(\gamma) \tag{2.2.3}$$

with $C(\gamma)$ the normalization constant such that $\|g_\gamma\| = 1$.

A M-decomposition of a signal y with dictionary $\mathfrak{D}$ is a linear combination of atoms such that :

$$y \approx \sum_{i=1}^{M} A_i g_{\gamma_i} \tag{2.2.4}$$

for some positive scalars $A_1, ..., A_M \in \mathbb{C}^+$.

Example 2.2.2. *For a discrete signal* $(y_i)_{i=1}^N$ *with corresponding sampling times* $(t_i)_{i=1}^N$*, we can define the set* Γ *by all the combinations between* n_s *values of spread* $s_1, ..., s_{n_s}$*,* N *arrival times as the values* $t_1, ..., t_N$ *and* n_ν *values of central frequencies* $\nu_1, ..., \nu_{n_\nu}$ *,* n_ϕ *values of phase* $\phi_1, ..., \phi_{n_\phi}$*. In this situation, the dictionary is therefore composed of* $n_s \times N \times n_\nu$ *atoms of size* N*.*

An algorithm to sequentially decomposed a vector y into $K \in \mathbb{N}$ Gabor atoms of $\mathfrak{D}$ has been proposed, called the *Matching Pursuit* (or MP) [MZ93].

Let R_k be the residual vector obtained after k iterations of the decomposition. This vector is calculated sequentially, by removing the found decomposition vector, with $R_0 = y$. It is possible to write:

$$y = \sum_{k=1}^{K} A_k g_{\gamma_k} + R_k, \forall k \in \mathbb{N}^* \tag{2.2.5}$$

for some $A_1, ..., A_K \in \mathbb{R}^+$.

Each iteration can be described, at iteration k, using the residual R_{k-1} as:

1. Find $\gamma_k = \arg\max_\gamma \langle R_{k-1}, g_\gamma \rangle$
2. Set $A_k = \langle R_{k-1}, g_{\gamma_k} \rangle$
3. Set $R_k = R_{k-1} - A_k g_{\gamma_k}$

The vector α which minimizes the least-square error $\|y - \alpha x\|^2$ is the scalar product $\langle y, x \rangle$ with $y, x \in \mathbb{C}$ and $\|x\|^2 = 1$

The search in the previous dictionary raise the issue of estimation of phase ϕ. Exploring n_ϕ values on an interval $\left[\phi_{\text{inf}}, \phi_{\text{sup}}\right]$ multiply the numbers of scalar product to compute. The approach proposed by Mallat and Zang [MZ93] is to perform the exhaustive search on a lighter dictionary, but composed of atoms belonging to $\mathbb{C}^N$. The complex Gabor function are employed, parametrized by the vector $\tilde{\gamma} = [s, u, \nu_c]^T$, and defined as:

$$\tilde{g}_{\tilde{\gamma}}(t) = \exp\left(-\frac{(t-u)^2}{s}\right) \exp\left(2i\pi\nu_c(t-u)\right) * \tilde{C}(\gamma) \tag{2.2.6}$$

with $\tilde{C}(\tilde{\gamma})$ the normalization constant such that $\|\tilde{g}_{\tilde{\gamma}}\| = 1$.

The phase ϕ is dropped from the vectors $\gamma \in \Gamma$, forming the parameters vectors $\tilde{\gamma} \in \tilde{\Gamma}$. The search operation performed is looking now for the vector $\tilde{\gamma}_k$ such that:

The residual vector R_k is still real, but the scalar product apply now in its complex form, since $\tilde{g}_{\tilde{\gamma}}$ is complex.

$$\tilde{\gamma}_k = \arg\max_{\gamma} \langle R_{k-1}, \tilde{g}_{\tilde{\gamma}} \rangle \tag{2.2.7}$$

The phase is then estimated as :

$$\phi_k = \arg\left(\langle R_{k-1}, \tilde{g}_{\tilde{\gamma}_k} \rangle\right) \tag{2.2.8}$$

The vector which parametrized the real Gabor atom is then reconstructed using this estimation : $\gamma_k = \left[\tilde{\gamma_k}^T, \phi_k\right]^T$.

The estimation of the real coefficient involve the normalization coefficient of the real atom g_{γ_k}:

$$A_k = C(\gamma) |\langle R_{k-1}, \tilde{g}_{\tilde{\gamma}_k} \rangle| \tag{2.2.9}$$

Following every search in the dictionary, an optimization method can then be used to tune the parameter vector $\theta = \left[\gamma_k^T, A_k\right]^T$ to minimize locally the quadratic error between R_{k-1} and $f_\theta(t)$.

A formulation of the time-frequency energy of the signal decomposed with MP has been proposed in paper [MZ93]. If f is of the form:

$$f = \sum_{i=0}^{\infty} A_i g_{\gamma_i} \tag{2.2.10}$$

then applying the WVD gives:

$$W_f(t,\omega) = \sum_{i=0}^{\infty} |A_i|^2 W_{g_{\gamma_i}}(t,\omega) + \sum_i \sum_{j \neq i} A_i A_j^* W_{g_{\gamma_i}, g_{\gamma_j}}(t,\omega) \tag{2.2.11}$$

The term $\sum_i \sum_{j \neq i} A_i A_j^* W_{g_{\gamma_i}, g_{\gamma_j}}(t,\omega)$ in equation 2.2.11 is assumed, according to them, to be the cross interference term. Consequently, they defined an energy formula without it as:

$$E_f(t,\omega) = \sum_{i=0}^{\infty} |A_i|^2 W_{g_{\gamma_i}}(t,\omega) \tag{2.2.12}$$

The energy representation in the time-frequency domain becomes the sum of independent positive components. It must be emphasized that each component should be well separated from the others, to reduce interactions and keep the assumption of zero cross-terms.

2.2.2 *Probabilistic optimization*

An important issue with the MP algorithm is the impossibility to correct vectors obtained in the past. Due to this constraint, at some iterations, atoms could be selected only to balance imprecise previous computed vectors. A major improvement would be to correct, after each iteration, the set of all selected atoms. Since this set grows linearly with time, applying an optimization method on the joint vector of parameters will rise the issue due to its high dimensionality.

The *Expectation Maximization* (EM) methods aims at breaking a hard to optimize problems into easier ones thanks to the introduction of auxiliary variables [DLR77; MK08].

The aim of the EM algorithm is to update the parameter set $\Theta = [\theta_1, .., \theta_J]$ to reduce the total cost:

$$\text{Cost}(\Theta) = \sum_{j=1}^{J} \text{Cost}(\theta_j) \tag{2.2.13}$$

The Expectation step (E-step) constructs all the J residual vectors. The Maximization step consists in evaluating each new parameter conditionally to its related residual. For important values of J, the drawback of this approach is the need to keep in memory many residuals. Moreover, inside each Maximization step (M-step), benefit of one parameter update of Θ is only observed the end, when the new total cost $\text{Cost}(\Theta)$ is calculated.

Description of one iteration of the EM algorithm:

1. Init $\Theta^{(k)} = \left[\theta_1^{(0)}, .., \theta_J^{(k)}\right]$
2. For $j = 1, .., J$:

$$\text{E-step} : \text{compute } r_j^{(k)}(\Theta^{(k)})$$

3. For $j = 1, .., J$:

$$\text{M-step} : \theta_j^{(k+1)} = \arg\min_{\theta} \text{Cost}(\theta, r_j^{(k)}(\Theta^{(k)}))$$

4. Check convergence criterion

The Space Alternating Generalized EM (SAGE) algorithm is an adaptation of the previous algorithm, which has better convergence properties [FH94]. Instead of updating the total cost at the M-step end, a new computation is performed each time a vector of Θ is updated.

Description of one iteration of the SAGE algorithm:

1. Init $\Theta^{(k)} = \left[\theta_1^{(0)}, .., \theta_J^{(k)}\right]$
2. Set $\Theta^{(k+1)} = \Theta^{(k)}$
3. For $j = 1, .., J$:

$$\text{E-step (SAGE): compute } r_j^{(k)}(\Theta^{(k+1)})$$
$$\text{M-step (SAGE): } \theta_j^{(k+1)} = \arg\min_{\theta} \text{Cost}(\theta, r_j^{(k)}(\Theta^{(k+1)}))$$

4. Check convergence criterion

The residual estimated at each step is dependent of the application.

A hard rule will place at each iteration the full part not already explicated by the functions on the residual:

$$r_j^{(k)}(\Theta) = f(\theta_j) + \left[z - \sum_{l=1}^{J} f(\theta_l)\right] \tag{2.2.14}$$

A soft rule will divide this part across all residuals:

$$r_j^{(k)}(\Theta) = f(\theta_j) + \frac{1}{J}\left[z - \sum_{l=1}^{J} f(\theta_l)\right] \tag{2.2.15}$$

Using an iterative optimization method corrects progressively the results of decomposition using MP algorithms. This approach perform well signal sparse in basis formed of Gabor atoms and even more general pulses [DS01a].

2.2.3 *Saturation model*

To face the issue of saturation presented in chapter 1, we propose the following approach. Saturation is model by the variable d_i, which equals 1 if the signal is inside the measurement window, and 0 outside.

$$d_i = \begin{cases} 1 & \text{if} \quad x_i \in]x_{\text{inf}}, x_{\text{sup}}[\\ 0 & \text{elsewhere} \end{cases} \tag{2.2.16}$$

To recover a tractable estimation problem, a practical solution is to remove the points where saturation occurs from the calculation of the objective function. Times of saturation are considered as missing values. We try to find the signal s such that:

$$\hat{\theta} = \arg\min_{\theta} \sum_{t} \left(y_i - f_\theta(t_i)\right)^2 d_i \tag{2.2.17}$$

or with matrix formulation:

$$\hat{\theta} = \arg\min_{\theta} \left(y - f_\theta(t)\right)^T \mathbf{D} \left(y - f_\theta(t)\right) \tag{2.2.18}$$

with

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & d_T \end{pmatrix} \tag{2.2.19}$$

Calculation of the cost is reduced to the measurements for which saturation does not occur. Partitioning of the data between saturated and regular measurements allows to solve a sub-problem, which relies on the independence assumption between perturbations. Estimation with only the envelope would require a more complex model for saturation. Separation of data is no longer possible because of interactions between values of the two partition to compute the envelope points.

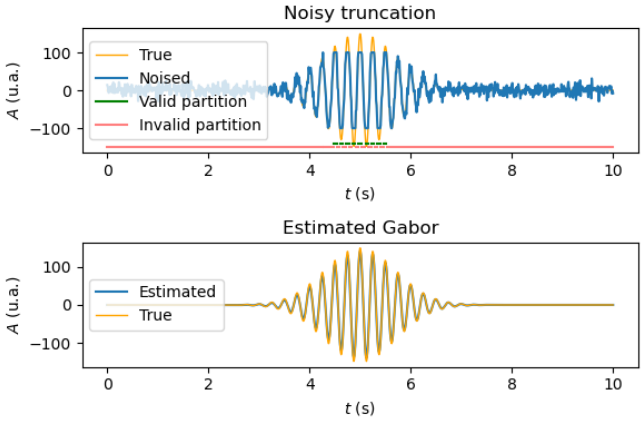

Figure 2.2: Simulation with $\theta = [1, 5, 4, 0, 150]$, $\sigma = 10$, and truncation at ± 100. Estimation gives $\hat{\theta} = [0.96, 5.0, 4.0, 0.56, 140]$. Invalid partition represents 6% of data.

Figure 2.2 illustrates the estimation procedure from a truncated signal. Although the estimated parameter vector is close to the real one, it gives a practical solution to this problem.

2.2.4 *Enforce sparsity with regularization*

Decomposition of atoms is difficult when pulses are interacting. Conditions of separability have been studied in [Lu+16], depending on

the distance between their central frequencies and arrival times. In the case of ultrasonic signals, interactions are widely due to temporal close receptions of the pulses.

To solve this issue, a regularization of the decomposition problem is proposed. It corresponds to adding a term in the loss function which penalize very spread pulses.

The previous problem involving only the Gabor function with random Gaussian noise was using the following cost function:

$$\text{loss}_{\text{MSE}}(\theta) = \|y - f_\theta(t)\|_2^2 \tag{2.2.20}$$

Regularization L_2 centred:

$$\text{loss}_{\text{reg}}(\theta, \lambda) = \text{loss}_{\text{MSE}}(\theta) + \lambda\|s\|^2 \tag{2.2.21}$$

Regularization L_2 non-centred:

$$\text{loss}_{\text{reg}}(\theta, \lambda, s_0) = \text{loss}_{\text{MSE}}(\theta) + \lambda\|s - s_0\|^2 \tag{2.2.22}$$

For the L_2 centred and non-centred regularization, the quantity to minimize can still take the form of a least-square problem. Consequently, the optimization algorithms already employed for the non-regularized loss are still available.

The regularization parameter λ depends on the measurement noise variance, and is therefore case-dependent. However, the centre spread parameter s_0 can easily be learned from clean signals with separable non-interacting pulses.

The previous regularization didn't enforce strictly positive values of s. A Gamma prior form regularization can be employed for this purpose:

$$\text{Loss}_{\text{reg}}(\theta, \lambda_1, \lambda_2) = \text{Loss}_{\text{MSE}}(\theta) + \lambda_1 \log s - \lambda_2 s \tag{2.2.23}$$

This formulation does not fulfil the conditions for use of the Gauss-Newton or Levenberg-Marquardt algorithms. The BFGS algorithm is a standard choice.

Figure 2.3 highlights the utility of regularization with an signal composed of two Gabor pulses. A simple optimization procedure could result in a local optimum with two degenerated atoms of high amplitudes, while the mixture gives low amplitude signal. Regularization allows discrimination of such degenerated situation. It does not necessary improved the estimation procedure (the algorithm is still stuck into a local optimum, but the global one has now significant lower cost due to the penalization of high values of s).

2.3 Filtering

Acquisitions from ultrasonic sensors can individually be used for damage detection. Features extracted with time-frequency representation, inserted in basic classification algorithms allow independent analysis of a medium area. However, links between multiple acquisitions are never

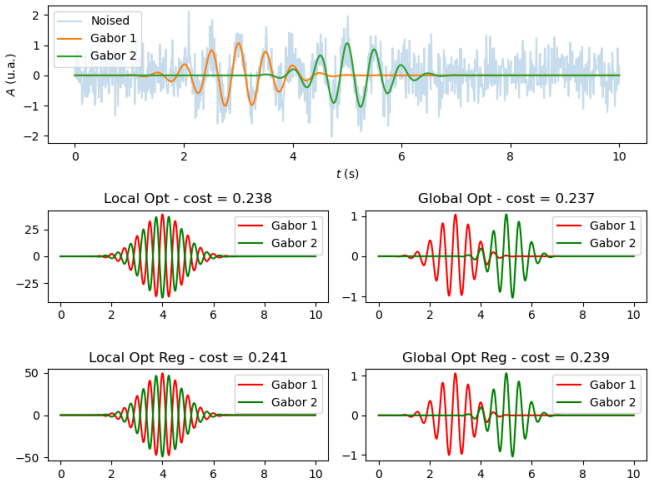

Figure 2.3: Simulation with $\theta_1 = [1, 3, 2, 0, 1]$, $\theta_2 = [1, 5, 2, 0, 1]$. Optimization is performed with and without regularization

exploited. The objective of this section is to include the evolution of features across acquisitions into a defect detection process.

2.3.1 *State space approach*

Operator is interesting into information gathered in a vector x, called a *state vector*, which belongs to a *state space* $\mathbb{R}^{n_x}$, for some $n_x \in \mathbb{N}^*$. This vector is generally not directly observable. Access of information about x is possible through a vector z called a *measurement* or *observation* vector, which belongs to a *measurement* or *observation space* $\mathbb{R}^{n_z}$, for some $n_z \in \mathbb{N}^*$. Knowledge about x is gathered in the *probability density function* (pdf) $p(x \mid z)$.

Sequential state space models assume that measurements are gradually obtained by the operator, for instance coming from a sensor. Let $(z_k)_{k=1..N}$ be the sequence of N *observations*.
The problem is transfered from the estimation of vector x to the pdf related to the variables x_k.

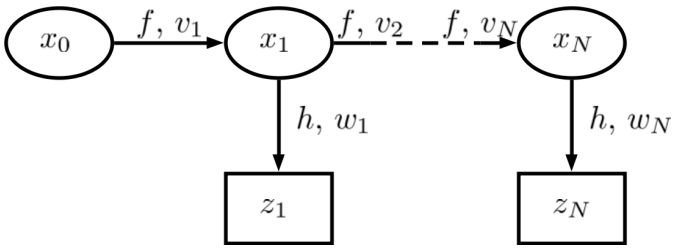

Figure 2.4: General state-space model

The following state-space model is assumed here:

- A deterministic function f, called the transition function, which take as arguments the state x_{k-1} and a random variable v_k to produce x_k;
- A deterministic function h, called the observation function, which take as arguments the state x_k and a random variable w_k to produce

the measurement z_k.

The filtering process is iterative and can be separated into two mains steps:

- Prediction: from the prior distribution $p(x_{k-1} \mid z_{1:k-1})$ (or $p(x_0)$), we compute the predicted distribution $p(x_k \mid z_{1:k-1})$ using the Chapman–Kolmogorov equation

$$p(x_k \mid z_{1:k-1}) = \int p(x_k \mid x_{k-1}) p(x_{k-1} \mid z_{1:k-1}) \, dx_{k-1} \tag{2.3.1}$$

- Update: from the previous predicted distribution, use of measurement z_k to compute the posterior distribution $p(x_k \mid z_k)$ with the Bayes' theorem

$$p(x_k \mid z_{1:k}) = \frac{p(z_k \mid x_k) p(x_k \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})} \tag{2.3.2}$$

The denominator term $p(z_k \mid z_{1:k-1}) = \int p(z_k \mid x_k) p(x_k \mid z_{1:k-1}) \, dx_k$ acts as a normalization term. It is sometimes referred as the Bayes factor or the likelihood of measurement z_k.

Example 2.3.1. *An important model is the Linear transition and measurement with additive zero-mean Gaussian noises, defined as:*

$$x_k = F x_{k-1} + v_k \tag{2.3.3}$$
$$z_k = H x_k + w_k \tag{2.3.4}$$

with:

$$v_k \sim \mathcal{N}(0, Q_k) \tag{2.3.5}$$
$$w_k \sim \mathcal{N}(0, R_k) \tag{2.3.6}$$

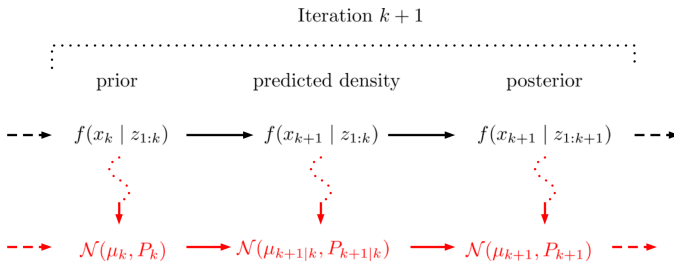

Figure 2.5: Kalman recursion

for some positive definite matrix Q_k and R_k, transition square matrix F and measurement matrix H. If a Gaussian prior is assumed on the state x_0, i.e. $x_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$, the Kalman equations allow to compute analytically the posterior distributions $p(x_k \mid z_k)$, which take the form of Gaussian distributions [Kal60; Sär13] (entirely characterized by their means μ_k and covariance matrices P_k, calculated recursively at each step k, see figure 2.5).

Sequential Monte Carlo techniques (or SMC) are a general class of methods designed to estimate sequentially the posterior distributions,

using stochastic approximations. For general models, an analytical form is not necessary available. This is the case with non-linear transition and measurement functions, non-additive or non-Gaussian random perturbations. For each step k, the posterior distribution given the measurements from step 1 to step k is approximated by a discrete sum of N_p Dirac functions of vectors paths $x_{1:k}^{(i)}, \forall i = 1, .., N_p$ called *particles*, with respective weights $W_k^{(i)}, \forall i = 1, .., N_p$, each of which depends on the corresponding particle:

$$p(x_{1:k} \mid z_{1:k}) \approx \sum_{i=1}^{N_p} W_k^{(i)} \delta_{x_{1:k}^{(i)}}(x_{1:k}) \tag{2.3.7}$$

Example 2.3.2. *The Sequential Importance Resampling (SIR) scheme, usually called the Particle Filter (PF), is an efficient method in SMC.*

Particle filters approximate the filtering distribution by a Dirac mixture with specific weights, as illustrated in figure 2.6:

$$p(x_k \mid z_{1:k}) \approx \sum_{i=1}^{N_p} w_k^{(i)} \delta_{x_k^{(i)}}(x_k) \tag{2.3.8}$$

For the next step, N_p sampled are simulated from the importance distribution q:

$$x_{k+1}^{(i)} \sim q(x_k \mid x_{1:k-1}^{(i)}, z_{1:k}), \forall i = 1..N_p \tag{2.3.9}$$

Their corresponding weights are then updating according to:

$$w_{k+1}^{(i)} = \frac{\tilde{w}_{k+1}^{(i)}}{\sum_{i=1}^{N_p} \tilde{w}_{k+1}^{(i)}}, \forall i = 1..N_p \tag{2.3.10}$$

with $\tilde{w}_{k+1}^{(i)}$ the unnormalized weights, defined as:

$$\tilde{w}_{k+1}^{(i)} = \alpha_{k+1}^{(i)} w_k^{(i)}, \forall i = 1..N_p \tag{2.3.11}$$

using the following importance ratio:

$$\alpha_{k+1}^{(i)} = \frac{p(z_{k+1} \mid x_{k+1}^{(i)}) p(x_{k+1}^{(i)} \mid x_k^{(i)})}{q(x_k \mid x_{1:k-1}^{(i)}, z_{1:k})}, \forall i = 1..N_p \tag{2.3.12}$$

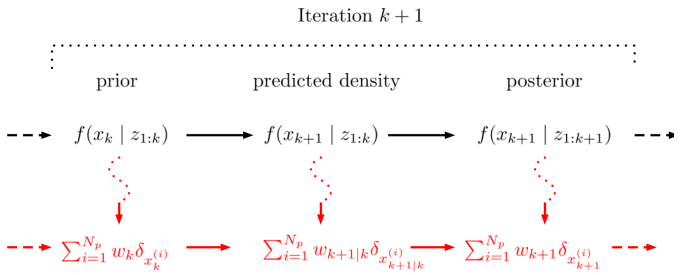

Figure 2.6: SIR recursion.

A resampling step is next applied at some iteration, to cope with the phenomenon called weight degeneracy, where all particle get weights close to 0, except some of them. However, the resampling method can

also induce one other phenomenon called the sample impoverishment, when only a small set of particles are resampled. The exploration of the state space is then concentrated on limited regions, reducing the performance of the filter. To avoid this, advanced resampling methods have been studied, for instance with deterministic selection of the particle, or clustering steps. This allows to bring more diversity in the particle population.

The likelihood of the data observed at step $k+1$ can therefore be estimated by:

$$p(z_{k+1} \mid z_{1:k}) \approx \sum_{i=0}^{N_p} w_{k+1}^{(i)} \alpha_{k+1}^{(i)} \tag{2.3.13}$$

2.3.2 *Smoothing*

Give a finite measurement vector $z_{1:N}$, the filtering recursions allow to compute sequentially the filtering densities $p(x_k \mid z_{1:k})$ for any given $k \in [\![1, N]\!]$. This operation can, in specific case, be done online (see examples 2.3.1 and 2.3.2). However, after reception of all measurement, we could also evaluate the full posterior distribution $p(x_{0:N} \mid z_{1:N})$, called the *smoothing* distribution.

Example 2.3.3. *A analytical form of the marginal smoothing solution is available in the Linear model with Gaussian perturbations. It has been found by the maximization of the marginal log-likelihoods of* $\log p(x_k \mid y_{1:N})$ *and* $\log p(x_k, x_{k+1} \mid y_{1:N})$ *[RTS65].*

The paper finds the marginal MLE (marginalization over all $x_j, \forall j \neq k$), but not the joint MLE $p(x_{0:k} \mid y_{1:N})$.

$$\bar{x}_{k|N} = x_{k|k} + C_k \left[\bar{x}_{k+1|N} - F \cdot x_{k|k}\right] \tag{2.3.14}$$

with

$$C_k = P_{k|k} \cdot F^t \cdot P_{k+1|k}^{-1} \tag{2.3.15}$$

and

$$\bar{P}_{k|N} = P_{k|k} + C_k \cdot \left(\bar{P}_{k+1|N} - P_{k+1|k}\right) \cdot C_k^t \tag{2.3.16}$$

The marginal distribution obtained is therefore :

$$p(x_k \mid y_{1:N}) = \mathcal{N}(x_k \mid \bar{x}_{k|N}, \bar{P}_{k|N}) \tag{2.3.17}$$

Example 2.3.4. *A simple solution of the smoothing problem is to keep the final posterior density calculated from the particle approximation [Kit96]. Following the classical SIR algorithm, an important difference is that instead of manipulating particles which are vectors with their weights, we propagate them with their lineage (theirs ancestors, or histories) . When conditions are met, the full particles are resampled with their own ancestors. Their dimensions increase with iterations. At final stage, the final approximation is:*

What [GDW04] calls this filter the "standard trajectory-based smoothing".

Instead of using at step k, $x_k^{(i)}, \forall i = 1, .., N_p$, we manipulate $x_{0:k}^{(i)}, \forall i = 1, .., N_p$

$$p(x_{0:N} \mid z_{1:N}) = \sum_{i=1}^{N_p} w_N^{(i)} \delta_{x_{1:k}^{(i)}}(x_{1:k}) \tag{2.3.18}$$

The main drawback of this approach is due to sample impoverishment: as particles are resampled, ones with low weights disappear from the population with their lineages, and particles tend to share the same ancestors. For long signals, the first steps have poor diversity. The characterization of the posterior at these stages will be extremely limited.

Example 2.3.5. *The marginal Forward Filter Backward Smoother (FFBSm) [DGA00] express the marginal smoothing distribution* $p(x_k \mid z_{1:N}), \forall k = 1, .., N$, *using Kitagawa's recursive formula [Kit87]. Here, from the previously calculated filtering distributions, the weights of particles are corrected to get a particle approximation with same supports for the filtering and marginal smoothing distributions:*

$$p(x_k \mid z_{1:N}) = \sum_{i=1}^{N_p} w_{k|N}^{(i)} \delta_{x_k^{(i)}}(x_k) \tag{2.3.19}$$

with $\forall i = 1, .., N_p$*:*

$$w_{k|N}^{(i)} \tag{2.3.20}$$

and $\forall i = 1, .., N_p, \forall k = 0, .., N-1$*:*

$$w_{k|N}^{(i)} = \sum_{j=1}^{N_p} w_{k+1|N}^{(j)} \frac{w_k^{(i)} p(x_{k+1}^{(j)}) \mid x_k^{(i)}}{\sum_{l=1}^{N_p} w_k^{(l)} p(x_{k+1}^{(j)}) \mid x_k^{(l)}} \tag{2.3.21}$$

This algorithm is deterministic given a run of a SIR algorithm. However, it has high complexity $\mathcal{O}(N_p^2)$*, which can only be reduced with the number of particles.*

Example 2.3.6. *The Forward Filter Backward Sampling (FFBSi) is a stochastic algorithm, providing a SIR run [GDW04]. It creates particle paths, starting from the final density* $p(x_N \mid z_{1:N})$*, and sample its ancestors using the intermediate filtering distributions, producing a vector* $\tilde{x}_{0:N}$*. This operation is repeated* S *times, to produce the following particle estimation, with similar form as the simple posterior smoother:*

$$p(x_{0:N} \mid z_{1:N}) = \frac{1}{S} \sum_{s=1}^{S} \delta_{\tilde{x}_{0:N}^{(s)}}(x_{0:N}) \tag{2.3.22}$$

The backward procedure to sample one path is described as:

- *choose* $\tilde{x}_N = x_N^{(i)}$ *with probability* $w_N^{(i)}$
- *for* $k = N-1, .., 0$*:*
  - $w_{k|k+1}^{(i)} \propto w_k^{(i)} p(\tilde{x}_{k+1} \mid x_k^{(i)})$
  - *choose* $\tilde{x}_k = x_k^{(i)}$ *with probability* $w_{k|k+1}$

A backward run has complexity $\mathcal{O}(N_p)$*, and is repeated* S *times. For increase of precision of the particle smoothing distribution, higher values of* S *must be used [Lin13]. The total complexity is consequently* $\mathcal{O}(SN_p)$*, which offers higher flexibility compared with the deterministic FFBSm.*

2.3.3 *Random Finite Set (RFS) filtering*

State space models based on vector measurements are useful when dealing with a single sensor, and single measurement information. The purpose of this part is to present formulation of the filtering equations using the Finite Set Statistics (or FISST) theory developed by Mahler [Mah07b], which allows manipulation of sets instead of vectors.

Filtering with state space models assumes one measurement vector at each step, used to gather information about the state. However, many applications do not provide a single vector during the measurement process. Examples are numerous in the tracking literature, in which a simple thresholding procedure on the data coming from the sensor creates multiple detections. The measurement vector produced by the state is embedded with random vectors in a set, without a priori knowledge about which one is derived from the state. In addition, there is still the possibility of non-detection of this measurement vector of interest. The probability of detection is generally described by the parameter p_D. The measurement set obtained at step k can be described as:

$$Z_k = T_k^D\left(\{x_k\}\right) \cup C_k \tag{2.3.23}$$

with T_k^D a set which contains the single measurement coming from the state x_k with probability $p_D(x)$, possibly depending on x, therefore in the case of detection, and which is empty with probability $1-p_D$, in the opposite case. The state x_k is referred as the *target* in the tracking literature. The set C_k is assumed to contain the random measurement detected with no link to the state, called the *clutter*.

Solving this estimation problem with an enumeration of all the assignment possibilities brings a number of cases which increases exponentially with steps. This situation, called an *hard-assignment* method, can be partially solved with a selection of several best assignments, or even only the best one, according to a specific metric. A probabilistic approach, such as the Probability Data Association Filter (PDAF) will perform *soft-assignment* [BL95]. Information about all measurement will impact the distribution leading the state space. The miss-detections are included in the update of the filtering distributions, escaping from the exponential complexity of the last method.

When multiple targets are present, several states evolve through steps, with possibility for each target to produce a measurement vector. If we assume n existing targets, with states $x_k^i, \forall i \in [\![1, n]\!]$, the measurement set can be expressed as:

$$Z_k = \cup_{i=1}^n T_k^D\left(\left\{x_k^i\right\}\right) \cup C_k \tag{2.3.24}$$

The Multiple Hypothesis Tracker (MHT) propagates different assignments with time, but needs a truncation to retain only the most relevant, due to the exponential growth of possibilities with steps [Cha+11]. The Joint Probabilistic Association Filter (JPADF) provide an estimation procedure similar to the single target PADF, if the exact number of target is provided [BL95]. However, it does not take

into account other phenomena which could affect the targets, like their *death*, or the appearance of new ones, called *target birth*.

The FISST theory aims at extending the notions of probabilities to sets [Mah07b]. It provides a mathematical basis to manage sets of vectors, which cardinalities can vary with time.

One notes $\mathcal{X}$ a compact subset of $\mathbb{R}^d$, and $\mathcal{F}(\mathcal{X})$ the collection of all the finite subsets of $\mathcal{X}$. A Random Finite Set (RFS) X on $\mathcal{X}$ is defined as a measurable mapping from a probabilistic space (Ω, A, P) to the collection $\mathcal{F}(\mathcal{X})$ [Mah07b]:

To make a link with previous subsection, $\mathcal{X}$ will act as the state space for individual targets ($x \in \mathcal{X}$).

$$X : \Omega \to \mathcal{F}(\mathcal{X}) \tag{2.3.25}$$

In other world, a RFS has stochastic cardinality, and its elements are random. We observe realizations of a RFS X.

Extending the concept of posterior distribution, it is possible to define state space models using RFS. With marginal distributions, one obtains multi-target filtering densities. Some important examples of RFS are:

Notation i.i.d for *independent identically distributed*

Example 2.3.7. *I.I.D Clusters Processes with*

- *p a probability density on $\mathbb{N}$*
- *f a probability density function on $\mathcal{X}$*

A set X of cardinality n, noted $X = \{x_1, .., x_n\}$, follows an i.i.d. process if its has $f(X)$ as its distribution, with:

$$f(X) = n!\, p(n) \prod_{i=1}^{n} f(x_i) \tag{2.3.26}$$

Example 2.3.8. *A Poisson process is a specific case of an i.i.d. clutter process, if p is a Poisson distribution with parameter λ:*

$$p(n) = e^{-\lambda} \frac{\lambda^n}{n!} \tag{2.3.27}$$

The density of the RFS is therefore

$$f(X) = e^{-\lambda} \lambda^n \prod_{i=1}^{n} f(x_i) \tag{2.3.28}$$

Example 2.3.9. *A RFS X following a Bernoulli process with parameter $r \in [0, 1]$ and probability density f defined on $\mathcal{X}$ has probability density*

$$f(X) = \begin{cases} 1 - r & , \quad X = \varnothing \\ r f(x) & , \quad X = \{x\} \\ 0 & , \quad |X| \geq 2 \end{cases} \tag{2.3.29}$$

Example 2.3.10. *A Multi-Bernoulli RFS $X = \bigcup_{i=1}^{M}$ is the union of M independent Bernoulli RFS X_i, each one defined by a parameter r_i and a probability density f_i. The probability density of $X = \{x_1, .., x_n\}$ is defined as :*

$$f(X) = n! \prod_{j=1}^{M} \sum_{1 \leq i_1 < .. < i_n \leq M} \frac{r_{i_1} p_{i_1}(x_1)}{1 - r_{i_1}} .. \frac{r_{i_n} p_{i_n}(x_n)}{1 - r_{i_n}} \tag{2.3.30}$$

The FISST theory allows to generalize the Bayesian recursion equation of filtering on multi-target densities. Noting $f_k(X \mid Z_{1:k})$ the multi-target filtering density at step k, the predicted multi-target density is obtained as:

$$f_{k+1|k}(X \mid Z_{1:k}) = \int f(X \mid X_k) f_k(X_k \mid Z_{1:k})\, dX_k \tag{2.3.31}$$

The filtering multi-target at step $k+1$ is then obtained using the update integration:

$$f_{k+1}(X \mid Z_{1:k+1}) = \frac{p(Z_{k+1} \mid X) f_{k+1|k}(X \mid Z_{1:k})}{p(Z_{k+1} \mid Z_{1:k})} \tag{2.3.32}$$

The denominator quantity

$$p(Z_{k+1} \mid Z_{1:k}) = \int p(Z_{k+1} \mid X_{k+1}) f_{k+1|k}(X_{k+1} \mid Z_{1:k}) \delta X_{k+1} \tag{2.3.33}$$

is called the Bayes factor. It is also interpreted as the likelihood of the measurement set Z_{k+1} given all previous measurement sets.

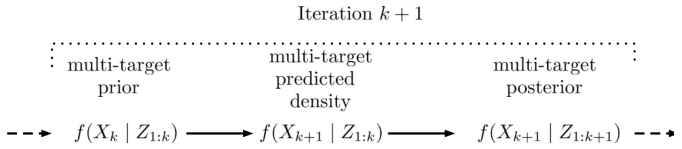

Figure 2.7: Multi-target density recursions

The operation scheme, shown in figure 2.7, is similar to the previous one, which use the classical recursion equations.

The previous operations 2.3.31 and 2.3.32 are not tractable in the general case, even with low number of targets. Less complex models or approximations have therefore been proposed.

The function D, defined on the single target space $\mathcal{X}$, called the *Intensity* or *Probability Hypothesis Density* (or PHD), is defined, for a multi-target density distribution $f(X)$, as [GMN97]:

$$\begin{aligned} D(\{x\}) &= \int_{Y \supseteq \{x\}} f(Y) \delta Y \\ &= \int_{Y \supseteq \{x\}} f(\{x\} \cup W) \delta W \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} f(\{x\} \cup \{x_1, .., x_n\}) dx_1 ... dx_n \end{aligned} \tag{2.3.34}$$

By extension, for any vector $x \in \mathcal{X}$, to lighten the notation, shortcut $D(x)$ is used instead of $D(\{x\})$.

An interesting property of this function is that, given a region S of $\mathcal{X}$, the integral $\int_S D_{k|k}(x)dx$ is the expected number of targets in S at step k. Moreover, we can define $N_{k|k} = \int_{\mathcal{X}} D_{k|k}(x)dx$ as the total expected number of targets at step k. It is indeed the first moment of the multi-target posterior of X_k, $f_{k|k}(X \mid Z_k)$ [Mah03].

Note that this function is not a probability density function, since it does not necessarily integrate to 1.

The *PHD filter* aims to provide the calculation of the intensity $D_{k|k}(x)$ associated to the intractable multi-target posterior for each time k. Its complexity is of order $\mathcal{O}(nm)$, with m as the number of measurement vectors, and n the number of targets.

Indeed, the expected number of targets is providing by the mean of the intensity function, but local information in a smaller region can also be calculated.

Appearance of targets at step k are modelled by a birth intensity γ_k. The recursions equations are given by [Mah13a]:

$$D_{k|k-1}(x) = \int p_S(y) f_{k|k-1}(x \mid y) D_{k-1}(y)\ dy + \gamma_k(x) \tag{2.3.35}$$

$$D_{k|k}(x) = [1 - p_D(x)] D_{k|k-1}(x) + \sum_{z \in Z_k} \frac{p_D(x) f(z \mid x) D_{k|k-1}(x)}{\kappa(z) + \int p_D(x) f(z \mid y) D_{k|k-1}(y)\ dy} \tag{2.3.36}$$

The intensity of the Poisson clutter is defined, for some clutter rate λ and distribution $c(z)$ as $\kappa(x) = \lambda\, c(z)$. A classical distribution used is the uniform law on an observation space volume V. It required consequently the definition of the limits of the observation space.

Equation 2.3.35 (the prediction equation) uses the probability of survival of each target, combined to their transition equation, with addition of births contribution. For equation 2.3.36 (the update equation), the left side is related to the target that are not detected by the sensors (missed detection). The right side mixes the contribution of detected targets and clutter.

The recursion equations are however hard to solve, without analytical expression for general equations. A tractable approximation using SMC methods has been proposed, which gives access to estimation of the intensity function [VSD05]. However, the state estimation requires more advanced techniques, to cluster of the particles. Gaussian Mixture Models fitting have been tested for this, but appeared less performant than the K-Means algorithms [CB07]. Because of the stochastic approximation with particles, the SMC-PHD filter shares the troubles of the Particle Filter: sample impoverishment and weight degeneracy, with need of resampling schemes.

But contrary to the Particle Filter, the weights of the SMC-PHD filter does not sum to 1, but to the expected number of targets $N_{k|k}$.

The PHD filter has been well studied both theoretically and practically. One important drawback is its instability over time. Since only the intensity is propagating through time, it is prone to have high variance over the number of targets with time. The Cardinalized PHD filter (CPHD filter) has been proposed to bring more stability to the estimation of the target number [Mah07a], for a slight increase of the algorithm complexity ($\mathcal{O}(nm^3)$ instead of $\mathcal{O}(nm)$ for the PHD filter). Beside the intensity, it also propagates the probability density of the cardinality of targets trough time. Analytical resolutions for recursions equations 2.3.35 and 2.3.36, using Gaussian mixtures models, has allowed development of deterministic versions of the PHD and CPHD filters, respectively the *Gaussian Mixture PHD* (GM-PHD) [VM06] and *Gaussian Mixture CPHD* (GM-CPHD) [VVC07] filters. The PHD takes the form of a sum of Gaussian components, which weights sum to the expected number of targets.

2.3.4 *Random Finite Set Bayesian tracker*

Multi-target filters provide for each time step information about the state set X_k. However, they do not provide any link between a specific target at step k and one other at step $k+1$: we do not have any association of targets through steps. The purpose of *trackers* is to provide such information. A multi-target tracker should give an estimate of the individual tracks for each single target through steps.

Many heuristics have been developed to extend the power of multi-target filters, to produce tracks. Although they can be fast and simple to implement, they do not provide any theoretical caution.

A simple heuristic to provide track estimates is to assign to each Gaussian component of the GM-PHD Filter a *tag*, which will be propagated through steps [PCV09]. While this seems simple to implement, attention must be paid to the tag management and track continuities (example: when target paths are crossing each other). From a more general point of view, a multi-target filter can be employed jointly with an association procedure, applied on the filtered states to create the tracks.

As noted from [Che18], the definition of *track* depends on the observer point of view:

- From the filtering point of view: it consist in constructing an association of filtered sets at each step k, without ability to modify the previous associations.
- From the posterior point of view, the associations can be performed globally, after reception of all the measurement sets from the sensors. Here, the online aspect is lost. Moreover, the exploration of the set of possible associations appears to be intractable for high numbers of measurement vectors in each set, or long inspection.

Even with no clutter, no death, birth or miss-detections, if Z_k and Z_{k+1} have cardinality n, the number of possible assignment is $n!$.

The online problem has also been treated with a particle filter, using as particles long vectors containing label associations for each track and their states [SVL07]. An offline approach with a fully Bayesian algorithm has been proposed, which appeared to outperform the GM-PHD filter, at the price of a costly calculation time [VVE11].

The association history of tracks can indeed be viewed as a long parameter vector, which posterior distribution is searched.

The tracking problem has been treated with RFS with some modifications. It is necessary to define the vectors on an extended space $\mathcal{X} \times \mathbb{L}$. The denoted state is noted $\mathbf{x} = (x, l)$ with $x \in \mathcal{X}$ and $l \in \mathbb{L}$. The label space $\mathbb{L}$ allows us to move from a filtering to a tracking procedure. A wanted property is that, for a specific identified target, one unique label will be related to it.

An extension of the PHD filter for labelled RFS has been proposed [Gom18], with performances superior to the GM-PHD filter with tags.

The delta Generalized Labelled Multi-Bernoulli (δ-GLMB) filter is the first Bayes multi-object tracker. It employs the Bayes equations to propagate uncertainty about the states and their labels.

The distinct label indicator is defined, for a subset $\mathbf{X}$ of $\mathcal{X} \times \mathbb{L}$, as the quantity $\Delta(\mathbf{X})$ which equals 1 if and only if the cardinality of the set of extracted labels from $\mathbf{X}$, noted $\mathfrak{L}(\mathbf{X})$, equals the cardinality of

And $\mathfrak{L}(\mathbf{x}) = l$.

$\mathbf{X}$. It ensures that all the labels of the states vectors of $\mathbf{X}$ are different:

$$\Delta(\mathbf{X}) = \delta_{|\mathbf{X}|}(|\mathfrak{L}(\mathbf{X})|) \tag{2.3.37}$$

The core principle is the formulation of a δ-GLMB RFS $\mathbf{X}$, with related density:

Using notation $p^X = \prod_{x \in X} p(x)$ with $p^{\varnothing} = 1$.

$$\pi(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \omega^{(I,\xi)} \delta_I(\mathfrak{L}(\mathbf{X})) \left[p^{(\xi)}\right]^{\mathbf{X}} \tag{2.3.38}$$

This density is formed of multiples *hypotheses*, each characterized by an tuple (I, ξ), with I a subset of labels from $\mathbb{L}$, supposed to exist for this hypothesis, an association history ξ, a positive weight $\omega^{(I,\xi)}$, and the probability densities related to the existing targets $p^{(\xi)}$, which depend on the association history, and are functions of both a vector from the state space and a label. Conditions on the set is that :

$$\sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \omega^{(I,\xi)} = 1 \tag{2.3.39}$$

The cardinality distribution is calculated as:

$$p(n) = \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \omega^{(I,\xi)} \delta_n(|I|) \tag{2.3.40}$$

The most important property is that this distribution has a closed form under the Bayes recursion. The set of new targets born at current step is distributed according to a labelled Multi-Bernoulli (LMB) RFS with label space $\mathbb{B}$.

A LMB RFS has density of the form (see [Reu+14]):

$$\pi(\mathbf{X}) = \Delta(\mathbf{X}) w\left(\mathfrak{L}(\mathbf{X})\right) \left[p^{(\xi)}\right]^{\mathbf{X}} \tag{2.3.41}$$

but can simply be described as a set of probability densities associated with a probability of existence $r \in [0, 1]$, or more explicitly $\left\{(r^{(l)}, p^{(l)})\right\}_{l \in \mathbb{B}}$.

With $\mathbb{L}_+ = \mathbb{L} \cup \mathbb{B}$, the filtering density has then the form:

$$\pi_+(\mathbf{X}) = \Delta(\mathbf{X}) \sum_{(I_+,\xi) \in \mathcal{F}(\mathbb{L}_+) \times \Xi} \omega^{(I_+,\xi)} \delta_{I^+}(\mathfrak{L}(\mathbf{X})) \left[p_+^{(\xi)}\right]^{\mathbf{X}} \tag{2.3.42}$$

The updated density has the form:

$$\pi(X \mid Z) = \Delta(X) \sum_{(I,\xi) \in \mathcal{F}(\mathbb{L}) \times \Xi} \sum_{\pi \in \Pi(I)} \omega^{(I,\xi,\pi)} \delta_I(\mathfrak{L}(X)) \left[p^{(\xi,\pi)}(\cdot \mid Z)\right]^X \tag{2.3.43}$$

with $\Pi(I)$ the set of all association possible between measurements and targets existing in I.

To obtain the set estimate, the MAP of the cardinality (equation 2.3.40) is estimated [VVP14]. Next the hypothesis related to this cardinality with higher weight can be used as a MAP estimate for the set [Pun17].

In practice, to implement these recursions, it is nevertheless necessary to perform truncations during the prediction step, and during the update step, by retaining only the predicted and updated hypotheses with higher weights. However, a faster implementation is possible, reducing also the loss of information induced by two successive trun-

cations. Indeed, the exact update filtering density can be directly written:

$$\pi_{Z^+}(X) = \Delta(X) \sum_{(I,\xi,I^+,\pi_+)} \omega^{(I,\xi)} \omega^{(I,\xi,I^+,\pi_+)} \delta_I(\mathfrak{L}(X)) \left[p_{Z^+}^{(\xi,\pi_+)} \right]^X \tag{2.3.44}$$

Therefore, after enumeration of all possibilities, one single truncation step is made. This operation saves the number of operations to perform at each step. Benefit is double: in terms of efficiency and speed. This technique was first dedicated to the LMB filter [Reu+14], an approximation of the δ-GLMB filter, but was next applied to it [HVV15].

Although the previous formulas look complex, in practice, it is sufficient to understand that this filter is only composed of sets of probability densities attributed to labels, which interacted in the previous steps with different associations of measurements. The history of associations is indeed not really important for sequential filtering purposes. Of course, this filter outperformed the classical PHD or CPHD filters, but is more demanding, with cubic complexity during prediction and update phases [Reu+13b].

In reality a tracker, or filter and labelled RFS

Using an assignment algorithm to select truncation, and using Gaussian distributions, the full filter stays completely deterministic. This is the version of the tracker used in this thesis to perform associations between states through time.

2.3.5 *Metric for comparison of sets*

For state space approaches with vectors, comparisons between estimated vectors and their true values are calculated with indicators based on some distances, depending on the problem. A common choice is for instance the MSE. Quantification of performances of filtering method is obtained with resulting scores. However, comparison between two RFS is far more complex. First, their cardinalities do not necessarily coincide. Secondly, the FISST theory does note allow addition operations between sets. Distance definition for RFS using vector distances is not trivial.

To compare estimated sets resulting from RFS filters, a specific metric has been designed, the *Optimal SubPattern Assignment* (OSPA) metric [SVV08]. It introduced two additional variables: the *cut-off* parameter $c \in \mathbb{R}^+$, and the order $p \geq 1$. For two vectors x and y, an intermediate metric is calculated, and defined as:

$$d^{(c)}(x,y) = \min(c, d(x,y)) \tag{2.3.45}$$

with d any valid metric (for instance the Euclidean ones).

Euclidian distance between vectors x and y is defined as:

$$d(x,y) = \|x-y\|_2^2 \tag{2.3.46}$$

For two sets $X = \{x_1, .., x_n\}$ and $Y = \{y_1, .., y_m\}$ with respective cardinalities n and m, such that $n \leq m$ and $n, m > 0$, the OSPA

distance between X and Y is defined as:

$$d_p^{(c)}(X,Y) = \left[\frac{1}{m}\left(\min_{\pi\in\Pi_m}\sum_{i=1}^{n} d^{(c)}\left(x_i, y_{\pi(i)}\right)^p + c^p(m-n)\right)\right]^{\frac{1}{p}} \tag{2.3.47}$$

with Π_m the set of permutations on $\{1,..,m\}$. If $n=0$ or $m=0$, the OSPA distance between X and Y is defined as $d_p^{(c)}(X,Y)=0$. For the last, where $n,m>0$ but $n>m$, the distance is defined as $d_p^{(c)}(X,Y)=d_p^{(c)}(Y,X)$.

As indicated by its name, the OSPA metric is a metric (it fulfils the conditions of identity, symmetry and the triangle inequality). It induces the calculation of a minimal cost over possible permutations between vectors, and requires therefore use of a minimal assignment algorithm. Parameter p acts as a level of penalization for outliers vectors: for high values of p, the distance will give more weight to extreme values far from the others in the sets. Parameter c is more difficult to interpret, but could be viewed as a cap value. OSPA distance nears 0 means that X and Y are close, and an OSPA distance near c is the worst situation, with X and Y highly dissimilar.

2.4 Conclusion

Purpose of this chapter was to introduce the techniques typically dedicated to ultrasonic measurement analysis. NDT processing focus on identification of physical features, which allow a diagnostic about the condition of the inspected medium. However, their sensitivity to perturbations prevents their direct use for field measurements. Specifically dedicated to complex measurement, a sparse decomposition scheme has been presented for extraction of sets of features from EMAT signals. However, this new representation does not allow an understanding of evolution of signals with time. To include this element, essential for mobile inspection, state space models offer an efficient and elegant answer. Filters are statistical objects which associate information with time. For vector states, they estimate hidden information, not directly accessible from measurements, and provide quantification of the uncertainty relative to this estimation. Their complexity depends on made assumptions about the evolution of state, notably their linearity. For multi-object representations, the FISST theory led to the design of advanced filters based on RFS, notably the PHD and δ-GLMB filters, which allow estimation on sets with uncertainty related to appearance or disappearance of objects. Creation of tracks, association of states across time, is thus made possible. They offer a promising perspective for joint use with the latter sparse representation methodology.

Statistical characterization of ultrasonic signals for NDT evaluation of rails

3

3.1 Introduction

Purpose of this chapter is to combine statistical methods to improve characterization of ultrasonic signals.

For acquisition k of the dataset (a total of N acquisitions are assumed obtained), the MP algorithm finds a Gabor vector set Θ_k with unknown cardinality. This set, in this chapter, is assumed to act as a measurement set:

$$Z_k = \Theta_k \tag{3.1.1}$$

Objective is thus to build a state space representation, which includes information about the evolution of the ultrasonic pulses, and estimate at each step k the hidden state vectors, or in other terms, provide an estimate about the state set X_k which conducts measurement set Z_k, from previous data.

In section 3.2, a new methodology to solve the complete parameter estimation problem of a general multi-target filtering and tracking process is presented, based on Bayesian processes. No specific a priori knowledge is assumed from the operator, and our process could therefore be applied on any practical situation, with attention put on the reasonable tractability of the problem, in terms of accuracy and computational resources limitations.

Section 3.3 demonstrates first the statistical validity of approximations made in the methodology. Simulations are next conducted to assess the performances in terms of accuracy and convergence behaviour.

Section 3.4 describes the main model formulation proposed to solve the ultrasonic pulse tracking problem. Laboratory experiments were performed to create situations of interest, in which the filters had to extract physical information from signals. With respect to the Gabor representation of ultrasonic signals, new models are elaborated.

Section 3.5 applied the methodology on laboratory experiments conducted with mobile sensors. Control and design of test confirms the utility of the RFS formulation before application in field rail inspection.

Section 3.6 described the mobile structure employed for field experiments.

Last section 3.7 applied the methodologies on measurements obtained on rails with head-checking at Eurotunnel.

3.2 Parameter tuning

3.2.1 *Parameters for sparse decomposition*

The general methodologies for decomposition of a signal with a Gabor dictionary has been used. However, the number of iterations to perform has to be chosen, depending on the signals to analyse. With a different formulation this issue refers to the selection of the stopping criterion of the decomposition recursion. In this section, the MP algorithm is applied to data coming from our measurements on the experimental rails. For improved performances, the SAGE algorithm is applied on all the atoms found at a specific iteration of the MP algorithm. The soft rule is applied for calculation of residuals (equation 2.2.15). To measure performances of decompositions, a specific metric is introduced for this section. Let y be the signal, decomposed into k atoms $\theta_1, .., \theta_k$ at step k of the MP algorithm. We defined the *Normalized Mean Square Error* (NMSE) at step k (NMSE_k) by the ratio between the energy of the residual signal using the k atoms (energy of the difference between y and its reconstruction using the available atoms) and the energy of the original signal y:

$$\text{NMSE}_k = \frac{\|y - \sum_{i=1}^{k} f(\theta_i)\|_2^2}{\|y\|_2^2} \tag{3.2.1}$$

Expected behaviour is to obtain a fast decrease of the curve with iterations of the MP algorithm. When the NMSE trend reaches a slow slope, every new atom brings little information, and efforts are lost in computational resources to improve the decomposition. For one set of acquisitions (for the $1\,m$ and $70\,cm$ rails), the score is averaged over all measured signals. A total of 15 iterations are performed to see the evolution of the score.

Figure 3.1 highlights the fast decay of the NMSE on the $1\,m$ rail, which does not content perturbation. The averaging over many acquisitions has removed random noises. After 5 iterations of the MP algorithm, the score has already reached a quasi-constant trend.

Observations of the $70\,cm$ rail, with electro-erosion defects show similar results. Even if the decrease during the first steps is weaker, after iterations 5 and 6, no clear improvement is found.

These results allowed us to select a proper stopping rule for the decomposition recursion. A threshold on the NMSE is set to $5\,\%$, to allow extraction of multiple ultrasonic pulses. In the case of noisy signals, after estimations of the atoms of interest, the MP algorithm could try to focus on the signal noise, producing many iterations to reach this threshold, with little improvement of the decomposition. To prevent this behaviour, and with the empirical knowledge that only few pulses are present in the signal, a maximum number of iterations is set to 10.

Our main objective is to extract a huge number of points from the measurements. Noise and perturbations will sometimes be kept by the decomposition method. However, application of filtering methods

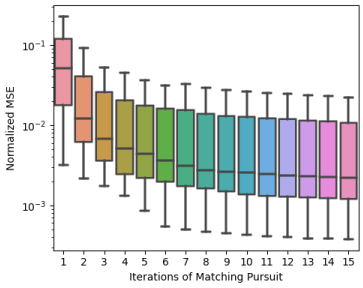

Figure 3.1: Evolution of the Normalized Mean Square Error of the decomposition with iteration of the Matching Pursuit algorithm for the $1\,m$ rail, with errors bars computed with all acquisitions.

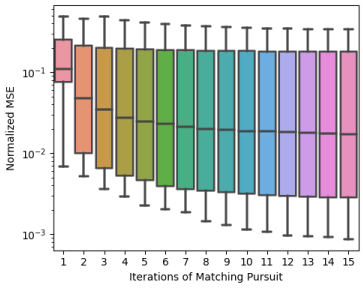

Figure 3.2: Evolution of the Normalized Mean Square Error of the decomposition with iteration of the Matching Pursuit algorithm for the $70\,cm$ rail, with errors bars computed with all acquisitions.

allows separation of atoms coming from ultrasonic pulses from these disturbances, which will be considered as clutter.

3.2.2 *Processing of unknown parameters in RFS filtering*

As many estimation procedures, filtering processes are conducted by parameters, which are assumed to be known *a priori*. Outputs of filters (let them be state spaces, tracks or cardinality estimations) depend on the initial values set by the operator, relying on its knowledge about the phenomenon inspected. Using parameters which do not reflect the reality could yield degenerated results or lack of estimations (for example, no tracks in the case of RFS filtering). Some algorithms have proven robustness against wrong parametrization, but even approximate guess can be hard to find in high dimension, when the number of parameter increases [YJH17].

Parameter inference is the topic dedicated to the estimation of parameters, or parameter vector ϑ, from data $\mathcal{D}$. Goal is either to provide a point estimate, or to quantify the uncertainty about parameters. This information is next included into the previous procedure to get the vector (or RFS) of interest x.

Point estimation (figure 3.3) involves the quantification of a cost related to a parameter vector to be optimized. The obtained vector $\hat{\vartheta}$ is next interpreted as the *a priori* known vector guiding the phenomenon.

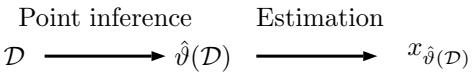

Figure 3.3: Estimation procedure with point parameter estimator.

In Bayesian inference, this quantification takes the form of a probability density $p(\vartheta \mid D)$, known as the posterior distribution of ϑ, which is next used (figure 3.4) for calculation of the expectation of x.

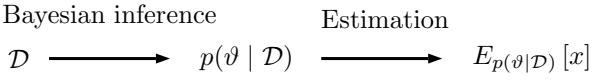

Figure 3.4: Estimation procedure with Bayesian inference.

In the case of RFS filtering, every procedure aiming at solving the parameter inference problem faces the issue of the high number of parameters to estimate. In addition to the ones coming from the state space model, the ones managing the evolution of targets appear: survival, death, occlusion and clutter.

Some partial solutions have already been studied, for the following parameter variables:

- Clutter (no dependence across steps): the Poisson model with uniform distribution on the observation space can be dealt with MLE when number of clutter points is assumed to largely exceed the cardinality of targets, from the Point Process theory. For a constant value of the Poisson clutter distribution λ across measurements, non-uniform distributions can perform better in situation

Estimation gives $\hat{\lambda} = \frac{1}{N}\sum_{k=1}^{N}|Z_k|$.

where clutter non-homogeneously spreads across the observation space. MLE procedure using EM formulations has been successfully tested with the specific case of mixtures of parametric distributions, for estimation of the clutter intensity parameters (both its λ and its related distribution). The PHD filter allows calculation of the derivative of the likelihood, which eases the minimization procedure [LHL10]. No specific hypothesis is made about the clutter intensity, but all other parameters are assumed known.

Mixture of Gaussian distributions with and uniform distribution which acts as background noise is a classical example [LHL10].

- Clutter (step dependent): in this situation, the clutter intensity will evolve across steps. Generally, in the literature, only the case of the variation of the intensity parameter λ is addressed. This allowed the development of robust algorithms which are able to adapt to different practical situations without further tuning. The model by itself integrates the uncertainty about this parameter, and provides, in specific case, a step dependent estimation. The idea is to extend the state space $\mathcal{X}$ with an additional state space $\mathcal{G}$ [MVV11; Che+12; Mah12], which contain clutter generators, each of which can produce a clutter measurement. These generators are then modelled with birth, survival and detection equation, like regular targets. Statistics made with these specific states provide information about the evolution of clutter across steps. Implementations have been proposed for the PHD, CPHD and CBMemBeR filters [Vo+13]. However, efforts are moved on the selection of the clutter generator models, adding new parameters to be estimated.

Augmented space $\bar{\mathcal{X}} = \mathcal{X} \times \mathcal{G}$.

- Probability of detection : similarly to the step-dependent clutter, this parameter can also be a hidden information. An new extension of the state space to include the probability of detection tracks or clutter generators allows the modelling of its uncertainty. Assuming a beta prior for the detection probability, and Gaussian models for the tracks and clutter generator, an analytical form of the CPHD filter is available [MVV11].

A beta distribution with parameters $a, b \in \mathbb{R}^{+*}$ has density defined by

$$f(x \mid a, b) = C(a, b)x^{a-1}(1-x)^{b-1}$$

on $[0, 1]$, with $C(a, b)$ the normalization constant.

- Target birth : the model used for the RFS filters assumes a prior model of birth intensity. However, current information about available measurement are not used to create new tracks, but to confirm or not the ones already born and created by the birth model. To save computational resources, it seems more interesting to initiate new tracks in the neighbourhood of existing measurements. Nevertheless, the basic transition equation of RFS filtering must not be applied directly, since a double use of information (for both creation of tracks and update of the density of PHD) would be performed. An efficient solution has been proposed for the PHD and CPHD filters [RCV10; Ris+12]. Using a mark $\beta \in [\![0, 1]\!]$ which augments the state space, with $\beta = 1$ denoting a newborn target, and $\beta = 0$ an existing one, it is possible to build an adaptive filter which uses measurements in its birth intensity model. Prediction and update equations are modified to include this change. A target born at step $k-1$ evolves into an existing one at step k. Moreover, a newborn

For a Gaussian Mixture birth intensity, the number of required components increased drastically with both the support of the targets and the dimension of the state space.

target at step k is necessarily detected ($p_D(x \mid \beta) = 1$ if $\beta = 1$). Measurements allow to reconstruct one part of the state vector, the other part remains uncertain and modelled by probability distributions. An approximate analytical form is still available for the PHD and CPHD filters, assuming an uniform probability distribution for the birth intensity model [HL10; Bea+12]. A different version was proposed for the LMB and δ-GLMB filters. Since birth are modelled by a LMB distribution, current measurements (set Z_k at step k) are involved in the target birth at step $k+1$ [LVN16]. Measurements which have been associated with a low number of targets inside hypothesis are more prone to create new tracks.

Augmenting the state space with such variable is indeed equivalent to considering two models for targets, and applying the multiple model recursion equations.

The generic case (ϑ contains combinations of the above cases, or other variables), has been only studied for the PHD filter. Following the theory developed for particle filtering, especially the Liu and West filter [LW01], an adaptive multi-target filter has been designed for accommodating abrupt changes of parameters, thanks to propagation of sufficient statistics related to the uncertainty on their values [Yan+16]. A different approach is to propagate multiple PHD filters, to compute at each step the likelihood of a set of parameter vectors $\{\vartheta_1, .., \vartheta_K\}$. One PHD filter is assigned to one of theses parameters. An additional step estimate the instantaneous parameter (for instance by selection of the highest obtained likelihood) [Sch+16]. A drawback of this method it the cost of running a massive number of filters. Moreover, no specific model is attributed to the parameter vector variations. A link with uncertainty quantification on parameters is not straightforward.

3.2.3 *Bayesian inference*

For Multi-target tracking, the obtained data takes the form of a measurement set collection, noted $Z_{1:N}$. Calculation of the posterior density $p(\vartheta \mid Z_{1:N})$ is based on the Bayes' rule:

$$p(\vartheta \mid Z_{Z_{1:N}}) = \frac{p(Z_{1:N} \mid \vartheta)p(\vartheta)}{p(Z_{1:N})} \tag{3.2.2}$$

with calculation of the *Bayes normalization factor* (or *evidence* of data $Z_{1:N}$):

$$p(Z_{1:N}) = \int p(Z_{1:N} \mid \vartheta)p(\vartheta)\, d\theta \tag{3.2.3}$$

The prior choice $p(\vartheta)$ is operator dependent, relying on previous experimentations, or knowledge of the admissible bounds for values of ϑ. A common non-informative prior is the uniform law over a bounded domain (ex: $p(\vartheta) = \mathcal{U}_{[\vartheta_{\text{inf}}, \vartheta_{\text{sup}}]}(\vartheta) = \frac{1}{\vartheta_{\text{sup}} - \vartheta_{\text{inf}}}$ in the $1D$ case).

For sampling from the distribution, following notation is preferred: $\vartheta \sim \mathcal{U}([\vartheta_{\text{inf}}, \vartheta_{\text{sup}}])$.

Unfortunately, in the case of multi-target filtering and tracking, no close form is directly available for the posterior distribution. Even worse, no expression for the Bayes normalization factor can be found in the general situation of a time-varying number of targets.

A solution of this common problem in Bayesian inference, when the likelihood can still be evaluated, is to use a Monte Carlo Markov Chain

(MCMC) algorithm. Purpose is to generate a sequence of samples, called a *chain*, whose stationary distribution converges to the posterior of interest [RC04].

The Metropolis Hastings (MH) algorithm only requires an access to the likelihood of the data, conditioned to a parameter sample ϑ, and to the prior. It generates a sequence $\{\vartheta_k\}$ approximately distributed from $p(\vartheta \mid Z_{1:N})$ by exploring the space with a proposal distribution q. It requires the marginal likelihood $\mathcal{L}_\vartheta(Z_{1:N})$, noted $\mathcal{L}_\vartheta$ to be evaluated, and the design of a proposal distribution q, which will master the sequence. The principle is indeed very simple: starting from a value ϑ_k, a proposal value ϑ^* is sampled from $q(\vartheta \mid \vartheta_k)$. After evaluation of the marginal likelihood for this specific value, we keep it with probability $\alpha(\vartheta^*, \vartheta_k)$, with the later defined as the *acceptance ratio* (or *acceptance probability*):

$$\alpha(\vartheta^*, \vartheta_k) = \frac{\mathcal{L}_{\vartheta^*}}{\mathcal{L}_{\vartheta_k}} \frac{p(\vartheta^*)}{p(\vartheta^k)} \frac{q(\vartheta_k \mid \theta^*)}{q(\vartheta^* \mid \vartheta_k)} \tag{3.2.4}$$

If the proposed value is not accepted, the value ϑ_k is kept for this iteration. Therefore, a proposed value with low likelihood can still be selected (although with low probability). The amount of accepted values over the total proposed in the chain (latest equals thus the length of the chain) is defined as the *acceptance rate*. For a good exploration of the space, this rate should neither be near 0 (all samples are rejected), nor near 1 (nearly all samples are accepted, and exploration is thus very low). Heuristically, the proposal is designed to get a rate between 0.23 and 0.44. It has indeed been shown that the optimal rate for an unidimensional Gaussian is around 0.44, and for higher dimensions, this rate reach the value of 0.23 [Bro+11]. However, as soon as the limit values of 0 and 1 are avoided, the chain will still converge, but with perhaps slower mixing (exploration of the space) [LLC10].

Although simple, this algorithm allows powerful inferences, and results ensured by strong properties.

3.2.4 *Practical considerations for MCMC sampling*

An interesting choice for the proposal distribution is a Gaussian centred at the current values ϑ : $\vartheta^* \sim \mathcal{N}(\vartheta, \Sigma)$ for some covariance matrix Σ. This proposal is said to be *symmetric*, since $q(\vartheta^* \mid \vartheta) = q(\vartheta \mid \vartheta^*)$ if Σ is independent of ϑ. Main benefit is to re-write the acceptance ratio as :

$$\alpha(\vartheta^*, \vartheta_k) = \frac{\mathcal{L}_{\vartheta^*}}{\mathcal{L}_{\vartheta_k}} \frac{p(\vartheta^*)}{p(\vartheta^k)} \tag{3.2.5}$$

This specific case of the MH algorithm is also known as the *random walk* MH.

However, a critical point is the starting of the chain, coupled with choice of the proposal covariance matrix. A possible solution is to let the MCMC explore the state space to draw samples near the areas of higher density, and discard those preliminary steps. This is called the *burn-in* or *warm-up* period. An other strategy, not to lose numerous

points, is to start from the posterior mode, and use a modification of its variance $\widehat{\mathrm{Var}}(\vartheta)$. The proposal covariance matrix is then defined as [She+15]:

$$\Sigma_{\mathrm{prop}} = \gamma^2 \frac{2.56^2}{d} \widehat{\mathrm{Var}}(\vartheta) \tag{3.2.6}$$

Not to confound this parameter with the birth intensity of RFS theory. Use of γ is only specific to this formula.

with γ a tuning parameter which equals 1 for a Gaussian posterior, and d the dimension of the vector ϑ. To estimate $\widehat{\mathrm{Var}}(\vartheta)$, if calculation time is not an issue, a long preliminary MCMC chain can be run (possibly not mixing well). Numerical differentiation using the log-likelihood at the mode is also possible, is an estimation of this mode is provided.

A different approach is to design a step adaptive proposal covariance matrix, which will evolve during the sampling procedure, to reach a suitable acceptance rate. Taking information from the previous samples, the covariance matrix is updated. Numerous strategies have been designed, depending on the nature of the proposal and practical considerations [AT08]. In particular, a robust algorithm has been created to handle posterior distributions with heavy-tails, without more restrictions than the previous adaptive algorithms [Vih12].

An important challenge in sampling with MCMC is the assessment of simulation performances. An operator could expect his chains to mix, combined with exploration of the whole space in case of multimodality of the posterior distribution. First aspect, mixing, can be analysed via inspection of the relationship of consecutive samples. If exploration is very slow (proposal covariance too small), or if the chain gets stuck at one point (low acceptance rate), correlation between samples will be high on intervals. Looking at the *autocorrelation* function is therefore a good indicator to measure mixing. Notably, if ϑ is a unidimensional variable, with $\{\vartheta_i\}_{i=1}^K$ the vectors chain and $r_t\left(\{\vartheta_i\}_{i=1}^K\right)$ the autocorrelation function of the chain at lag t, the Integrated Autocorrelation Time (IACT) is defined as 1 plus the area of the autocorrelation function:

The acceptance rate can locally be slow because of oversized proposal covariance matrix. Nearly all samples are then rejected

$$\mathrm{IACT}\left(\{\vartheta_i\}_{i=1}^K\right) = 1 + 2\sum_{t=1}^{+\infty} r_t\left(\{\vartheta_i\}_{i=1}^K\right) \tag{3.2.8}$$

An alternative indicator is the *Effective Samples Size* (ESS), defined as:

$$\mathrm{ESS} = \frac{K}{\mathrm{IACT}} \tag{3.2.7}$$

Mixing properties between two MCMC chains is made possible with this indicator, and the one minimizing the IACT preferred.

One other important test is to run multiple chains, and compare them with a score. The Potential Scale Reduction Factor (PSRF) has been designed to compare variances between and within chains [GR92]. It provides an unique scalar $\widehat{R} \geq 1$. Good mixing chains should give a score near the optimal value of 1. It can also be used during the sampling, as a stopping rule to continue the chain until a delimited value of $\widehat{R}$ is obtained. This score was original shaped for univariate variables [GR92], but a multivariate version has been proposed thereafter [BG98]. A simpler solution is to compute the score for each dimension, and stop only when all scores are close to 1 [Bro+11].

3.2.5 *Optimal parameter search with Bayesian Optimization*

Running a MCMC chain, despite the amount of information it brings about the parameter posterior, is a long task to perform. Its sequential nature makes it hard to speed up. Moreover, manipulation of RFS does not allow to compute an equivalent of the *a posteriori* expectation of the filtered states, since RFS can not be added together. A simpler strategy would be to find only one parameter vector to be used for the filtering or tracking processes, and get the resulting estimation of the states or the tracks, conditioned to this specific parameter vector. Point inference for PHD filters has been tested only for the SMC method. Optimization is performed by maximization of the likelihood or log-likelihood of the measurement sets, computed recursively in parallel of the filtering process. Therefore, estimation time of $\mathcal{L}_\vartheta(Z_{1:N})$ for a specific ϑ is sensitive to the complexity of the employed algorithm, number of targets and density of clutter. Search for the optimal, or approximately optimal point, is prone to take a huge amount of time. In this part, we present a methodology to provide an estimate of the parameter vector which maximizes the measurement likelihood with few evaluations of the objective function (and thus few filtering runs).

A powerful methodology for maximization or minimization with a costly or noisy objective function is the Bayesian Optimization (BO) scheme. Since the objective function is costly to evaluate (in terms of computational resources or times), a surrogate model is build to fit the available evaluation of the objective function [Sha+16; Ber+11]. A probabilistic model is generally employed, to include a measurement of the uncertainty about fitting, and possible prediction in the parameter space. Optimization is however not performed directly on this approximation function. It is included in an acquisition function, which makes a balance between exploitation of the known regions of the parameter space, and exploration of the space which could also result in better parameters. Various acquisition functions have been designed: the Probability of Improvement, methods based on the Upper Confident Bound for instance. A well-studied function is the Expected Improvement, which has been shown to outperform other methodologies, but keep an intuitive form [SLA12].

Problem formulation is: the deterministic function to maximize is written $\tilde{y}$, which depends on the variable $x \in \mathbb{R}^d$, with d the dimension of x. Some data has already been acquired, taking the form of a set of n evaluations $\{y_i\}_{i=1}^n$, with $y_i = y(x_i), \forall i = 1, .., n$ for a set of observations $\{x_i\}_{i=1}^n$. Data is summarized as the set $\mathcal{D} = \{y_i\}_{i=1}^n$. The function y is the stochastic version of $\tilde{y}$, to include non-perfect measurements. Based on the maximum of available evaluations $y^{\max} = \arg\max \{y_i\}_{i=1}^n$, a realization of the *Improvement* for new observation x is defined as [JSW98]:

for example, each evaluation can be perturbed by a random noise ϵ_i, such that $y(x_i) = \tilde{y}(x_i) + \epsilon_i$.

$$\mathrm{I}(x) = \max\left(y(x) - y^{\max}\right) \tag{3.2.9}$$

The *Expected Improvement* (EI) is defined as the expectation of the Improvement over y. If the distribution of y depends on hyperparam-

eter vector ρ and data $\mathcal{D}$, it is defined as:

$$\mathrm{EI}(x) = \int \max\left(y(x) - y^{\max}\right) p(y \mid x, \mathcal{D}, \rho)\, dy \tag{3.2.10}$$

Inside the BO loop, from a set of previous parameter points, the next one to evaluate is selected by maximizing the EI over the parameter space. Indeed, the EI maximization procedure is assumed to be largely lighter than optimizing the objective function.

Choice of surrogate function is critical, since it conditioned the tractability of the EI calculation. Either parametric and non-parametric approximation have been used for this task. A common choice which coupled well with this acquisition function are Gaussian Processes (GP). This non-parametric methodology allows modelling of distributions of functions, from a Gaussian prior [RW06]. The posterior distribution, conditioned to a dataset $\{(x_i, y_i)\}_{i=1}^{n}$ for observations $x_i \in E$ and targets $y_i \in \mathbb{R}$, allows inference on a new observation x^*, giving its related expected target value $\bar{y}(x)$, with expression of the linked variance $\sigma(x^*)$. Formally, a symmetric positive function K is used to model the covariance of the GP. If the vector y gathers all acquired targets, and the matrix X the related observations, a new observation x^* will have predicted target y^*. The joint vector $[y, y^*]^T$ is assumed to follow the prior:

A 0 mean is generally assumed to ease calculation but any $m \in \mathbb{R}^n$ is admissible.

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X,X) & K(X,x^*) \\ K(x^*,X) & K(x^*,x^*) \end{bmatrix}\right) \tag{3.2.11}$$

Consequently, conditioning to this joint prior, following properties results [RW06]:

To lighten the notation, reference to ρ and D is dropped for these two variables, but should not be forgotten.

$$p\left(y^* \mid X, y, x^*\right) = \mathcal{N}\left(y^* \mid \bar{y}(x^*), \sigma^2(x^*)\right) \tag{3.2.12}$$

with

$$\bar{y}(x^*) = K(x^*, X)K(X, X)^{-1}y \tag{3.2.13}$$

$$\sigma^2(x^*) = K(x^*, x^*) - K(x^*, X)K(X, X)^{-1}K(X, x^*) \tag{3.2.14}$$

The analytical form of the EI, which also depends on hyperparameter vector ρ, is described in this specific situation by:

$$\mathrm{EI}_\rho(x) = \left(\bar{y}(x) - y^{\max}\right)\phi\left(\frac{\bar{y}(x) - y^{\max}}{\sigma(x)}\right) + \sigma(x)\varphi\left(\frac{\bar{y}(x) - y^{\max}}{\sigma(x)}\right) \tag{3.2.15}$$

An interesting property of the GP modelling is the possibility to automatically select its hyperparameters, by selection of the vector ρ which maximizes the marginal likelihood. The acquisition function is thus $x \mapsto EI_{\rho^{\mathrm{MLE}}}(x)$.

A Bayesian solution of the equation is also possible through marginalization over values of ρ [SLA12]. Under a prior $p(\rho)$ over hyperparameters, the posterior density $p(\rho \mid \mathcal{D})$ is generally not known in a close form. Running a MH algorithm to produce P values targetting this distribution is more tractable. To bypass the fastidious tuning aspects,

Notes: we are here in the context of the inference from the sets of the filter's parameters, and data are the corresponding value of the objective function (log-likelihood of the measurements). Marginalization is performed over the latent functions f according the the GP modelling.

a Slice sampling algorithm have been proposed to perform efficient sampling of GP hyperparameters [MA10], with better performances than the classical MH. The chain's parameter have little impact of the convergence properties, and can be randomly fixed. From these procedures, a set $\{\rho_i\}_{i=1}^{P}$ is obtained, giving the following EI formula:

$$\widetilde{\mathrm{EI}}(x) = \frac{1}{P}\sum_{i=1}^{P} \mathrm{EI}_{\rho_i}(x) \tag{3.2.16}$$

The optimization of the acquisition function is next performed with any procedure, stochastic or deterministic. Numerical differentiation is possible to estimate the Jacobian, taking advantage of the fast evaluation of the acquisition function. Since the EI (either MLE of Bayesian form) can have multiples local maxima, the search should be repeated.

3.2.6 *Proposed methodology for parameter estimation of RFS filters*

Using the PHD or δ-GLMB filters, maximization of the data likelihood $\mathcal{L}_\vartheta(Z_{1:N}) = \prod_{k=1}^{N} f_\vartheta(Z_k)$ is not straightforward. Unfortunately, a close form for calculation of this quantity is not accessible under assumptions made by these two filters. However, formula of calculation can be used to perform estimations of their values.

For the PHD filter, if the predicted multi-target density $f_{k+1|k}\left(X \mid Z_{1:K}\right)$ is approximately Poisson, using notation $Z_k = \{z_i\}_{i=1}^{m}$, then [Mah03; Mah13b]:

$$f(Z_k) = \exp\left[-\lambda - \int D_{k|k-1}(x)p_D(x)\,dx\right]\prod_{i=1}^{m}\left(\kappa(z_i) + \int D_{k|k-1}(x)p_D(x)f(z_i \mid x)\,dx\right) \tag{3.2.17}$$

Under the Poisson assumption, calculation of the likelihood of measurement set Z_k has still complexity $\mathcal{O}(nm)$ with n the number of targets.

However, for the δ-GLMB filter, evaluation formula of likelihood of the set Z_k gives [Reu+13a; Mah13b]:

$$f(Z_k) = f_c(Z_k)\sum_{(I,\xi)\in\mathcal{F}(\mathbb{L})\times\Xi}\sum_{\pi\in\Pi(I)}\omega^{(I,\xi)}\left[p^{(\xi,\pi)}(\cdot \mid Z_k)\right]^I \tag{3.2.18}$$

with the term related to clutter (with Poisson assumption) $f_c(Z_k)$ defined as:

$$f_c(Z_k) = e^{-\lambda}\prod_{i=1}^{m}\kappa(z_i) \tag{3.2.19}$$

The form of the likelihood for the δ-GLMB filter is the same as with resolution of the Bayesian recursion equations. The overall complexity of evaluation grows exponentially with targets, measurements and steps. In this specific case, use of only terms obtain after truncation will result in a biased estimate of the likelihood. The bias will highly depend on the number of truncations allowed.

Our proposed methodology to find the best parameter vector ϑ which maximizes the data likelihood is to use the Poisson approximation. Final objective is to use the δ-GLMB filter for track estimations, but the PHD filter is used as a simplification, whose only objective is to provide an efficient evaluation of the data likelihood. This strategy is sub-optimal, but is a practical solution to the tuning of the tracker.

However, for long measurements (high number of measurement sets, N), evaluation of the data likelihood remains a long task. We propose to use a second level of approximation to find the parameter vector $\vartheta^\star$ which maximizes the data likelihood estimated with the Poisson approximation. Summary of the full process in presented in figure 3.5.

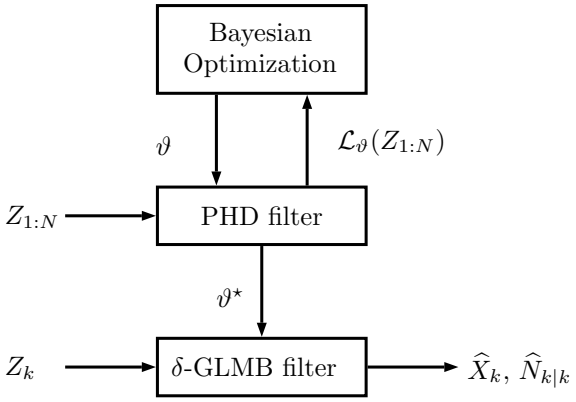

Figure 3.5: Proposed procedure for estimation of the optimal model and filter vector $\vartheta^\star$, using the Poisson approximation, and Bayesian Optimization methods.

Dataset for the BO recursion will be composed of proposed parameter vectors with their associated likelihood values, estimated from PHD filter runs on the total measurement sets $Z_{1:N}$. The final parameter vector $\vartheta^\star$ is the only vector applied to the δ-GLMB filter. Estimates of states, cardinalities and tracks are next performed, producing the algorithm output. Consequently, from only the measurement sets $Z_{1:N}$, this methodology estimates the model and filter parameter, and provides estimations of targeted tracks.

All the steps are presented in algorithm 1. The procedure needs preliminary evaluations of the likelihood for parameters vectors spread on the parameter space, to create an initial dataset $\mathcal{D}$. Using $n_0 \in \mathbb{N}^*$ points uniformly sampled in this space is possible. These points allow an efficient training of the GP in the first BO loop. An operator only have to select a number of initial parameter vectors, and the maximum number of BO recursions n_{iter}, depending on the time or resources available. Optimal parameter vector is next selected from the final dataset, and used by the δ-GLMB filter to provide tracks and cardinalities.

3.3 Bayesian inference in multi-target scenario

3.3.1 *Simulation scenario for performances assessment*

The aim of this section is to demonstrate the performances of the above methods on a well documented scenario.

Algorithm 1: Automatic procedure for track estimation with the δ-GLMB filter.

Input : Measurement sets $Z_{1:N}$, Initial dataset $\mathcal{D} = \left\{(\vartheta_{0,i}, \mathcal{L}_{\vartheta_{0,i}})\right\}_{i=1}^{n_0}$

Output: Optimal model an filter parameter vector $\vartheta^\star$, track estimates $X_{1:N}$, cardinality estimates $\left\{N_{k|k}\right\}_{k=1}^{N}$

Bayesian Optimization:

for $i = 1, .., n_{\text{iter}}$ **do**

- Fit a Gaussian Process f with dataset $\mathcal{D}$
- Use f to construct $EI_{\rho^{\text{MLE}}}$ (or $\widetilde{\text{EI}}$)
- Select $\vartheta_i = \arg\max_\vartheta EI_{\rho^{\text{MLE}}}(\vartheta)$ (or $\arg\max_\vartheta \widetilde{\text{EI}}(\vartheta)$)
- Run a PHD filter to estimate $\mathcal{L}_{\vartheta_i} = p(Z_{1:N} \mid \vartheta_i)$
- Update dataset $\mathcal{D} = \mathcal{D} \cup \{(\vartheta_i, \mathcal{L}_{\vartheta_i})\}$

Set $\vartheta_\star = \arg\max_{\vartheta \in \mathcal{D}} \mathcal{L}_\vartheta$

δ-GLMB filter recursion:

for $k = 1, .., N$ **do**

- Joint predict and update to estimate $\pi(X \mid Z_k, \vartheta_i)$
- Estimate $\widehat{X}_k$ and $N_{k|k}$ from $\pi(X \mid Z_k, \vartheta_i)$

This simulation has been intensely exploited in the RFS filtering literature to analyse the performances of new filters, and compare them with existing state-of-the-art results [Vo+13; VVP14; VVH17]. Here, known targets are assumed to move in a $2D$ environment, described by two axes (x-axis and y-axis). Each target is described by a four dimensional vector $x \in \mathbb{R}^4$, containing the positions (p_x and p_y) and velocities (v_x and v_y) according to the two axis $x = [p_x, v_x, p_y, v_y]^T$. Targets tracks are simulated during $N = 100$ steps. Time interval between each step is set to $\Delta t = 1\,s$. Transition function is assumed to be linear with Gaussian perturbation, with transition matrix F and covariance matrix Q:

$$F = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } Q = \sigma_Q^2 \begin{bmatrix} \Delta t^4/4 & \Delta x^3/2 & 0 & 0 \\ \Delta t^3/2 & \Delta t^2 & 0 & 0 \\ 0 & 0 & \Delta t^4/4 & \Delta x^3/2 \\ 0 & 0 & \Delta t^3/2 & \Delta t \end{bmatrix} \tag{3.3.1}$$

with $\sigma_Q^2 = 0.1\,m/s^{-2}$. Measurements available are the noisy position, with measurement matrix H and perturbation matrix R such that:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } R = \sigma_R^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.3.2}$$

with $\sigma_R = 10\,m/s$. In addition, clutter points are generated according to a Poisson RFS with clutter rate $\lambda = 10$ and uniform distribution over the volume $[-1000, 1000] \times [-1000, 1000]$. A total of 12 targets will appear, but at maximum 10 at the same time (2 will die).

All tracks are displayed in figures 3.6, where the real paths are shown

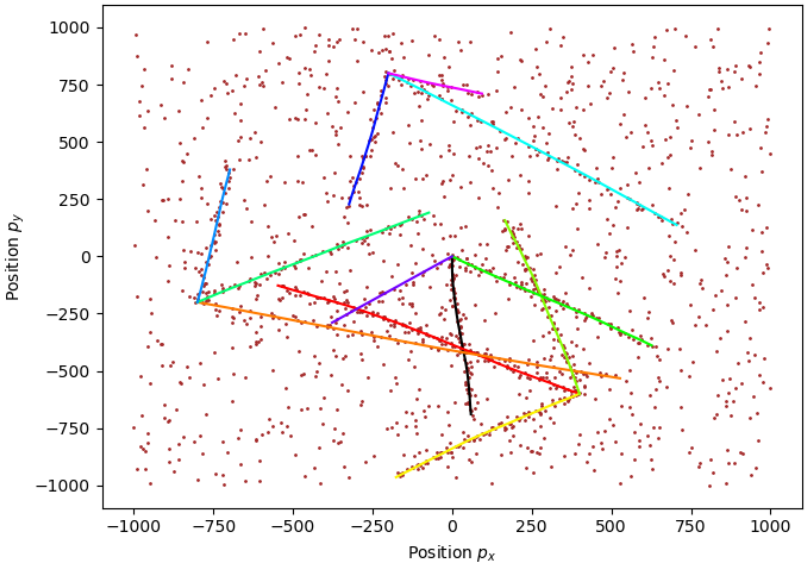

Figure 3.6: $2D$-representation of the tracks with clutter.

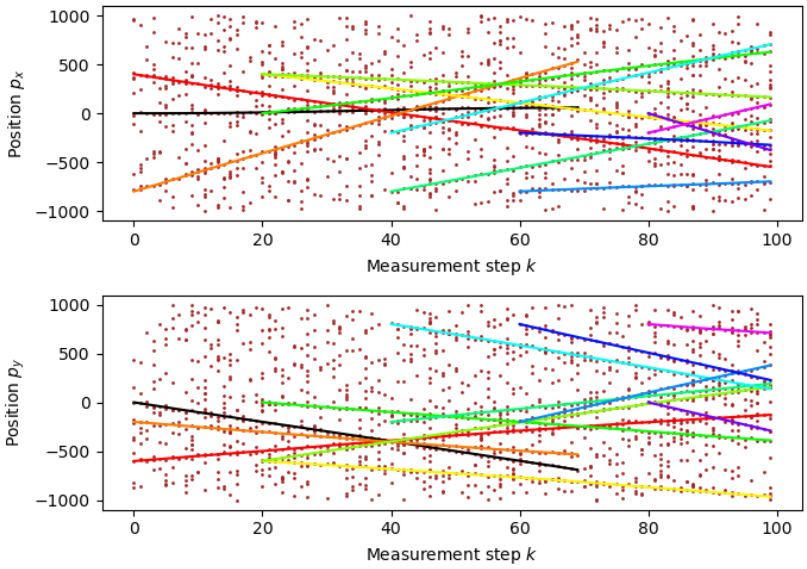

Figure 3.7: Step dependent representation of the tracks with clutter on x-axis and y-axis.

on the $2D$-plan in colours, with the cluttered measurements shown as the points. Figure 3.7 shows the evolution of displacement of tracks on each dimension with time (or steps) with the same rules of colours. For a human operator, distinction of individual tracks (without colours) is made difficult by both the presence of clutter and the noisy paths of tracks, justifying the use of a RFS filter.

3.3.2 *Kernel choice for Bayesian Optimization*

To face numerous cases in the parameter optimization procedure of algorithm 1, the GP covariance function has to be chosen. A common and flexible choice is the *Matérn* covariance function [RW06], which generalizes the square-exponential function. Adding a white noise term to this kernel allows also to solve instabilities and is useful when evaluations are not deterministic. If the parameter vector ϑ has dimension $d \in \mathbb{N}^*$, our *anisotropic* covariance function between vector $\vartheta_1 = [\vartheta_{1,1}, .., \vartheta_{1,d}]^T$ and vector $\vartheta_2 = [\vartheta_{2,1}, .., \vartheta_{2,d}]^T$ is defined as:

$$K(\vartheta_1, \vartheta_2) = C\frac{2^{1-\nu}}{\Gamma(\nu)}\left(\sqrt{2\nu}\sum_{i=1}^{d}\frac{(\vartheta_{1,i}-\vartheta_{2,i})^2}{l_i}\right)K_\nu\left(\sqrt{2\nu}\sum_{i=1}^{d}\frac{(\vartheta_{1,i}-\vartheta_{2,i})^2}{l_i}\right)+W \tag{3.3.3}$$

with C a positive constant, Γ the Gamma function, K_ν the modified Bessel function of second kind of order ν, ν a positive hyperparameter, l a stricly positive vector and W a positive scaler which models a white noise. A common choice for hyperparameter ν is 2.5 or 1.5 (selected for the following applications). Other values make the evaluations of the function hard. The vector l models a scale hyperparameter for every dimension of ϑ, causing the anisotropic property of the kernel. For our application the complete EI formula 3.2.16, with 1000 sampled hyperparameters vectors (and 100 kept).

Selecting a vector l with identical components gives the *isotropic* kernel, which lower the complexity of hyperparameter tuning, but can be damaging in high dimension.

3.3.3 *Posterior distribution of parameter*

The birth intensity is described as:

$$\gamma(x) = \sum_{i=1}^{4} w_b \mathcal{N}\left(x \mid \mu_b^{(i)}, P_b\right) \tag{3.3.4}$$

with w_b, $\mu_b^{(1)} = [0,0,0,0]^T$, $\mu_b^{(2)} = [-800,0,-200,0]^T$, $\mu_b^{(3)} = [-200,0,8000,0]^T$, $\mu_b^{(4)} = [400,0,-600,0]^T$ and $P_b = \mathrm{diag}([100,100,100,100])$.

Estimation of likelihood with PHD filter is based on the Poisson approximation of the multi-target densities. For the GLMB filter, the likelihood is approximated by its truncation, depending on the number of retained hypotheses. Parameter vector (unknown) gathers the clutter Poisson parameter λ and the measurement noise parameter σ_R^2. Goal is to provide estimation of the posterior density $p(\vartheta \mid D)$ with $\vartheta = \left[\lambda, \sigma_R^2\right]$. To reach this objective, a MCMC algorithm is run, with adaptive exploration strategy during the burn-in phase. The real vector has thus value $\vartheta^{\mathrm{True}} = [10, 100]$.

To run the MH chain, Uniform prior and half-Gaussian wide priors are used respectively for λ and σ_R^2:

$$p(\lambda) = \mathcal{U}([0, 25]) \tag{3.3.5}$$
$$p(\sigma_R^2) = \text{Half-}\mathcal{N}(100) \tag{3.3.6}$$

The density Half-Gaussian distribution with variance σ^2 is defined on the positive real axis $\mathbb{R}^+$ as twice the density of the Gaussian distribution with mean 0 and variance σ^2.

The random-walk proposal is applied on the log-transform of the variable σ_R^2, to prevent negative values. Consequently, the acceptance ratio takes the following form:

$$\alpha(\vartheta^*, \vartheta_k) = \frac{\mathcal{L}_{\vartheta^*}}{\mathcal{L}_{\vartheta_k}} \frac{p(\vartheta^*)}{p(\vartheta^k)} \frac{\sigma_R^{2,*}}{\sigma_{R,k}^2} \tag{3.3.7}$$

The δ-GLMB filter uses truncations of 25 and 100 components. Posterior distributions resulting from the PHD MCMC chain has supports centred on the true values of parameters. The GLMB filter, however, shows posterior distributions with high biases, even in this low dimension case ($d = 2$), as shown in comparison between figures 3.8 and 3.9. In this situation, it seems to be more interesting to use the PHD filter for the simulation of numerous filtering processes. In addition, contrary to the δ-GLMB filter, it has only linear complexity in the measurement set cardinality. Reducing the bias on the likelihood could require the computation of a higher number of hypotheses, increasing the computational cost. The Poisson approximation is therefore enough for this specific application.

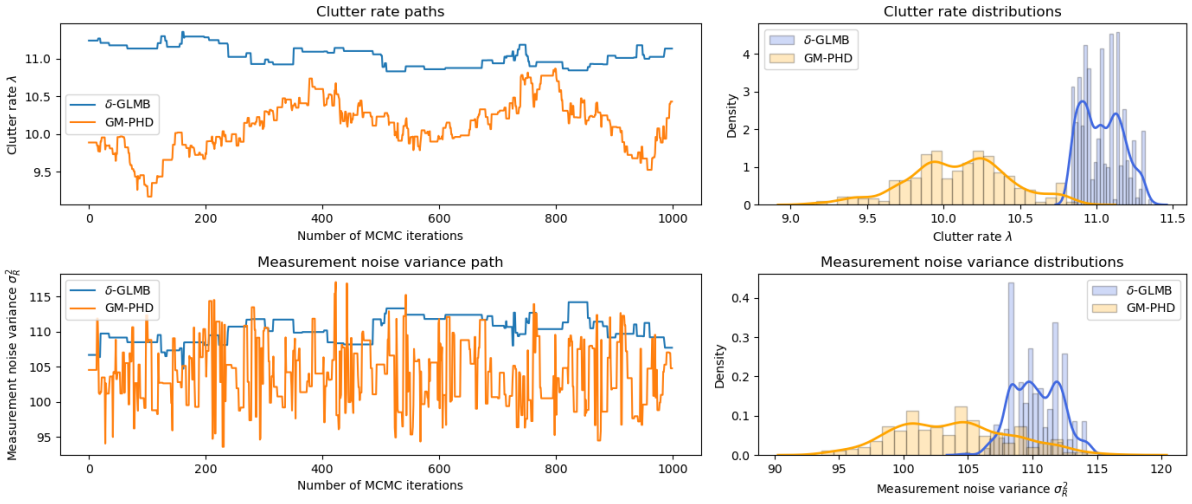

Figure 3.8: MCMC chain with 25 terms truncation for the GLMB filter.

Observation of cardinalities highlights the lost induced by a low number of truncation terms with the GLMB filter. Even if the posterior cardinalities have less variance than the ones obtained from the PHD filter, the GLMB filter is not able to track all targets with a low number of retained hypotheses, see figures 3.10 and 3.11. The PHD filter posterior cardinalities have higher variances but approach the real cardinalities with less resources.

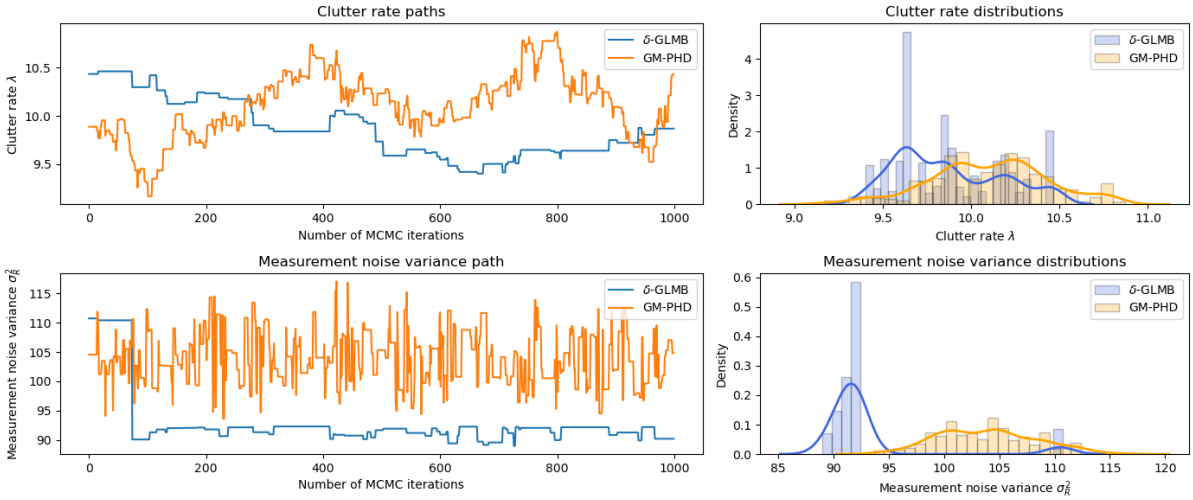

Figure 3.9: MCMC chain with 100 terms truncation for the GLMB filter.

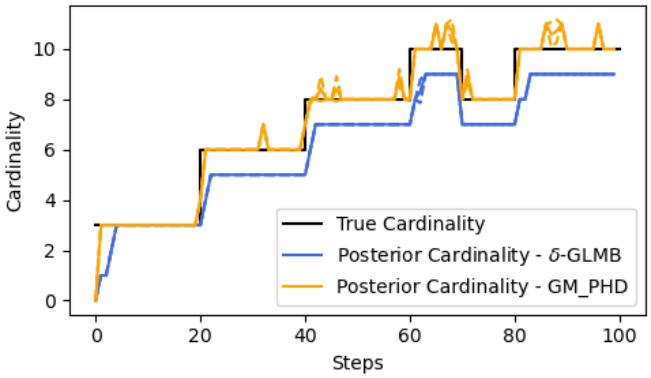

Figure 3.10: Cardinalities with 25 terms truncation for the GLMB filter.

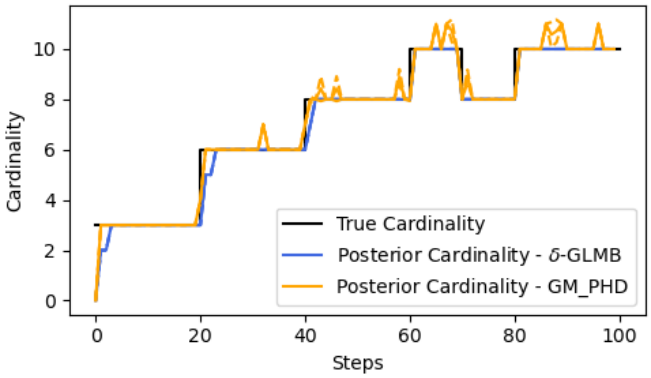

Figure 3.11: Cardinalities with 100 terms truncation for the GLMB filter.

3.3.4 *Fast optimal parameter search on simulation*

Contrary to the previous case, the measurement noise parameter is split into 2 different parameters: one for x-axis (σ_x^2) and one for y-axis (σ_y^2). Search is performed in a space of dimension 3 $\vartheta = \left[\lambda, \sigma_x^2, \sigma_y^2\right]^T$. The covariance function for the GP is the Anisotropic Exponential Square kernel (one scale variable per dimension, therefore 3 in our case). Hyperparameters are tuned using the MLE method. A maximum of 50 iterations, during the BO process, is performed. Optimization of the acquisition function is made with random search combined with use of the BFGS algorithm.

Here, it is sufficient to use the PHD filter to evaluate the log-likelihood. After twenty evaluations, a value is obtained $\lambda = 9.8$, $\sigma_x^2 = 10^{1.05}$ and $\sigma_y^2 = 10^{1.04}$. Use of δ-GLMB filter is therefore not necessary, if one is looking only for a rough estimation of the vector parameter.

3.4 Ultrasonic pulses tracking

3.4.1 *General model for pulses propagation*

In this section, a state space model is proposed to represent ultrasonic pulses across acquisitions. A major point is the ability of surface waves (and particularly Rayleigh waves) to interact with cracks. A phenomenon of diffraction of the Rayleigh wave has already been studied [Jia+07], and dependences between reflected pulse amplitudes and the crack depth highlighted. Relationships between angle of an inclined crack and amplitudes of refracted and reflected pulses has been analysed in the context of laser generated Rayleigh waves [Ni+10]. However, methodologies to analyse these particular phenomena are not suited to mobile inspection. Nevertheless, possibility to extract information from reflections allows the construction of a state space which includes the possibility of identifying these events from regular pulses.

In ultrasonic signals, events take the form of pulses distributed across a time interval (duration of measurement). Decomposition with Gabor wavelets allowed the transformation of raw measurement, or acquisition, into a set of Gabor functions parameters. Each event, or pulse, has a specific characterization inside one acquisition. However, ultrasonic sensors do not guarantee perfect measurements. Depending on the sensor quality, measurement system or properties and shape of the inspected medium, the raw signal can carry disturbances, deteriorating the decomposition. A common solution to bypass short disturbances, such as measurement noise, is to average numerous acquired signals. If noise is Gaussian and centred, independent of the measurement sampling, the mean signal is the best estimator, which precision increases with the number of signals included.

In ultrasonic testing with wedges, a little variation of the sensor's orientation can induce a miss detection of several pulses, depending on the directivity of the propagating waves.

However, a major drawback of dynamic inspection is the impossibility to perform multiple acquisitions on a fixed position of the material.

On linear shape inspected mediums, such as rails, each acquisition will deliver unique information about a local area. Averaging is no longer a valid solution, since pulses could have changing characteristics across acquisitions. Hence, it is necessary to model these specific behaviours, to recognize the pulses and follow them as the sensors move.

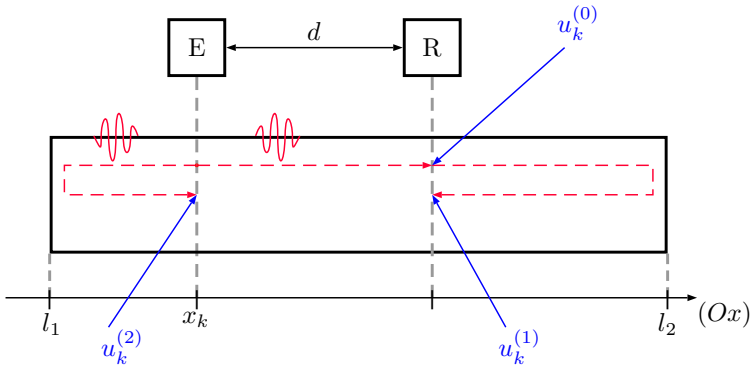

Figure 3.12: Pulses propagation on a bounded medium (limited on the x-axis between l_1 and l_2); E: emitter; R: receiver; d: distance between E and R; $u_k^{(i)}$: ToF for pulses for modes $i = 0, 1, 2$.

In a bounded medium, ultrasonic pulses propagating are reflecting by the borders. Figure 3.12 illustrates this phenomenon with a solid delimited by two planes, put on the x-axis, at positions l_1 and l_2, such that $l_1 < l_2$. After generation by the emitter (E) at position x_k, the pulse travels a path of length d to the receiver (R). It next continues its travel across the medium, and undergoes a reflection because of a plane orthogonal to its propagation direction. For an acquisition indexed by the integer $k \in \mathbb{N}$, the ToF of the pulse coming from the emitter and directly passing under the receiver is noted $u_k^{(0)}$. This pulse is called the *direct* pulse. After reflection on border l_2, the pulse produces a new excitation of the receiver, called the *front echo*, with ToF noted $u_k^{(1)}$. In addition, the emitter could not be perfectly uni-directional. During the generation of the pulse at x_k, a small amount of energy could generate an ultrasonic pulse which propagates in the opposite direction of the direct pulse. A possible reflection on border l_1 could result in one other pulse, called the *back echo*, if the pulse reaches the receiver.

for non-orthogonal planes, angle of reflection follows the Snell-Descartes law

Some relationships between the different ToF are easily obtained, like $u_k^{(0)} \leq u_k^{(1)}$ and $u_k^{(0)} \leq u_k^{(2)}$. Especially, for both positions of emitter and receiver between the two boundaries ($l_1 \leq x_k \leq l_2 - d$), for a constant pulse velocity v, and an emission at time t_0, the following relationships are raised:

$$d = v(u_k^{(0)} - t_0) \tag{3.4.1}$$

$$d + 2(l_2 - x_k) = v(u_k^{(1)} - t_0) \tag{3.4.2}$$

$$d + 2(x_k - l_1) = v(u_k^{(2)} - t_0) \tag{3.4.3}$$

Between two successive acquisitions, one at step $k-1$ and a next at step k, ToF for front and back echoes are modified. With notation $\Delta x = x_k - x_{k-1}$, if $\Delta x > 0$ (displacement of the emitter-receiver device in the direct direction of the x-axis, the front echo will be detected

earlier $u_k^{(1)} < u_{k-1}^{(1)}$, and the back echo later $u_{k-1}^{(2)} < u_k^{(2)}$. The direct one should not have moved $u_k^{(0)} \approx u_{k-1}^{(0)}$.

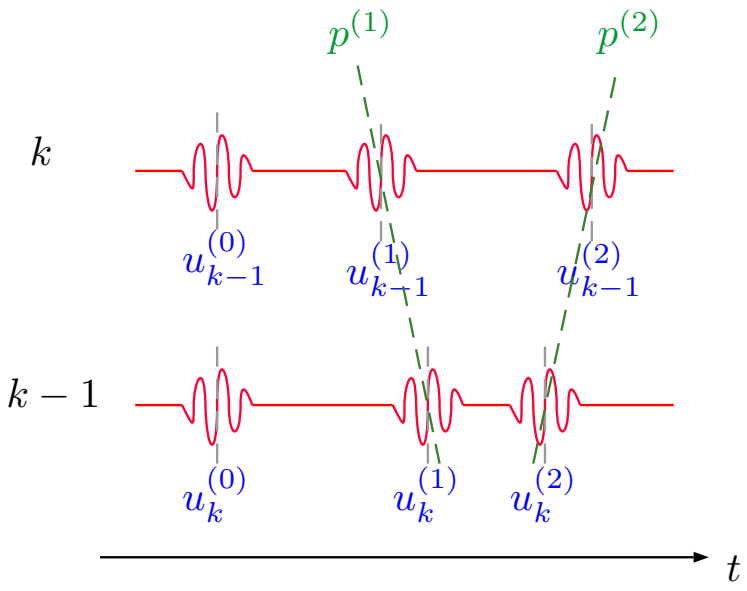

Figure 3.13: Change of arrival times for direct pulse (index 0), front (index 1) and back (index 2) echoes with between two consecutive ultrasonic acquisitions.

To analyse the change of ToF, we define the *pace* p_u as:

$$p_u = \frac{\Delta u}{\Delta x} \tag{3.4.4}$$

For the front echo, it is written:

$$\begin{aligned} p_u^{(1)} &= \frac{u_k^{(0)} - u_{k-1}^{(0)}}{\Delta x} \\ &= \frac{d + 2(l_2 - x_k) - d - 2(l_2 - x_{k-1})}{v} \frac{1}{\Delta x} \\ &= -\frac{2}{v} \end{aligned} \tag{3.4.5}$$

Similarly, for the back echo:

$$p_u^{(1)} = +\frac{2}{v} \tag{3.4.6}$$

This indicator, which is indeed a derivative of a time quantity according to a variation of a space position, is directly proportional to the inverse of the pulse velocity v. A knowledge of Δx allows to retrieve this important information.

Of course, for complex geometries, multiples reflections could arise. If different kinds of waves are generated, recorded signal would be a mixture of direct pulses with their reflections. In the following, we make the distinction between these identified types of pulses:

- Direct pulses, identified as the *mode* 0, which have constant ToF across acquisitions, and therefore null paces;
- Front echoes, resulting from reflections, with the latest on the front border (relatively to the sensors), identified as the *mode* 1. Their paces are negative;

- Back echoes, resulting from reflections, with the latest on the back border, identified as the *mode* 2. Their paces are positive.

Unfortunately, the decomposition of ultrasonic signals does not include instantaneous information about the pulses natures (direct or echo), and their paces (and consequently their propagation velocities). This motivates the use of the filtering and tracking methods, to recover these hidden variables and to decompose the joint set of acquisitions into specific groups of pulses (thus by modes).

A noticeable result can easily be obtained with consideration of echo paths. If the distance between the emitter and the receiver, noted $\overline{ER}$ and the current position of the emitter are known, it is possible to estimate the position of the defect.

With front echoes, replacing l_2 by the position of the defect x_D gives:

$$x_D = x_k + \frac{v\,(u_k^{(1)} - t_0) - \overline{ER}}{2} \tag{3.4.7}$$

Similarly, for back echoes, replacing l_1 by x_D gives:

$$x_D = x_k - \frac{v\,(u_k^{(2)} - t_0) - \overline{ER}}{2} \tag{3.4.8}$$

These relationships can therefore be used when a track for a back or front echo has been identified in the process, giving a precision on the defect location.

3.4.2 *Ultrasonic pulse model*

A further specification of the model used in this thesis is now presented. A measurement set Z produced by a decomposition algorithm is assumed to contain the noisy versions of the real states. A measurement vector $z \in Z$ can be written:

$$z = \left[\widetilde{u}, \widetilde{A}, \widetilde{\nu}, \widetilde{\varsigma}\right] \tag{3.4.9}$$

where $\widetilde{y}$ refers to the noisy measurement of a state y.

Inclusion of dynamic information between steps is performed by the first order approximation 3.4.4. Using the pace p_u as an information available in the state space, the following approximation results from previous equation:

$$u_k = u_{k-1} + \Delta x\, p_{u,k} \tag{3.4.10}$$

The take into account evolution of amplitude with time, a pace p_A related to state value A is introduced, which follows the same hypotheses:

$$A_k = A_{k-1} + \Delta x\, p_{A,k} \tag{3.4.11}$$

Possibility of variations of the amplitude allows to track the decay induced by a pulse reflected by a distant defect. Moreover, slow

variations of the lift-off could also induce a variation of the signal amplitude. Allowing a small change with time for all pulses (direct pulses included) gives more flexibility to the tracking scheme.

To prevent overflow errors, the variable $\vartheta = \sqrt{s}$ is preferred over s. A space vector x belonging to the state set X has therefore the form:

$$x = [u, p_u, A, p_A, \nu, \varsigma,]^T \tag{3.4.12}$$

The chosen model is the linear one with Gaussian perturbations (from example 2.3.1).

The transition matrix between states, for equation 2.3.3 is written:

$$F = \begin{bmatrix} 1 & \Delta x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta x & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.4.13}$$

Measurement matrix for equation 2.3.4 is defined as:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.4.14}$$

The state perturbation covariance matrix selected (2.3.5) is the discrete piecewise model [Sär13]:

$$Q = \begin{bmatrix} q_u\Delta x^3/3 & q_u\Delta x^2/2 & 0 & 0 & 0 & 0 \\ q_u\Delta x^2/2 & q_u\Delta x^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & q_A\Delta x^3/3 & q_A\Delta x^2/2 & 0 & 0 \\ 0 & 0 & q_A\Delta x^2/2 & q_A\Delta x & 0 & 0 \\ 0 & 0 & 0 & 0 & q_\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & q_\varsigma \end{bmatrix} \tag{3.4.15}$$

with four parameters q_u, q_A, q_ν, q_ς, respectively related to variables u, A, ν and ς.

Finally, noise measurement matrix (equation 2.3.6) is defined, with four more positive parameters r_u, r_A, r_ν, r_ς, as:

$$R = \begin{bmatrix} r_u & 0 & 0 & 0 \\ 0 & r_A & 0 & 0 \\ 0 & 0 & r_\nu & 0 \\ 0 & 0 & 0 & r_\varsigma \end{bmatrix} \tag{3.4.16}$$

All state values are assumed to belong to $\mathbb{R}$. Although this contradicts the physical need to obtain positive estimations of amplitude A, frequency ν and spread ς, since measurements are ensured to follow this condition with the MP algorithm, this issue is mitigated. The distinction between modes (0, 1 or 2) is entirely dependent on the value of the pace p_u. To prevent switching of modes between steps, little uncertainty should be related to estimation of this variable. This is traduced in practice by use of low amplitude initial covariance coefficients for the related variable.

Variables dedicated to the model to be estimated are therefore: q_u, q_A, q_ν, q_ς, r_u, r_A, r_ν and r_ς. To save some resources, and use available information from visual inspections of B-scans, survival probability is fixed to $p_S = 0.90$, and detection probability $p_D = 0.95$, therefore constant for our experiments.

For our experimentations, u and ς are expressed in μs, and ν in MHz. With issue related to gain, A is not expected to exceed $100\,u.a.$. Moreover, Δx is expressed in m.

3.4.3 *Birth design*

We benefit from the developments of adaptive birth strategies for the GM-PHD [Bea+12] and δ-GLMB [LVN16] filters. They allow the birth of new tracks around current existing measurements. Here, the distributions of new tracks are initiated by a single Gaussian component. For the GM-PHD filter, newborn targets will evolve into persistent targets at the next step, but are already included at current step. For the δ-GLMB filter, new targets will only be initiated during the next step, depending on the likelihood of each measurements (measurements with low likelihood have more chance to initiate new tracks than measurements already explained by existing tracks).

Adaptive births strategies create new parameters specific to each filter, to be tuned. However, since there are only few parameters changing, we suggest repeating for some iterations our optimization procedure, to find the best birth parameter for the δ-GLMB filter, but keeping the common parameters shared with the GM-PHD filter.

In addition, since no information about the mode of the new track is available, each measurement initiates three new targets: one for each pulse mode. At step k, for each vector $z_k^{(i)} \in Z_k$, such that $z_k = \left[\widetilde{u}_k, \widetilde{A}_k, \widetilde{\nu}_k, \widetilde{\varsigma}_k\right]^T$, three Gaussian distributions are initiated, defined by their mean vector:

$$\mu_b^{(i)} = \left[\widetilde{u}_k, p_u^{(i)}, \widetilde{A}_k, p_A^{(i)}, \widetilde{\nu}_k, \widetilde{\varsigma}_k\right]^T, i \in \{0, 1, 2\} \tag{3.4.17}$$

and the common covariance matrix P_b, which is diagonal, with diagonal equals to:

$$\mathrm{diag}(P_b) = [r_u, 1, r_A, 1000, r_\nu, r_\varsigma]^T \tag{3.4.18}$$

Depending on the pulse mode i, values of p_u and p_A are adapted from the knowledge we have about attenuation with distances and

Mode i	Arrival time pace $p_u^{(i)}$	Amplitude pace $p_A^{(i)}$
0 (Direct)	0	0
1 (Front)	-678	100
2 (Back)	678	-100

Table 3.1: Variables of the adaptive birth mean which are specific to each pulse mode.

velocity of the Rayleigh wave: front echoes (mode 1) arrive earlier with step, and their amplitudes increase (attenuation is reduced as paths are shortened), and opposite results are expected for back echoes (mode 2). Values are gathered in table 3.1.

3.4.4 *Mode identification*

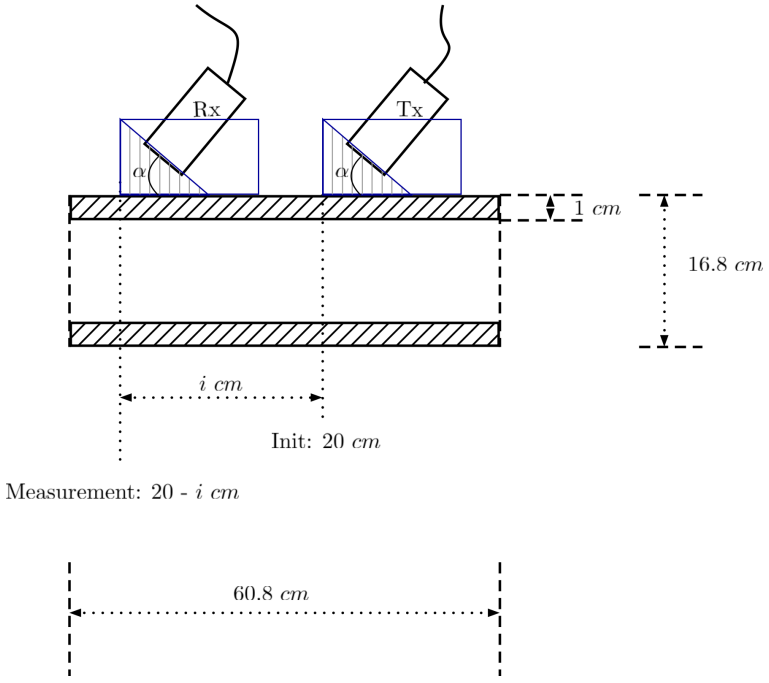

Figure 3.14: Configuration for test on a steel pipe with wedge ultrasonic transducers in the same direction.

This configuration favour the echoes

A steel pipe is inspected, with two wedges transducers (with 35° angles) in the same direction, see figure 3.14. Since the inspection is performed near the borders of the pipe, echoes are present is the measured signals, as shown in figure 3.15. To estimate the pulses velocities (direct pulses and echoes), it is necessary to include the mode of the pulse inside the estimation process. Since no information is a priori available about the pulse mode, the filter will have to decide the mode by applying recursive decisions. Detector is the Matching Pursuit algorithm, which produced the sets of Gabor parameter vectors. Information selecting for tracking are the amplitude of the pulses and their ToF. A RFS filter with multiple models is constructed, with a large Gaussian prior put for velocity direct pulses, and a different for the echo pulses. Model parameters are selected using the Bayesian Optimization procedure.

A total of 8 acquisitions are obtained. All the tracks are shown

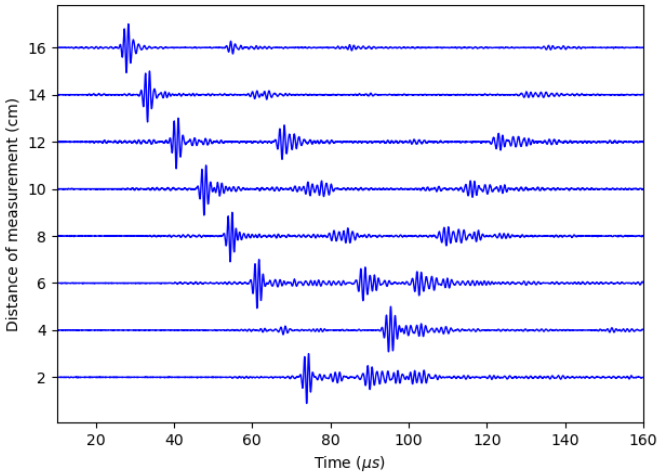

Figure 3.15: Signals from the steel pipe experiment.

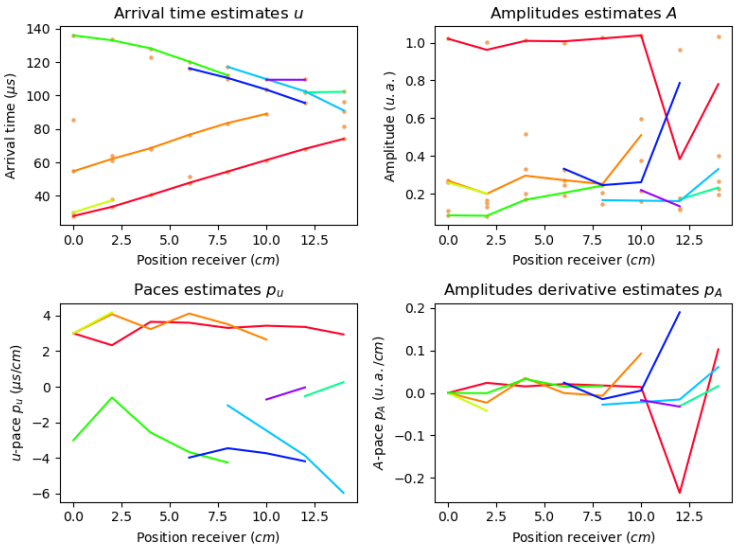

Figure 3.16: Tracks obtained from arrival times (in μs) and amplitudes of pulses (in arbitrary units $u.a.$) of the experiment on steel pipe.

in figure 3.16. From the tracks, the mean velocities are extracted. Results are summarized in the following table, with ranking based on the length of tracks (or lifespan), and showing their estimation. For short tracks, estimations are not reliable. Longer tracks show propagation velocities in the range of $3000 m/s$, consistent with the velocity of ultrasonic waves in metal.

Lifespan (steps)	Type	Velocity (m/s)
8	Direct	3125
6	Direct	2915
5	Echo	3547
4	Echo	2607
3	Echo	2760
2	Direct	2785

Table 3.2: Characteristics of pulses on steel pipe experiment.

3.4.5 *Velocity estimation on rail*

The previous methodologies are now demonstrated on a laboratory experiment conducted on the $1\,m$ rail. Purpose is to identify the echoes coming from reflections on defects. Previous methodology is applied to estimate optimal parameter vector. The δ-GLMB filter provides the final estimations of states and cardinalities.

For assessment of performances, the indicators used are: the cardinalities related to each mode, and OSPA (with parameters $p = 1$ and $c = 100$).

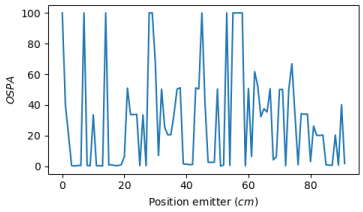

Figure 3.17: OSPA metric between set states from manual labelling and estimation from the δ-GLMB filter.

First observation is related to the OSPA: the metric exhibits some peaks at starts and ends of tracks. For some short tracks, relative to the signals between central defects of the rail, the tracker does not succeed in initiating tracks. This is certainly due to high variations of amplitudes in the areas.

Cardinalities obtained from the tracker and manual labelling exhibit the same trends (figure 3.20). However, cardinalities resulting from the tracker has slightly higher cardinalities.

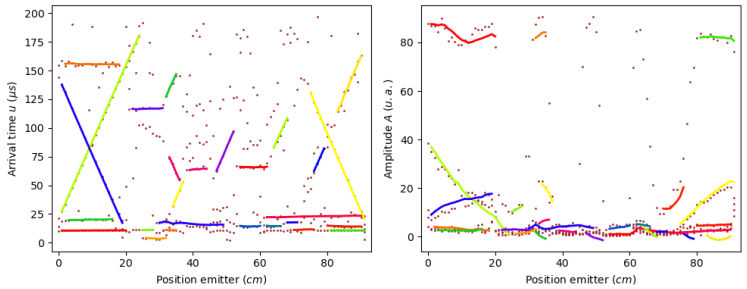

Figure 3.18: Arrival times and amplitudes estimated from the δ-GLMB filter.

Qualitatively, the tracker succeeded in reproducing the association performed by the manual labelling of points. Indicators highlighted the difficulties of the tracker to manage the short paths between defects. However, estimations are efficient on stable areas, and reflection

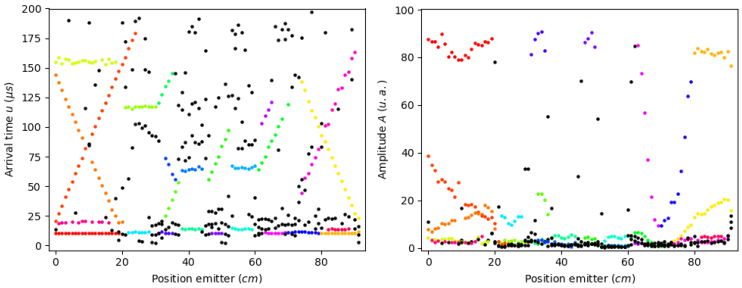

Figure 3.19: Arrival times and amplitudes resulting from manual labelling of points.

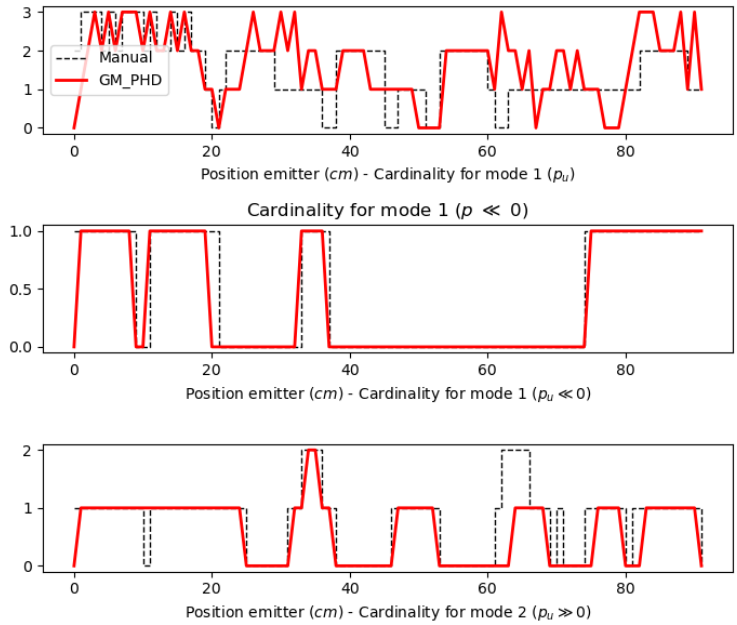

Figure 3.20: Estimated cardinalities for each mode - from manual labelling and tracker estimations.

echoes has also been clearly identified. For the next applications of rail inspection, the same model parameters will be used. Only made adaptations will be related to the clutter rate λ, which will depend on the experimental environment.

3.5 Sensitivity of sensors toward movement

3.5.1 *Laboratory experiments of mobile inspections*

RAILENIUM test bench allows automatic displacement of the measurement device over a limited distance, up to $2\,m$, at speed limited to $5\,m/s$. Previous experiments made with this equipment only performed qualitative analyses of inspections [Nda16]. The main benefit of measuring EMAT signals on laboratory is the control of the experiment parameters. The configuration of sensors and tested rails is known, allowing a comparison of results from immobile experiments with dynamic ones. Objective is to obtain with mobile measurements detection and characterization performances close to the previous obtained from clean signals in the ideal situation of immobile sensors with non-zero lift-off. Because of the non-negligible lift-off, ultrasonic pulses are expected to have lower amplitudes. Possible solution could be either to strengthen the power signal used by the generator for pulse emission, or to increase the gain. However, the later will also amplify external perturbations, which have the effect of generating noisy signals. For white Gaussian noise, frequency filtering allows a fast denoising of the signal. However, transformations of the received signal by hardware filters could create some artefacts in the signal with frequencies near the central frequency of EMAT pulses (around $500\,kHz$). Sparse decomposition with MP algorithm is expected to extract the main pulses of interest, and possibly minor events not related to pulse propagations.

To prevent electrical issues, a plastic or paper protection is always inserted between the rail and sensors, therefore the lift-off does not strictly equal $0\,mm$

This section will compare some features extracted from both immobile and mobile inspection of the two laboratory rails. A comparison of estimated ToF will be first perform, which allows a preliminary comparison of signals. Next, filtering methods with Gabor decompositions of signals will be applied to measurements. Since the mobile inspection does not allow a comparison of acquisitions at the same rail locations, use of OSPA distance is not performed. However, estimations of cardinalities for each pulse mode is still possible, with little adaptations of results.

3.5.2 *Time of flight degradation with speed*

Three examples of measurements performed on the $70\,cm$ rail are performed: one with immobile inspection, one with mobile inspection devices at low speed ($0.5\,m/s$) and medium precision (PRF = 25) and a last with same speed inspection but higher precision (PFR = 100). Distance between emitter and receiver is set to $10\,cm$, to measure a nominal ToF of $39.4\,\mu s$.

The CWT serves as a basis for identification of the ToF, with analysis made on the frequency bandwidth $[400\,kHz, 700\,kHz]$. Expected theoretical changes of ToF for $6\,mm$ depths defects are around $4.2\,\mu s$, and for $3\,mm$ depths around $2.2\,\mu s$.

Total path to add for the Rayleigh wave $\approx 2 * \text{depth} + \text{tickness}$

A display of the three related B-scans (figure 3.21) highlights the visual similarities between experiences. Even with moving sensors, echoes and reflections from defects are observed in signals.

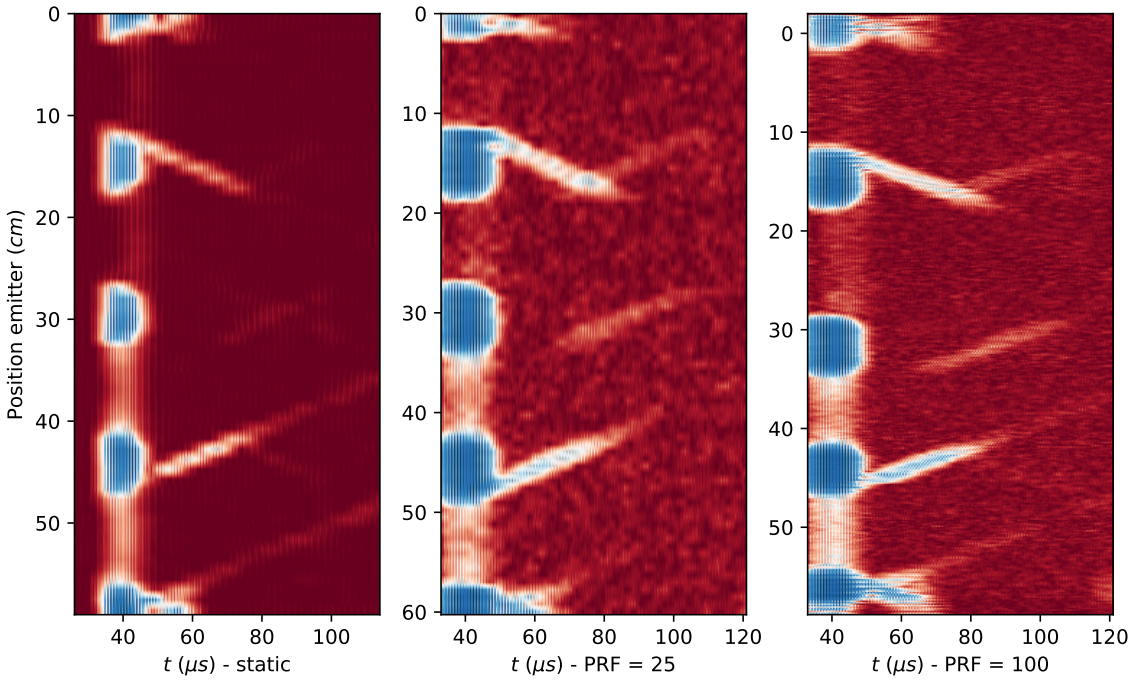

Figure 3.21: B-scan of the $70\,cm$ rail for static acquisitions, dynamic acquisitions with PRF = 25 and dynamic acquisitions with PRF = 100.

To study the estimations made of ToF, damaged areas are identified manually: when a defect is located between emitter and receiver, the estimations of ToF are gathered and referred to this specific defect. To prevent side effects, defects have to be located between magnets (not below them).

	Static		PRF = 25		PRF = 100	
	CWT	MP	CWT	MP	CWT	MP
Defect $6\,mm$, $\alpha = 0°$	5.0	5.2	10	38	7.0	25
Defect $6\,mm$, $\alpha = 45°$	4.2	4.3	3.2	3.2	4.1	8.1
Defect $3\,mm$, $\alpha = 45°$	0.53	0.60	0.87	0.76	0.83	0.95
Defect $3\,mm$, $\alpha = 0°$	0.44	0.51	1.0	0.59	1.2	0.92

Table 3.3: Mean difference of time of flights with reference value, in μs, for the $70\,cm$ rail for static acquisitions, dynamic acquisitions with PRF = 25 and dynamic acquisitions with PRF = 100, with wavelet and matching pursuit base estimators, for different defect depths and orientations.

Mean values estimated with the two methods for each areas are listed in table 3.3. A first observation is the discrepancy of estimations for $3\,mm$ depth defects with the theoretical values, even in the low noise situation (immobile measurements). This could be a consequence of the low depth of cracks: the wave does not necessarily have to increase its total path, since propagation at high depth in the ma-

terial is still possible. Next, for mobile acquisitions, both estimations are very poor for larger depth. Attenuation of pulses is nearly total, and little information is available in measurements. The MP based algorithm does not find the best atom which minimizes the error with theoretical ToF (2∗depth+tickness). This issue is however reduced for inclined cracks. Attenuation could be less important in this situation, allowing a higher portion of the pulse energy to pass the crack.

These experiment has shown the inconsistency of ToF estimators with noisy measurements and low amplitudes signals.

3.5.3 *Estimation of time of flight with data association*

Since estimations of ToF performed poorly with noised signals, challenge is to reconstruct a stable estimation using several acquisitions. Filters allow identification of pulses tracks and separation between modes. If at step k the estimates RFS obtained with cardinality N_k is noted $\widehat{X}_k = \left\{x_k^{(i)}\right\}_{i=1}^{N_k}$, with $x_k^{(i)} = \left[u_k^{(i)}, p_{u,k}^{(i)}, A_k^{(i)}, p_{A,k}^{(i)}, \nu_k^{(i)}, \varsigma_k^{(i)}\right]^T$, then the best vector x_k^* for ToF estimation is selected as the one with higher amplitude:

$$x_k^* = \arg \max_{x_k \in \widehat{X}_k} A_k \tag{3.5.1}$$

The ToF estimate u_k^* for the measurement is directly provided by the related ToF of x_k^*.

Compared to previous estimations of ToF, the later takes advantage of the RFS filter. If a empty set $\varnothing$ is provided as estimation at step k, no information is given by the process. If an existing track for direct pulses has already been initiated, and still exists at current step, a possible estimate can be obtained. On the contrary, if no tracks for direct pulses are coming from the filter, this could traduce the lack of information to provide association between vectors of the sparse decomposition. Instead of offering an unrepresentative value, absence of estimate prevents wrong interpretation of results.

Figure 3.22 demonstrates the interest of the filtering techniques combined with the state space representation. In the left B-scan, a direct picture of measurement, without any processing part, gathers the searched information. Direct transmissions with related attenuations due to defects are mixed with reflections. After sparse decomposition, and estimation of the pulses characteristics, identification of modes is performed using the pace of each estimated vector p_u. Reconstruction of envelopes for all pulses is possible, since the state space representation includes the shape of the Gaussian function related to Gabor functions. Because the phase is not included in the state space, recovering the complete signal with its oscillations is therefore not possible. Middle and right B-scans of figure 3.22 are the reconstruction of each individual modes. Paths of pulses are clearly identified in this static measurement configuration.

Figure 3.23 exhibits similar decomposition performances in comparison with the static configuration. Number of acquisitions is fixed to

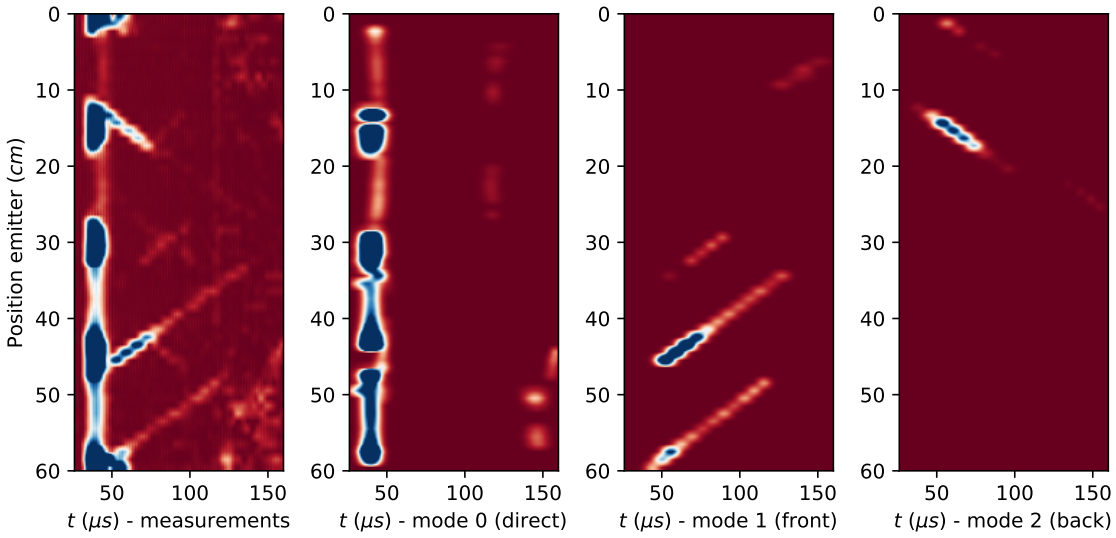

Figure 3.22: B-scan of the 70 cm rail for static acquisitions and decompositions into modes: direct, front echoes and back echoes.

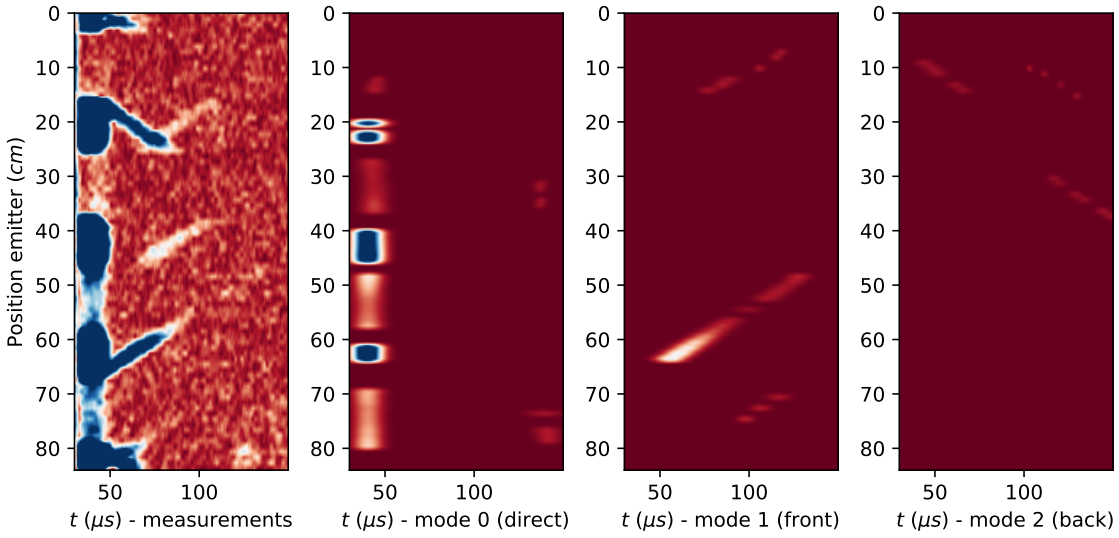

Figure 3.23: B-scan of the 70 cm rail for mobile acquisitions and decompositions into modes: direct, front echoes and back echoes with PRF = 25.

25 to obtain a repartition of measurement across the rail similar to the static case. Here, identification of modes and separation is still possible, despite the degradation of measurements with mobile sensors. In particular, reflection pulses are separated from direct transmissions. Use of their characteristics for improvement of defect location estimation is therefore a valid technique for mobile inspection of rails.

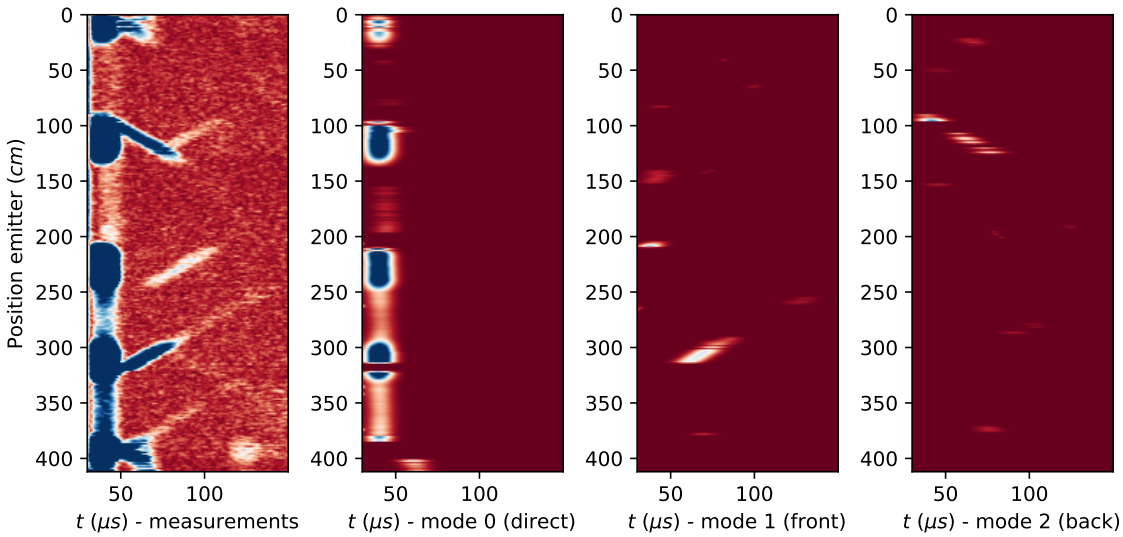

Figure 3.24: B-scan of the 70 cm rail for static acquisitions and decompositions into modes: direct, front echoes and back echoes with PRF = 100.

The last case has paradoxically the worst performances. The number of acquisitions per second has been increased, from 25 to 100, and improvements of the decomposition was expected. Even if direct transmission is still identified, reflections are hardly extracted. Their tracks are short, and poorly distributed across the B-scan. A possible explanation is the simplicity of adaptation of the optimal model and filter parameters to this situation. A modification of transition matrices could induce deeper changes of parameters. Fitting new data with a static measurement configuration before a new training phase should theoretically fix this issue, but the practical feasibility of such measurements is discutable. A different explanation could be the vibrations of the sensors during the displacement. This variability is integrated over time for large intervals between each acquisition, but could be observable for smaller ones.

	Static	PRF = 25	PRF = 100
Defect 6 mm, $\alpha = 0°$	4.4	5.7	5.6
Defect 6 mm, $\alpha = 45°$	5.7	4.2	4.7
Defect 3 mm, $\alpha = 45°$	0.94	1.3	1.4
Defect 3 mm, $\alpha = 0°$	-0.049	1.1	1.5

Table 3.4: Mean difference of time of flights with reference value, in μs, for the 70 cm rail for static acquisitions, dynamic acquisitions with PRF = 25 and dynamic acquisitions with PRF = 100, with wavelet and matching pursuit base estimators, for different defect depths and orientations.

Comparison with previous results (table 3.3) is straightforward, using the same methodology. For each set of vectors identified at step k, the subset of direct pulses, noted $\widehat{X}_k^0$, the pace p_u is identified, which

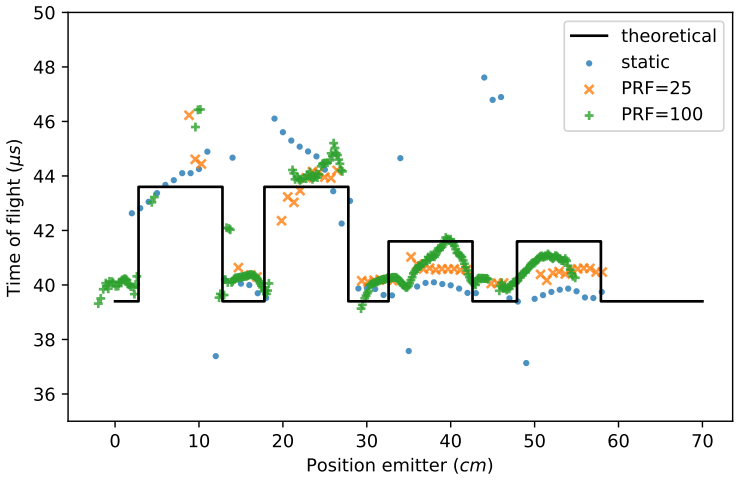

Figure 3.25: Comparison of time of flights estimates after filtering process for the 70 cm rail for static acquisitions, dynamic acquisitions with PRF = 25 and dynamic acquisitions with PRF = 100.

should be close to 0:

$$\widehat{X}_k^0 = \left\{ [u, p_u, A, p_A, \nu, \varsigma]^T \in \widehat{X}_k \mid p_u \approx 0 \right\} \tag{3.5.2}$$

The best vector x_k^* for ToF estimation is selected as the one belonging to this set with higher amplitude:

$$x_k^* = \arg\max_{x \in \widehat{X}_k^0} A \tag{3.5.3}$$

Here, the ToF estimates are more consistent. Figure 3.25 gives all points obtained from the δ-GLMB filter. Again, since the estimations come form the δ-GLMB filter, values are not always accessible for each step. However, this absence of estimations is traduced by better mean estimate, as shown in table 3.4. Here, mean values are not corrupted with exceptional values, given by the previous methods when signal amplitude is very low or too noisy. For 6 mm defects, ToF are even closer to theoretical values than the wavelet method. For 3 mm defects, estimated values are near the expected value, and higher than the ones from previous experiments.

Association between measurements allows thus to increase estimates of features extracted, providing more reliable use of EMAT signals, even when measurement process worsen.

3.5.4 *Defect location estimation with echoes*

For the previous associations made, extraction of echo tracks allowed estimations of defects locations. For each individual extracted track, estimations of the defect are averaged, creating one single point. Borders have positions 0 cm and 70 cm, and electro-erosion defects positions 12.8 cm, 27.8 cm, 42.6 cm an 57.9 cm. Tables 3.5, 3.6 and 3.7 gather estimations for the static experiment, dynamic with $PRF = 25$ and dynamic with $PRF = 100$. For purposes of clarity, only the 7 tracks with higher lengths are reported (targets which survived more

than 7 steps), if the total number of echo tracks is too high. In addition, to ease comparisons between estimated and true defect locations, for track, the closest defect location is displayed.

	Length	Mode	Estimate (cm)	Closest (cm)
Path 1	12	2 (front)	56.5	57.9
Path 2	12	2 (front)	69.0	70
Path 3	9	1 (back)	12.3	12.8
Path 4	5	2 (front)	45.9	42.6
Path 5	4	1 (back)	-1.33	0
Path 6	4	1 (back)	8.44	12.8
Path 7	4	2 (front)	31.8	27.8

Table 3.5: Characteristics of extracted echoe tracks, with their length, mode, estimation of position of defects and actual position of the closest defect, for the $70\,cm$ rail for static acquisitions.

	Length	Mode	Estimate (cm)	Closest (cm)
Path 1	14	2 (front)	57.7	57.9
Path 2	6	1 (back)	11.1	12.8
Path 3	5	1 (back)	6.5	12.8
Path 4	4	1 (back)	-2.02	0
Path 5	3	2 (front)	25.6	27.8
Path 6	3	2 (front)	26.5	27.8
Path 7	3	2 (front)	71.9	70

Table 3.6: Characteristics of extracted echoe tracks, with their length, mode, estimation of position of defects and actual position of the closest defect, for the $70\,cm$ rail for dynamic acquisitions with $PRF = 25$.

	Length	Mode	Estimate (cm)	Closest (cm)
Path 1	15	2 (front)	57.2	57.9
Path 2	8	2 (front)	29.9	27.8
Path 3	6	1 (back)	11.7	12.8
Path 4	5	1 (back)	12.4	12.8
Path 5	5	2 (front)	38.3	42.6
Path 6	5	1 (back)	48.3	42.6
Path 7	5	2 (front)	58.8	57.9

Table 3.7: Characteristics of extracted echo tracks, with their length, mode, estimation of position of defects and actual position of the closest defect, for the $70\,cm$ rail for dynamic acquisitions with $PRF = 100$.

Remarkably, estimations provided by the echo tracks are nearly all close to a true defect, with precision around $1\,cm$. These experiments have therefore proven the possibility to identify and characterize defects with use of reflection pulses. Such tracks in the B-scan indicate the presence of cracks on the inspected surface, and estimated information allows estimation of their location on the rail.

3.6 Mobile device inspection

3.6.1 *Measurement trolley configuration*

To perform dynamic real-time acquisition on rail, development of a moving measurement trolley involved multiple fields: mechanical and electrical engineering. From a signal processing point of view, each previous level is a possible source of interference, perturbation or corruption of the data coming from sensor. In comparison with a static measurement experiment in a safe environment, such as laboratories,

dynamic acquisition of ultrasonic signals contain observations with such abnormal events, and potentially weak knowledge about their origin or underlying generation processes.

A mechanical structure maintain three transducers, one emitter and two receivers on a relative position above the rail. Reference is not the rail surface by itself, but the trolley moving on the rail, guided by the rail head.

Two different positions of the sensors have been used:

1. Inspection of the top rail head (left figure 3.26). This scenario is the case similar to laboratory experiments, and dedicated to squat or large cracks detection.

2. Head-check inspection (right figure 3.26). The inspected areas are the internal edges of the head rail, where head-checks are more prone to emerge, under the rolling area.

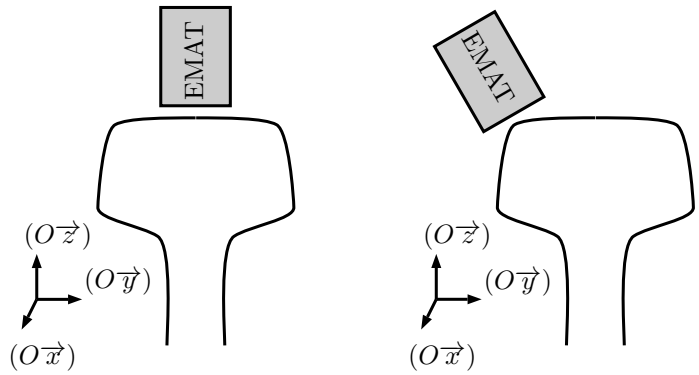

Figure 3.26: EMAT sensors configuration. Left: top rail. Right: internal edge of the rail.

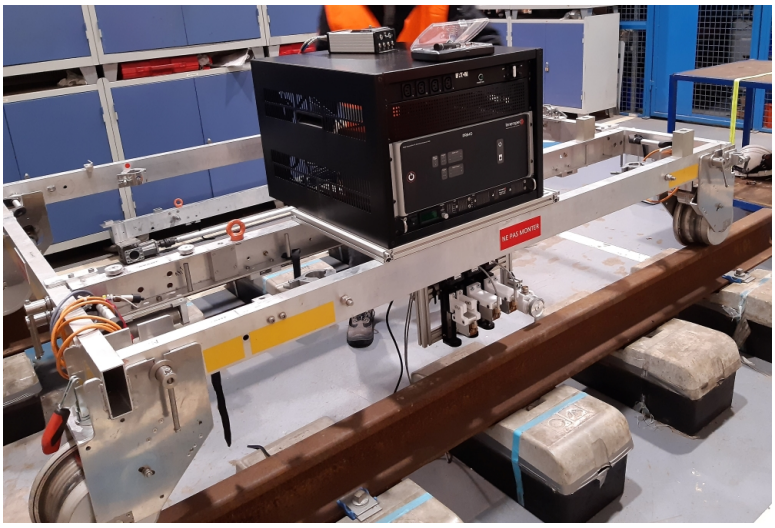

Figure 3.27: Measurement trolley used for Eurotunnel experiments - RAILENIUM.

The measurement trolley is shown in figure 3.27. The only contact between the full structure and rails are the wheels. Height of sensors (lift-off) must be manually adjusted. An odometer is in direct contact with a rail, but does not provide any constraint. Its purpose is only to

send the system the distance travelled by the trolley. Using this information, the measurement system performs one acquisition every $1\,cm$, to reproduce the precision level obtained in laboratory. Displacement of the trolley has to be made manually. Ensuring low speed is necessary to prevent sliding of the odometer wheel. Speed up to $1\,m/s$ have been obtained.

During field inspection, it is possible to use one additional sensor, which will act as a second receiver. Figure 3.28 specifies the configuration of sensors: receiver 1 is close to the emitter, and receiver 2 located in front of the device.

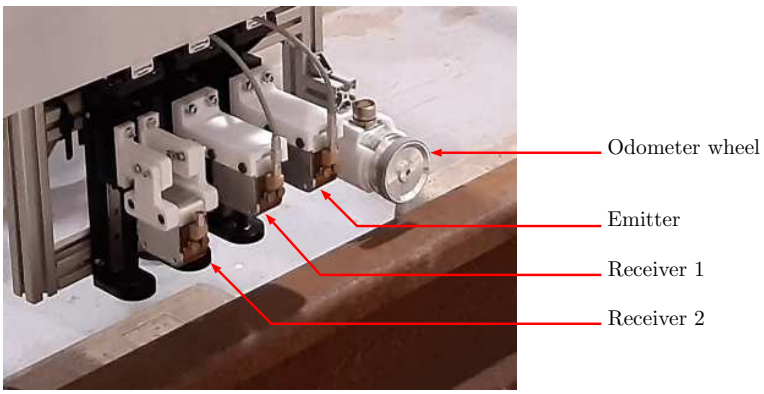

Figure 3.28: Configuration of the sensors before their positioning on the rail - RAILENIUM.

3.6.2 *Field measurements with the trolley*

The measurement trolley presented in previous subsection (figure 3.27) has been used for field measurements on tunnel rails of Eurotunnel (figure 3.29). The following sections present analyses of measurements from EMAT sensors during these experiments, with application of our methodologies.

3.7 Head-check detection with multiple sensors

3.7.1 *Acquisition comparisons*

Analysis made with two receivers, positioned sequentially after the emitter bypasses a common issue with pulses analysis: the lack of reference for amplitude and ToF. If the distance between the two emitter is fixed, the difference of ToF is kept constant over acquisitions. Similarly, ratio of signal amplitudes remains equally constant, even with low variations of lift-off. Issues can nevertheless be faced, when one or several sensors register very low amplitude signals, no measurement, or low SNR situations. Denoising is therefore a critical step, but it has to be coupled with indicators to assess the validity of calculated quantities.

Considering two simultaneous acquisitions made at step k, the ToF of the main pulses are noted, for two receivers, u_k^1 and u_k^2, and their respective amplitudes A_k^1 and A_k^2. The second receiver is assumed to

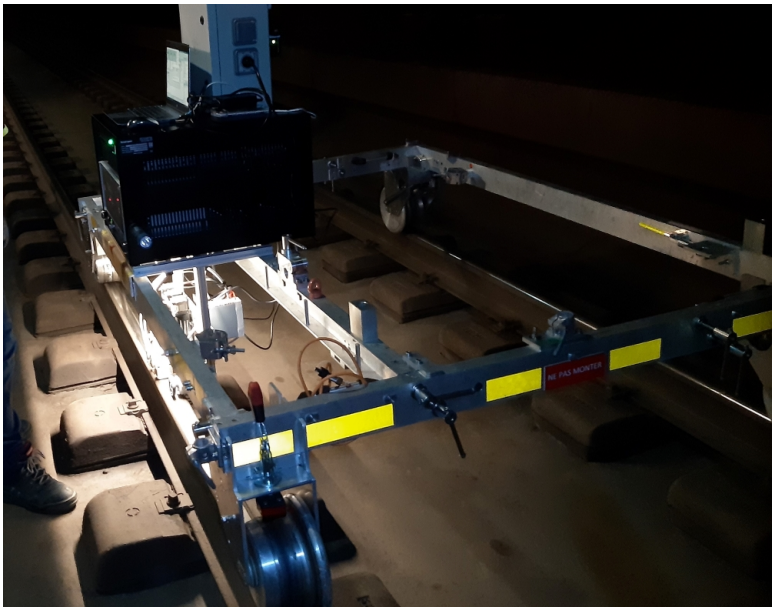

Figure 3.29: Measurement trolley built by RAILENIUM during tunnel experiments at Eurotunnel - RAILENIUM.

be the furthest from the emitter, with $u_k^1 \leq u_k^2$. Depending on the individual gain of each sensor, the same relationship is not applicable to amplitudes. The difference of ToF is noted:

$$\Delta u_k = u_k^2 - u_k^1 \tag{3.7.1}$$

The ratio of amplitudes, which can be inferior to 1 (amplitude of the second sensor superior to the amplitude of the first sensor), is noted:

$$r_k = \frac{A_k^1}{A_k^2} \tag{3.7.2}$$

3.7.2 *Head-check detection on a small rail portion*

This experiment was conducted on a area of $2.39\,m$. First half portion, until approximately $1.20\,m$, is healthy. The remaining rail is subject to head-check, until the end of the inspected area. Distance between emitter and first receiver is set to $10\,cm$, and distance between the two receivers to $10\,cm$. B-scans (figure 3.30) illustrate the slow decay of the signal, with no visual change of ToF. The gain of receiver 2 has been increased to reach amplitudes similar to the ones of receiver 1. However, a negative consequence is that the noise content is therefore higher.

Figure 3.31 illustrates the naive use of the direct output of the decomposition algorithm. Comparison indicators are calculated directly by taking the atoms with higher amplitudes. Modifications of indicators are observable from the middle to the end of area, but abnormal values could also be links to wrong estimations of the atom's parameters. Vertical red line indicates the approximate change of rail condition: left side is considered as healthy, and right side as subject to head-checking.

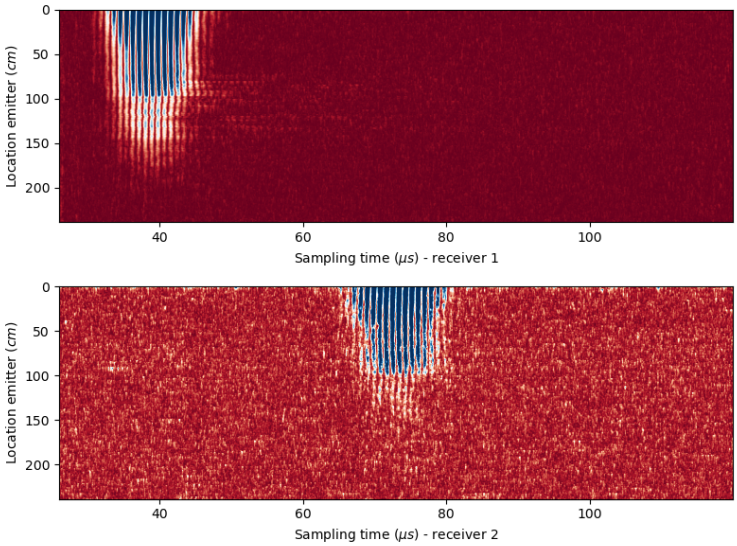

Figure 3.30: B-scan of acquisitions of the two receiver on the small area subject to head-checking.

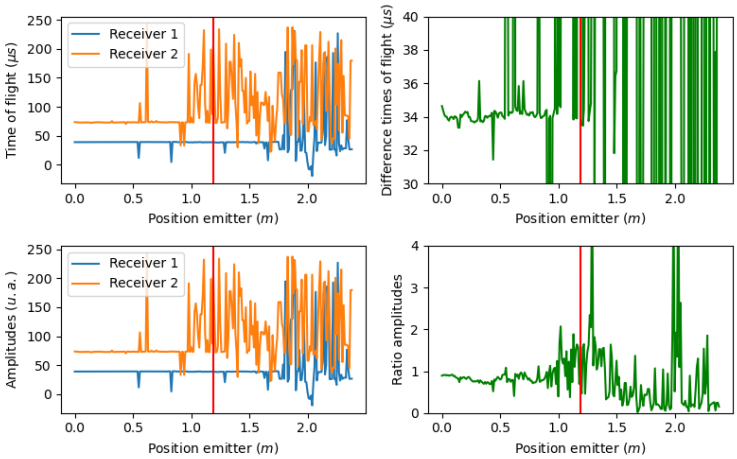

Figure 3.31: Estimation of comparison indicators on the small area subject to head-checking: left side: healthy, right side: damage.

These short examples clearly identify the major issue with use of classical feature extraction methods. The absence of information in the signal, here the presence of ultrasonic pulse, is not detected. Consequently, estimations could be made on noises or external perturbations of sensors.

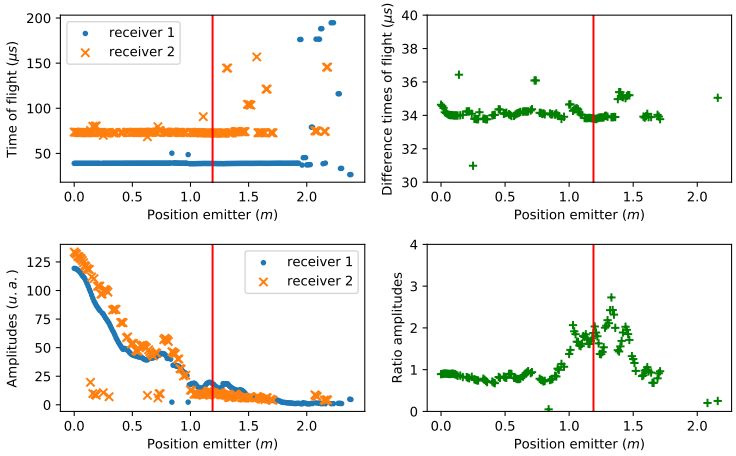

Figure 3.32: Estimation of comparison indicators on the small area subject to head-checking after filtering processes: left side: healthy, right side: damage.

After analysis with filters, nearly no more tracks are identified after position 1.7 m on the rail (position of the emitter). The RFS filter does not succeed in finding any pulse track. This indicator suggests that the signal has completely been lost after this point. Analysis of figure 3.32 indicates a slow change of amplitude during the first half of the rail, but with a constant amplitude ratio. Time of flights do not evolve on this part. However, amplitude ratio is doubled at the centre of the rail, where head-check appears on the rail surface. Time of flight difference undergoes a slight increase in the head-check region,

3.7.3 *Head-check detection on large rail portion*

A larger area of rail has been inspected to analyse performances of the sensors and processing algorithm with a harder task. Measurements are performed with two receivers, on the internal rail edge on an interval of 2.8 m (for a total of 280 acquisitions per receiver). The related B-scan is shown in figure 3.33.

The inspected rail can be separated into three main parts:

- Healthy section, from 0.0 m to 0.6 m. This portion suffers a light head-check, but is assumed to be the reference area. In figure 3.34 a top view of the rail is shown. Head-check is observable at the image bottom. Inspection is realized on this part of the rail.

- Medium head-check, from 0.6 m to 1.7 m. Here, head-check developed on a 1 cm band near the rolling surface. Cracks are separated by short distances, from 2 mm to 5 mm. They appear to be parallel, with maximum orientation of 45° (0° refers to a crack which goes directly from the edge of the rail to the centre, and 90° to a crack in

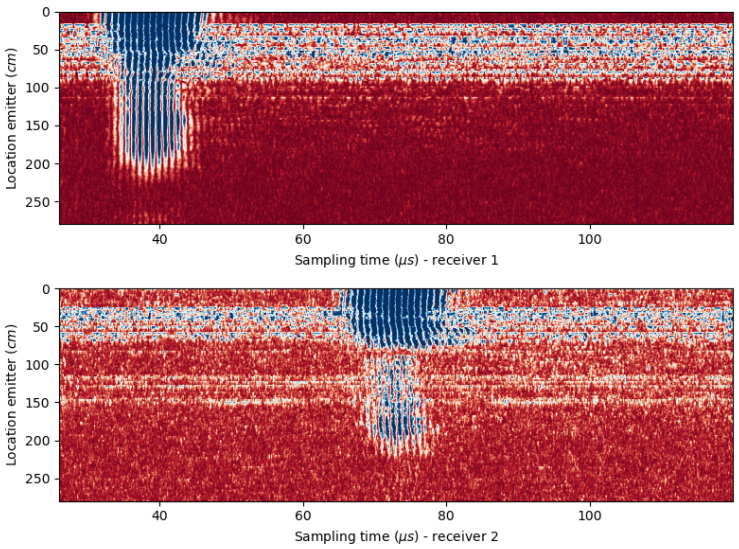

Figure 3.33: B-scan of acquisitions of the two receiver on the small area subject to head-checking.

Figure 3.34: Top view of healthy rail section for a large inspection. Some grinding marks are observable on the internal rail edge (bottom edge of the image).

the longitudinal direction of the rail). This phenomenon is shown in figure 3.35.

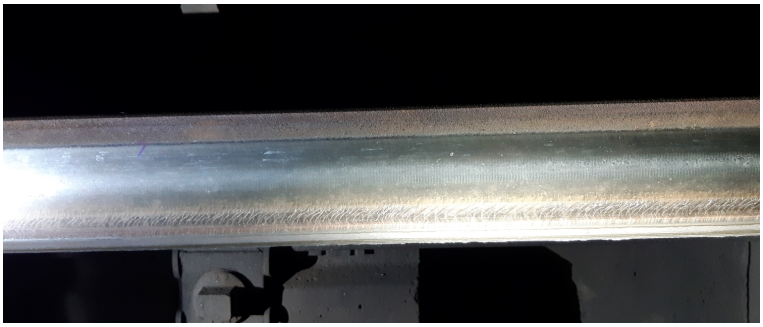

Figure 3.35: Medium head-check rail section for a large inspection, with thin cracks close to each other.

- Severe head-check, from 1.7 m to 2.8 m. Cracks developed after the rolling surface (near edges), and are separated by medium distances, from 5 mm to 10 mm. The defects spread on both the edge and the top of the rail: the cracks present a curvature, and get orientation in the longitudinal direction of the rail when they approach the centre of the rail. Figure 3.36 illustrates this behaviour.

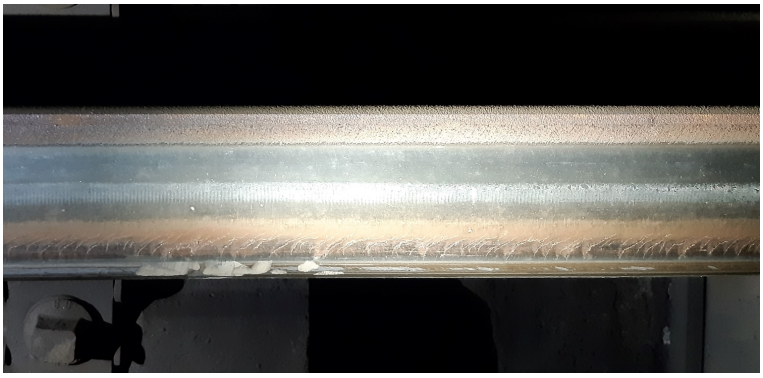

Figure 3.36: Severe head-check rail section for a large inspection, with wide and curved cracks.

Again, 3.37 illustrates the instability of indicators constructed directly with results of the decomposition. Measurement noise corrupts the signal, producing abnormal estimations of ToF. Amplitude is less impacted by this phenomenon, but shows instabilities for the two receivers. Since the second sensor receives a signal with higher attenuation (in comparison with the first sensor), higher disturbances are expected for it.

Results of the filtering process show important dependences between the computed features and the rail condition, as demonstrated in figure 3.38. Amplitude is subject to a strong attenuation, to reach a constant level on the middle part of the rail, subject to medium head-checking. The final part causes the strongest attenuation of the signal, but a pulse track is still identified. Access to measurements allows a clear separation between the three parts based on analyse of the signal attenuation. Difference of ToF only indicates abnormal points at the beginning of the second part, with medium head-checking. This could

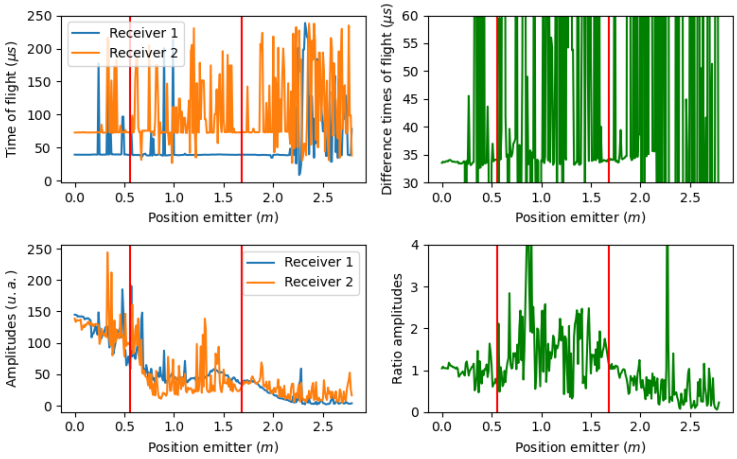

Figure 3.37: Estimation of comparison indicators on the large area subject to head-checking. Left part: healthy, middle part: moderate damage, right part: severe damage.

indicate a detection of the change of material properties. The slight increase in the third part, with severe head-checking, could be related to the increase of path for the Rayleigh wave, to circumvent deeper cracks. Ratio of amplitudes shows a strong increase during the second part of the rail. The absorption of energy made by cracks could highlight a non-linear effect of attenuation, specific to cracks close to each other.

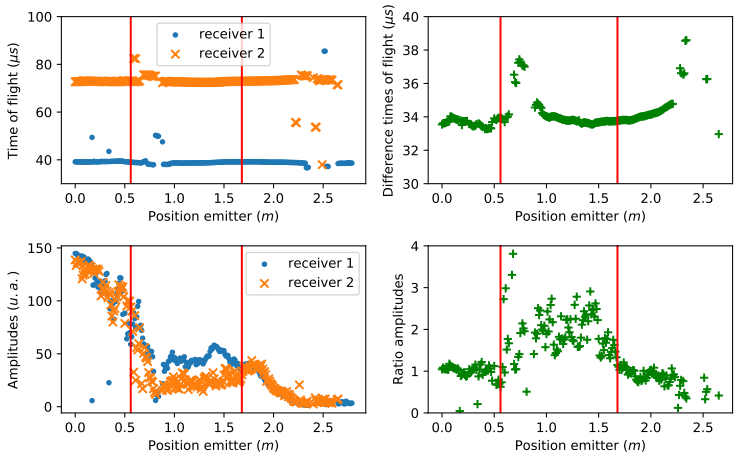

Figure 3.38: Estimation of comparison indicators on the large area subject to head-checking after filtering processes. Left part: healthy, middle part: moderate damage, right part: severe damage.

3.8 Conclusion

This chapter has introduced our methodology for inspection of rails with statistical and probabilistic techniques. First, a measurement signal is decomposed into a sum of Gabor atoms, creating as measurement set with related parameters. This set is next injected in an association scheme, performed by a RFS tracker. We show the relevance of inspection of ultrasonic signal and separation of pulses according to their behaviours across acquisitions. Echoes resulting from reflections on cracks can then be identified. The tracker allows separation between pulses, and use of their information for characterization

of defects. In this chapter, we proposed a full methodology to solve the parameter choice problem in filtering estimations, in a complete automatic aspect. The main objective, maximizing the measurement likelihood, requires few operator decisions. Going through the Poisson approximation of the intractable multi-target posterior densities allows fast and accurate estimations. Simulations based on a well-known and documented example demonstrated the validity of both approximation and methodology.

Our methodology and formulation have been tested and approved on simulated laboratory situations. Our applications on the machined and electro-erosion rail samples with EMAT sensors demonstrated the ability of trackers to improve detection of defects, and estimations of key features for their characterization. Field experiments conducted on the tunnel rails allowed the application of our methodology to more challenging data. Damaged areas have been inspected, with different severity levels of head-checking. Measurement perturbations is managed by both the sparse decomposition method and the filtering process. Use of multiple sensors exhibits important changes between signals when head-check is interacting with ultrasonic pulses.

The filtering analysis of signal could however be applied on many other infrastructures. Our application targeted inspection of rails, but different parts of the railway infrastructure are subject to vibrations. Filtering methods, using state space modelling, are applicable to, for instance, inspection of railways sleepers, which have recently raised interest [Hoa17]. Mediums like concrete, irons or woods are indeed subject to vibratory perturbations, which could be perturbed by defects, cracks or changes of the material property.

Local characterization of signal for non-contact defects detection

4

4.1 Introduction

Previous chapter has analysed ultrasonic signals with estimation methods which concentrate information into a small set of features, or even into a single indicator, such as the ToF. However, general signals resulting from physical process do not necessarily fulfil the conditions to be summarized with such low dimensional representation. Notably, with EMAT signals, the instantaneous frequency $\nu(t)$ computed from the analytic signal exhibits specific variations around the envelope local maxima (it follows a trend similar to the one of the envelope). For signals which could be not sparse in a Gabor dictionary, it would be more useful to extract local information which characterizes locally the signal, to be next used by a future process. This chapter aims at studying a generalization of the previous methods, which could be employed in vibratory applications. Rail analysis would be one specific application, but no complex behaviour of the signals has been registered with our sensors. It is thus necessary, to experience such phenomena, to study other vibratory systems. We had the opportunity to work with the University of Waterloo (ON, Canada) on *resonant-column* experiments, which exhibit highly non-linear behaviours of soil samples. Our objective is still to offer convenient estimation methods, with little decision to be made about the tuning aspect. We took the methodologies resulting from the previous chapter, and adapt our processes. Section 4.2 introduces the fundamental equations governing vibratory signals and simulation processes related to it. Section 4.3 presents the estimation tools developed to answer the feature extraction problem from the previous equation. Last section 4.4 summarizes the results on resonant-column experiments, interpretation and comparison with existing methodologies.

4.2 From set to vectors

4.2.1 *The resonant-column for soil characterization*

Resonant-column testing is a powerful tool for dynamic characterization of soil specimen under different confinements. Application in geotechnical sciences allow the understanding of soil behaviours during extreme events, such as earthquakes [FM20].

To measure non-linear behaviours of the material, a soil specimen is placed on a confined environment, with air pressure controlled by an external pump. The specimen is constrained by two plates. The bottom one is fixed, and the top one free. An accelerometer and two

coils are connected to the top plate. A generator is used to control the two coils with an external field, allowing an operator to control the vibration of the plate. A moment can therefore be transmitted to the specimen. The full system is represented in figure 4.1. After reaching a stationary situation, coil excitation is interrupted, and the specimen becomes subject to a *single degree of freedom* (SDOF) model. Depending on the maximum level of excitation selected for the stationary situation, its characteristics change.

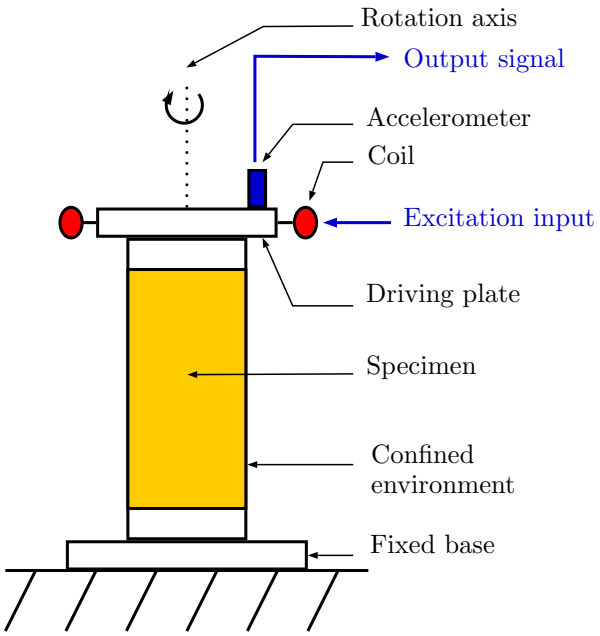

Figure 4.1: Resonant Column experiment with soil specimen and external excitation by coils, with measurement coming from the accelerometer sensor

4.2.2 *The Single degree of freedom model*

The soil sample can be represented as a mass subject to rotational movements. A basic rotating system is characterized by its angular position α which is function of time t, an inertia J, a stiffness K, an external moment M and a viscous coefficient V, as shown in figure 4.2. The latter coefficient, if positive, is responsible for energy dissipation. The system is defined by the following equation:

$$J\ddot{\alpha}(t) + V\dot{\alpha}(t) + K\alpha(t) = M(t) \tag{4.2.1}$$

The system is mostly found in the literature with its translation form in figure 4.3, with measure of position $x(t)$ instead of angle $\alpha(t)$.

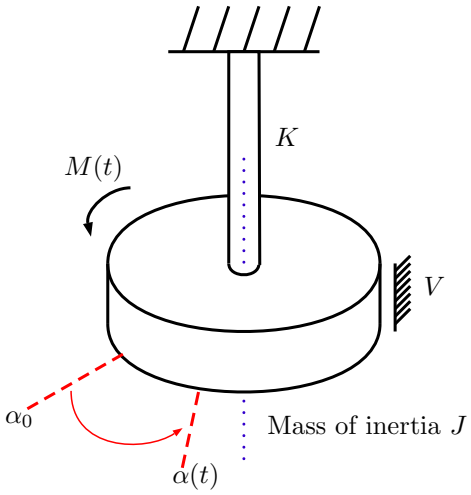

Figure 4.2: Rotational system with inertia J, viscous coefficient V, stiffness K, initial angle α_0 and external moment $M(t)$. This system is equivalent to the translation one described in figure 4.3.

With no external force applied on the system, this equation becomes the SDOF equation:

$$J\ddot{\alpha}(t) + V\dot{\alpha}(t) + K\alpha(t) = 0 \tag{4.2.2}$$

which is written in the reduced form:

$$\ddot{\alpha}(t) + 2D\omega\dot{\alpha}(t) + \omega^2\alpha(t) \tag{4.2.3}$$

with $D = V/2\sqrt{JK}$ the damping ratio and $\omega = \sqrt{K/J} = 2\pi f$ the system frequency. The set $\{D, f\}$ is referred to as the dynamic parameters of the system. An analytical solution is available, depending on the sign of the discriminant:

$$\Delta = 4\omega^2\left(D^2 - 1\right) \tag{4.2.4}$$

Value of the damping ratio D influences the sign of this discriminant. Case obtained with the condition $D > 1$ is called the aperiodic behaviour, and the specific case $D = 1$ the critical aperiodic behaviour. The case of interest is the one with $D < 1$, which causes the discriminant to be negative, and thus the roots to be complex, noted by λ_+ and λ_-, defined as:

$$\lambda_\pm = -D\omega \pm i\sqrt{\omega^2\left(1 - D^2\right)} \tag{4.2.5}$$

Analytical solution for α takes the form, with two variables $C_+, C_- \in \mathbb{C}$:

$$\alpha(t) = C_+ e^{\lambda_+ t} + C_- e^{\lambda_- t} \tag{4.2.6}$$

To obtain a particular solution (find values of C_+ and C_-), it is necessary to include solutions of this equation. Generally, origin values of α and its first derivative are used ($\alpha(0) = \alpha_0$ and $\dot{\alpha}(0) = \dot{\alpha}_0$), providing:

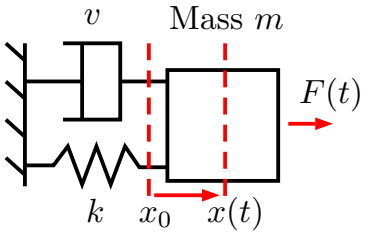

Figure 4.3: Translation system with mass m, viscous coefficient v, stiffness k, initial position x_0 and external force $F(t)$.

Not to be confounded with the PHD function

$$C_+ = \frac{\dot{\alpha}_0 - \alpha_0\lambda_+}{\lambda_+ - \lambda_-} \tag{4.2.7}$$

$$C_- = \frac{\dot{\alpha}_0 - \alpha_0\lambda_-}{\lambda_- - \lambda_+} \tag{4.2.8}$$

The following expression for α can be derived:

$$\alpha(t) = \frac{e^{-D\omega t}}{\omega\sqrt{1-D^2}}\left[\alpha_0\omega\sqrt{1-D^2}\cos\left(\omega\sqrt{1-D^2}\right) + (\dot{\alpha_0} + \alpha_0 D\omega)\sin\left(\omega\sqrt{1-D^2}\right)\right] \tag{4.2.9}$$

The equation with only one cosine is preferred:

$$\alpha(t) = Ae^{-D\omega t}\cos\left(\omega\sqrt{1-D^2} + \phi\right) \tag{4.2.10}$$

with the following variable calculated from the trigonometric identities:

$$A = \text{sign}(\alpha_0)\sqrt{\dot{\alpha_0}^2 + 2\alpha_0\dot{\alpha_0}D\omega + \alpha_0^2\omega^2} \tag{4.2.11}$$

$$\phi = \arctan\left(-\frac{\dot{\alpha_0} + \alpha_0 D\omega}{\alpha_0\omega\sqrt{1-D^2}}\right) \tag{4.2.12}$$

As a reminder, the sign function *sign* is defined, $\forall x \in \mathbb{R}$ by :

$$sign(x) = \begin{cases} 1 & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x = 0 \\ -1 & \text{for} \quad x < 0 \end{cases}$$

4.2.3 *Estimation of constant dynamic parameters* $\{D, f\}$

The measurement of attenuation is contained inside the damping ratio. Estimation of this dynamic parameter is simply achieved by calculating the attenuation of two consecutive peaks. If we note t_1, t_2 the times of occurrence of these peaks, and α_1, α_2 their amplitudes, a rough estimation of the damping is given by:

$$\hat{D} = \frac{\log(\alpha_1) - \log(\alpha_2)}{\omega(t_2 - t_1)} \tag{4.2.13}$$

Unfortunately, since the estimation is reduced to comparison of only two points of a signal, it is therefore easily disturbed by measurement noises. For low amplitude perturbations, with known frequency, an extraction of the signal envelope combined with a decreasing exponential curve provides a slightly more robust estimation of D. However, since the Hilbert transform does not preserve Gaussian distribution, it should not be used in low SNR scenarii. A least-square fitting, with estimation of the full vector $[A, D, f, \phi]^T$ is more robust, although more demanding in terms of computational resources.

Bayesian inference of the dynamic parameters is also possible with sequential filtering methods. For a linear transition model, a Kalman filter provides estimation of the angle positions $\alpha_k, k = 1, .., N$ and possibly its derivatives (depending on the model complexity), and an estimation of the likelihood of the measurement vector $z_{1:N}$ for a given parameter vector $[D, f]^T$. Since the recursion equations for the Kalman filter allow a direct calculation (analytically and exact), a MH algorithm provides an estimation of the posterior density $p(D, f \mid z_{1:N})$. Only choices remaining to the operator is the determination of the covariance matrices Q and R, and the initial state uncertainty (modelled by the distribution $\mathcal{N}(x_0, P_0)$). The full procedure with a random-walk proposal in this specific case is described in algorithm 2.

These filter's parameter can also be included in the estimation process, in this case the MH sample vectors $\vartheta = [D, f, Q, R, X_0, P_0]^T$

Algorithm 2: Metropolis-Hastings sampling with likelihood estimation from Kalman filters

Input : Number of iterations K, Prior distribution $p(\vartheta)$, Initial value ϑ_0, Proposal covariance matrix Q^{prop}
Output: Posterior samples $\{\vartheta_k\}_{k=1}^{K}$
Data: Measurements $z_1, \ldots, z_N$
Initialization : $\mathcal{L}_0 = \text{Kalman}(\vartheta_0, z_{1:N})$
$k = 1$
while $k \leq K$ **do**
  $\vartheta^{\text{prop}} \sim \mathcal{N}(\vartheta_{k-1}, Q^{\text{prop}})$
  $\mathcal{L}_{\text{prop}} = \text{Kalman}(\vartheta_{\text{prop}}, z_{1:N})$
  $\alpha = \frac{\mathcal{L}_{\text{prop}} p(\vartheta_{\text{prop}})}{\mathcal{L}_{k-1} p(\vartheta_{k-1})}$
  $u \sim \mathcal{U}([0,1])$
  if $u < \alpha$ **then**
    $\vartheta_k = \vartheta_{\text{prop}}$
    $\mathcal{L}_k = \mathcal{L}_{\text{prop}}$
  else
    $\vartheta_k = \vartheta_{k-1}$
    $\mathcal{L}_k = \mathcal{L}_{k-1}$
  $k = k + 1$

4.2.4 *Variables of interest in resonant-column experiments*

Instead of using directly the angular position α, geotechnical sciences have preferred the dimensionless variable γ called the *shear strain*. It can be calculated as:

$$\gamma = \frac{d}{h} \tag{4.2.14}$$

with h the height of the specimen, and d the maximum displacement at the radius r of the specimen, with therefore $d = \alpha\, r$. Resonant-column testing are dedicated to study of low level strain ($\gamma \leq 1e^{-2}$).

One other important variable is the *shear modulus* G, defined as the product of the density of the medium ρ and the square shear wave velocity V_S. The latter is in the context of resonant-column experiments proportional to frequency f (or ω). The maximum shear modulus $G_{\max}$ is therefore associated to the maximum frequency $f_{\max}$ found during experiments. A variable of interest is the ratio of shear modulus and maximal shear modulus, defined as the *shear modulus ratio*:

$$\frac{G}{G_{\max}} = \frac{\omega^2}{\omega_{\max}^2} = \frac{f^2}{f_{\max}^2} \tag{4.2.15}$$

Traditional results from resonant-column experiments have shown hyperbolic relationships between damping ratio, shear modulus ratio and strain. Damping is expected to approach 0 for small γ, and increases with this variable. On the contrary, shear modulus ratio is close to 1 for small strains and decreases for higher values. These relationships, *degradation curves*, are shown in figures 4.4 and 4.5.

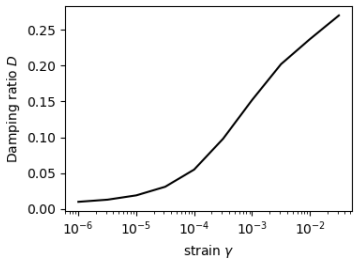

Figure 4.4: Evolution of damping ratio with strain, for rock with plastic index of 0, data from [VD91].

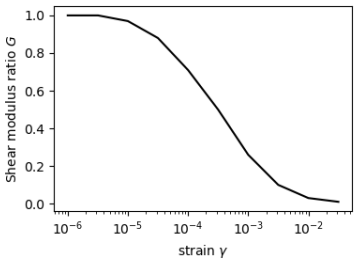

Figure 4.5: Evolution of shear modulus ratio with strain, for rock with plastic index of 0, data from [VD91].

4.2.5 *Single degree of freedom step model*

Starting from the SDOF equation, it is possible to directly deduce state space transition equations. Including information from the angular position and its first derivative is sufficient for a first order approximation. Let $x = [\alpha, \dot{\alpha}]^T$ be the continuous state vector. A transition function will find the best way to transit from the state vector x_k at step k to the following one x_{k+1} at step $k+1$. Time interval between the two steps corresponds to the value Δt.

The starting point for discretization is the general form of the continuous differential equation:

$$\dot{x}(t) = \mathbf{f}(t, x(t)) \tag{4.2.16}$$

with $\mathbf{f}$ any function which expresses the derivative of the state vector for a specific time $t \in \mathbb{R}$. An additional condition is generally added, in the form of the equality $x(t_0) = y_0$ for some $t_0 \in \mathbb{R}$.

The Euler method is a first order approximation to solve the differential equation 4.2.16, widely used in sequential estimation frameworks applied to analysis of mechanical structures or cinematic phenomena. The solution vector at step k is approximated by:

$$x_k = x_{k-1} + \Delta t\, \mathbf{f}(t_{k-1} + \Delta t, x_{k-1}) \tag{4.2.17}$$

In the SDOF equation, the function $\mathbf{f}$ is independent of time t_{k-1}, linear, and entirely defined by the matrix F_0 such that:

$$F_0 = \begin{bmatrix} 0 & 1 \\ -(2\pi f)^2 & 4D\pi f \end{bmatrix} \tag{4.2.18}$$

Using previous matrix F_0, the transition equation 4.2.17 becomes:

$$x_k = F^{\text{Euler}}\, x_{k-1} \tag{4.2.19}$$

with the transition matrix defined as:

$$F^{\text{Euler}} = I_2 + \Delta t\, F_0 \tag{4.2.20}$$

Where I_n referred to the identity matrix of dimension $n \times n$, for $n \in \mathbb{N}^*$, with 1 on the diagonal, and 0 elsewhere.

This first order approximation is sufficient for a very short time interval Δt. However, quality of estimation decreases with higher values, possibly causing biases in simulation aiming at solving the equation 4.2.16 for specific time steps, and even divergences.

The Runge-Kutta methods allows higher order of approximation [But00]. Despite the induced higher computational cost, the achieved estimations are much more accurate than with the Euler method. Here, the fourth order method (RK4) will be presented:

$$k_1 = \mathbf{f}(t_{k-1} + \Delta t, x_{k-1}) \tag{4.2.21}$$

$$k_2 = \mathbf{f}(t_{k-1} + \frac{\Delta t}{2}, x_{k-1} + \frac{\Delta t}{2} k_2) \tag{4.2.22}$$

$$k_3 = \mathbf{f}(t_{k-1} + \frac{\Delta t}{2}, x_{k-1} + \frac{\Delta t}{2} k_3) \tag{4.2.23}$$

$$k_4 = \mathbf{f}(t_{k-1} + \Delta t, x_{k-1} + \Delta t\, k_3) \tag{4.2.24}$$

$$x_k = x_{k-1} + \frac{\Delta t}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{4.2.25}$$

Assuming D and f to be constant during the interval Δt, the transition equation 4.2.17 becomes:

$$x_k = F^{\mathrm{RK4}}\, x_{k-1} \tag{4.2.26}$$

with the transition matrix defined as:

$$F^{\mathrm{RK4}} = I_2 + \Delta t\, F_0 + \Delta t^2 \frac{F_0^2}{2} + \Delta t^3 \frac{F_0^3}{6} + \Delta t^4 \frac{F_0^4}{24} \tag{4.2.27}$$

Of course, when an analytical form is available for equation 4.2.16, it should always be preferred to these sequential resolution methods, which are dependent on the machine precision, and propagate any error across steps, with risk of amplifying it.

4.3 Estimation of time-varying dynamic parameters

4.3.1 *Non-linear filtering*

Unfortunately, the previous hypothesis stating that the dynamic parameter are constant over all the experience is no longer admissible in several cases. Notably, in resonant column experiments, complex phenomena induce a variation of damping and frequency during the first oscillations, just after the system is released (first moment of validity of the SDOF model). Even worst, no abrupt change happens, meaning it is impossible to identify independent blocks of measurement, each one related to one specific and constant dynamic parameter. Estimation of the trend of these parameters should therefore be made jointly with the estimation of the system state, at each step k.

Including filter's parameters inside the vector state to be estimated is a common technique, which aims at slowly converging to the parameter vector posterior distribution. Since the relationship between the state space and the parameter state is complex, a particle filtering is the preferred technique used in this situation. However, in practice, the design of such filter is difficult, and a pure random-walk model for the parameter evolutions doesn't necessarily imply a convergence to the real posterior. A Gaussian proposal with decaying covariance is the simplest solution to this issue. The Liu and West filter aims at reducing the variance in the sequential parameter posteriors with calculation of sufficient statistics, and showed improved performances in comparison with the previous technique [LW01].

Perturbing the parameter space with a Gaussian noise is also called *jittering*.

Paradoxically, techniques which transform the problem into estimation of time-varying parameters were developed to manage the static parameter cases. Results were then concentrating into obtaining single values related to them, with no consideration of the step influence. Our objective is thus to provide estimation processes to perform estimation of slow varying dynamic parameters with quantification of the uncertainty and identification of their evolution.

A SDOF model combined with random-walk transition model for dynamic parameters gives:

$$D_k = D_{k-1} + \epsilon_k^D \tag{4.3.1}$$
$$f_k = f_{k-1} + \epsilon_k^f \tag{4.3.2}$$
$$x_k = F\left(D_k, f_k\right) x_{k-1} + \eta_k \tag{4.3.3}$$

with $\epsilon_k^D \sim \mathcal{N}(0, \sigma_D^2)$, $\epsilon_k^D \sim \mathcal{N}(0, \sigma_f^2)$ and $\eta_k \sim \mathcal{N}(0, Q_x)$.

To ensure positiveness of the dynamic parameters, a model with variations on the logarithm scale is chosen. It includes a trend variable, to allow more flexibility:

The variables r^D and r^f should more be seen as increase or decrease rates.

$$\log\left(D_k\right) = \log\left(D_{k-1}\right) + \Delta t\, r_{k-1}^D + \epsilon_k^D \tag{4.3.4}$$
$$\log\left(f_k\right) = \log\left(f_{k-1}\right) + \Delta t\, r_{k-1}^f + \epsilon_k^f \tag{4.3.5}$$
$$r_k^D = r_{k-1}^D + \epsilon_k^{r^D} \tag{4.3.6}$$
$$r_k^f = r_{k-1}^f + \epsilon_k^{r^f} \tag{4.3.7}$$
$$x_k = F\left(D_k, f_k\right) x_{k-1} + \eta_k \tag{4.3.8}$$

with $\epsilon_k^{r^D} \sim \mathcal{N}(0, \sigma_{r^D}^2)$ and $\epsilon_k^{r^f} \sim \mathcal{N}(0, \sigma_{r^f}^2)$.

The transition equation of the dynamic parameter vector

$$p\left(u_k \mid u_{k-1}\right) = p\left(D_k \mid D_{k-1}\right) p\left(f_k \mid f_{k-1}\right) \tag{4.3.9}$$

is then replaced by:

$$p\left(u_k \mid u_{k-1}\right) = p\left(D_k \mid D_{k-1}, r_{k-1}^D\right) p\left(f_k \mid f_{k-1}, r_{k-1}^f\right) \tag{4.3.10}$$

For a positive linear variation of D or f, the trends r^D or r^f will undergo an abrupt variation from 0, and slowly return to this specific value. The main interest of such construction is to get first order information from the local evolution of the dynamic parameter, similar to the estimation of velocity with a Kalman filter.

4.3.2 *Rao-Blackwellization*

An important remark related to particle methods is the stochastic aspect involved in the sequential calculation of distributions and statistics. Contrary to analytical methods, such as Kalman filters, every quantity resulting from the particle approximation is a random variable. Notably, data likelihood is only approached at each run of the particle filter, meaning that several parallel calculations would result in different values. The stochastic approximation for parameter ϑ is noted $\widetilde{\mathcal{L}}_\vartheta$ instead of $\mathcal{L}_\vartheta$. First, this specific aspect induces a computational issue for optimization algorithms aiming at maximizing the likelihood. The evaluations of the objective function to optimize could thus be interpreted as being perturbed by a random noise. Under several assumptions established on this corruption, regular methods are still available. In some conditions, stochastic approximations of the gradient can speed-up their convergence. Particle smoothing methods are important classes to improve estimations of their values, with the

draw-back of an increase of computational ressources. Secondly, the Bayesian inference scheme with MH algorithm requires an exact calculation of the data likelihood. Fortunately, it has been proven that the chain recursion can be preserved even with an acceptance ratio obtained from noisy likelihoods, under mild hypotheses [AR09]. The major one needs the estimator of the likelihood to be unbiased. However, the chain can be struck during several iterations if a realization of the likelihood reaches abnormal high values (higher than the expected value which could be computed from repetitions), causing convergence issues. Better performances are expected with reduction of the variance of the likelihood estimator, just as in the case of optimization processes.

A theoretical solution to reduce variance is to increase the number of particles. This approach is supported by asymptotic theorems (limits as $N_p \to \infty$) [Dou+15]. However, this solution is understandably not manageable in practice. Furthermore, it has also been experienced that the number of particle should increase jointly with the data length (or number of measurement N), at least linearly. On a computer with specific memory and processing speed, only a limited number of particles could therefore be handled.

Since one iteration of the particle is $\mathcal{O}(N_p)$, scaling N_p with N would result in an overall complexity of $\mathcal{O}(N^2)$ for a complete filter run.

An interesting property of mechanical system is the linear transition between subsets of the state space vectors. Under well-defined perturbations, such as Gaussian noises, analytical expression for sufficient statistics are available. Consequently, variances of estimators dedicated to these specific variables are reduced, in comparison with the ones obtained from the particle approximation. Such approach is called *Rao-Blackwellization*: introduction of analytical calculations inside the filtering recursions, to bypass the drawbacks of Monte Carlo approximations [DGA00].

Our previous models, also called *conditionally linear model*, can therefore be summarized as:

$$u_k \sim p(u_k \mid u_{k-1}) \tag{4.3.11}$$

$$x_k \sim \mathcal{N}(F(u_{k-1})\, x_{k-1}, Q(u_{k-1})) \tag{4.3.12}$$

$$z_k \sim \mathcal{N}(H(u_k)\, x_k, R(u_k)) \tag{4.3.13}$$

The linear part of the model is handled by Kalman filtering recursions. The non-linear sub-space, related to variable u_k, is the only part of the algorithm which requires particle approximation. This filter is denoted as the Rao-Blackwellized particle filter (RBPF). Using N_p particles, the posterior distribution of the state vector is rewritten as :

$$p(x_k, u_k) \approx \sum_{i=0}^{N_p} w_k^{(i)}\, \mathcal{N}\left(x_k \mid \mu_k^{(i)}, P_k^{(i)}\right)\, \delta_{u_k^{(i)}}(u_k) \tag{4.3.14}$$

with $\mu_k^{(i)}$ and $P_k^{(i)}$ the posterior mean and covariance matrix calculated from the Kalman recursion for particle i, with non-linear space

vector $u_k^{(i)}$ and weight $w_k^{(i)}$. Each particle consists now on the tuple $\left(w_k^{(i)}, \mu_k^{(i)}, P_k^{(i)}, u_k^{(i)}\right)$, which gather the sufficient statistics directly. The full recursion is described in algorithm 3. It should be noted that the resampling step is still necessary, related to the variable *criterion*, which processes the particle weights to decide the need of resampling.

Algorithm 3: Rao-Blackwellized Bootstrap Particle Filter

Input : Prior distribution $p(x_0)$, Number of particles N_p
Output: Posterior distributions $p(x_1|z_1)$, ... , $p(x_N|z_{1:N})$
Data: Measurements z_1, ... , z_N
while $k \leq N$ **do**
  Prediction:
  for $i \leftarrow 1$ **to** N_p **do**
    $u_k^{(i)} \sim p(u_k|u_{k-1}^{(i)})$
    Compute $F(u_{k-1}^{(i)})$ and $Q(u_{k-1}^{(i)})$
    Compute $\mu_{k|k-1}^{(i)}$ and $P_{k|k-1}^{(i)}$
  Update:
  for $i \leftarrow 1$ **to** N_p **do**
    Compute $H(u_k^{(i)})$ and $R(u_k^{(i)})$
    Compute $\mu_k^{(i)}$ and $P_k^{(i)}$
    $\alpha_k^{(i)} = p(z_k|u_k^{(i)}, z_{k-1})$
  for $i \leftarrow 1$ **to** N_p **do**
    $\widetilde{w}_k^{(i)} = \widetilde{w}_{k-1}^{(i)} \alpha_k^{(i)}$
  for $i \leftarrow 1$ **to** N_p **do**
    $w_k^{(i)} = \frac{\widetilde{w}_k^{(i)}}{\sum_{j=1}^{N_p} \widetilde{w}_k^{(j)}}$
  $k = k + 1$
if *criterion* **then**
  Resample $\left\{\left(\mu_k^{(i)}, P_k^{(i)}, u_k^{(i)}\right)\right\}_{i=1}^{N_p}$ and $\left\{w_k^{(i)}\right\}_{i=1}^{N_p}$

4.3.3 *Bayesian inference with Rao-Blackwellized particle filters*

As previously discussed, posterior distributions of the filter's parameters (given measurements) are accessible from MH methods. However, an important result has been demonstrated for particles filters, giving the possibility to compute the joint posterior distribution of both parameters and state space vectors $p(\vartheta, x_{1:N} \mid z_{1:N})$ [ADH10]. Marginalization over the parameter space gives the smoothing distribution $p(x_{1:N} \mid z_{1:N})$. The form of the MCMC algorithm is similar to the MH one, but hypotheses made on the filtering recursion are more restrictive. The posterior and smoothing distributions are particle approximations obtained from the successive iterations.

The algorithm is denoted as the *Particle Marginal Metropolis Hastings* (PMMH) algorithm [ADH10]. It consists in sampling from a proposal a parameter vector, running the particle filter with storage of

all particle paths with their ancestors, and next sampling one particle from the final ones obtained. The version of the PMMH employed for our application is a modified version to handle the RBPF, noted as the Rao-Blackwell PMMH (RBPMMH), and described in algorithm 4. After each estimation of the likelihood with the filter recursions, only the non-linear subset of the state space is however sampled, producing a particle path with ancestors [Pet+11; NPY11]. Conditionally to this vector, it is however possible to perform a smoothing estimation of the linear variables using a RTS smoother.

Algorithm 4: Particle Marginal Metropolis-Hastings recursion with Rao-Blackwellized Particle Filter

Input : Prior distribution $p(\vartheta)$, Initial parameter value ϑ_{k-1}, likelihood estimate $\widetilde{\mathcal{L}}_{k-1}$ and particle $\widetilde{u}_{1:N}^{(k-1)}$, Proposal covariance matrix Q^{prop}

Output: New parameter value ϑ_k, likelihood estimate $\widetilde{\mathcal{L}}_k$ and particle $\widetilde{u}_{1:N}^{k}$

$\vartheta^{\text{prop}} \sim \mathcal{N}(\vartheta_{k-1}, Q^{\text{prop}})$;

$\widetilde{\mathcal{L}}_{\text{prop}}, \left\{\mu_k^{(i)}, P_k^{(i)}, u_{1:N}^{(i)}\right\}_{i=1}^{N_p}, \left\{w_N^{(i)}\right\}_{i=1}^{N_p} = \text{RBPF}(\vartheta_{\text{prop}}, z_{1:N})$;

Sample $\widetilde{u}_{\text{prop}} \sim \left\{u_{1:N}^{(i)}\right\}_{i=1}^{N_p}$ with probabilities $\left\{w_N^{(i)}\right\}_{i=1}^{N_p}$;

$\widetilde{\alpha} = \frac{\mathcal{L}_{\text{prop}} p(\vartheta_{\text{prop}})}{\mathcal{L}_{k-1} p(\vartheta_{k-1})}$;

$u \sim \mathcal{U}([0,1])$;

if $u < \alpha$ **then**

$\vartheta_k = \vartheta_{\text{prop}}$;

$\widetilde{\mathcal{L}}_k = \widetilde{\mathcal{L}}_{\text{prop}}$;

$\widetilde{u}_{1:N}^{(k)} = \widetilde{u}_{\text{prop}}$;

else

$\vartheta_k = \vartheta_{k-1}$;

$\widetilde{\mathcal{L}}_k = \widetilde{\mathcal{L}}_{k-1}$;

$\widetilde{u}_{1:N}^{(k)} = \widetilde{u}_{1:N}^{(k-1)}$;

4.3.4 *Noisy Metropolis-Hastings sampling*

The issue of convergence of the PMMH sampler, which highly relies on the accuracy of the estimator of the data likelihood, is still present with the RBPF, even if it allows variance reduction. The MCMC chain can still be stuck for several iterations, and only a complete analysis of the chain could assess this sampling issue [FS11]. To overcome such problems, modified versions of the MH sampler has been developed, called *noisy* MH. The Monte Carlo Within Metropolis (MCWM) sampler aims at escaping from blocking situations by forgetting the previous calculated value of the likelihood [Bea03; AR09]. To derive the acceptance ratio, it is therefore necessary to compute likelihood for both proposed parameter vector and previous one. Of course, this forgetting property removes some interesting theoretical advantages of the MH sampler, and does even not guaranty the convergence of the MCWM

sampler. However, for practical simulation examples, it appeared to perform well, producing only a higher variance into the obtained posterior distributions [McK+14].

Our approach of the non-linear sequential estimation problem is to use the methodologies presented in chapter 3 to find an optimal parameter vector, and use it to initiate a MCWM sampler, which aims at sampling from the posterior distribution $p(\vartheta, u_{1:N})$, similarly to the PMMH sampler. The full method is presented in algorithm 5. We called it the Noisy PMMH sampler (NPMMH), used with a Rao-Blackwellized particle filter. Main benefit is thus to preserve the exploration of the parameter space, and escape from possible chain blockages. However, performing new runs of the filter at each iteration will also improve the diversity of sampled particles, since no repetition of the same one (in case of no acceptance of the proposed parameter vector) is expected.

Algorithm 5: Noisy Particle Marginal Metropolis-Hastings recursion with Rao-Blackwellized Particle Filter

Input : Prior distribution $p(\vartheta)$, Initial parameter value ϑ_{k-1}, Proposal covariance matrix Q^{prop}

Output: New parameter value ϑ_k, likelihood estimate $\widetilde{\mathcal{L}}_k$ and particle $\widetilde{u}^k_{1:N}$

$\vartheta^{\text{prop}} \sim \mathcal{N}(\vartheta_{k-1}, Q^{\text{prop}})$;

$\widetilde{\mathcal{L}}_{\text{old}}, \left\{\mu_k^{(i)}, P_k^{(i)}, u_{1:N}^{(i)}\right\}_{i=1}^{N_p}, \left\{w_N^{(i)}\right\}_{i=1}^{N_p} = \text{RBPF}(\vartheta_{k-1}, z_{1:N})$;

Sample $\widetilde{u}_{\text{old}} \sim \left\{u_{1:N}^{(i)}\right\}_{i=1}^{N_p}$ with probabilities $\left\{w_N^{(i)}\right\}_{i=1}^{N_p}$;

$\widetilde{\mathcal{L}}_{\text{prop}}, \left\{\mu_k^{(i)}, P_k^{(i)}, u_{1:N}^{(i)}\right\}_{i=1}^{N_p}, \left\{w_N^{(i)}\right\}_{i=1}^{N_p} = \text{RBPF}(\vartheta_{\text{prop}}, z_{1:N})$;

Sample $\widetilde{u}_{\text{prop}} \sim \left\{u_{1:N}^{(i)}\right\}_{i=1}^{N_p}$ with probabilities $\left\{w_N^{(i)}\right\}_{i=1}^{N_p}$;

$\widetilde{\alpha} = \frac{\mathcal{L}_{\text{prop}}\, p(\vartheta_{\text{prop}})}{\mathcal{L}_{\text{old}}\, p(\vartheta_{\text{old}})}$;

$u \sim \mathcal{U}([0,1])$;

if $u < \alpha$ **then**

  $\vartheta_k = \vartheta_{\text{prop}}$;

  $\widetilde{\mathcal{L}}_k = \widetilde{\mathcal{L}}_{\text{prop}}$;

  $\widetilde{u}^{(k)}_{1:N} = \widetilde{u}_{\text{prop}}$;

else

  $\vartheta_k = \vartheta_{\text{old}}$;

  $\widetilde{\mathcal{L}}_k = \widetilde{\mathcal{L}}_{\text{old}}$;

  $\widetilde{u}^{(k)}_{1:N} = \widetilde{u}_{\text{old}}$;

4.3.5 *Smoother with Rao-Blackwellized filter*

Paradoxically, if the Rao-Blackwellization eases computation of the filtering distribution, it does not allow a direct implementation of the smoothing part. The Forward Backward Rao-Blackwellized Particle Smoother (or RBPF) is a stochastic algorithm, based on a run of a PF [SBG12]. Conditionally to the full particle states at each state

with their weights, it provides one full particle path. This filter is an adaptation of the FFBSi for RBPF. Since it provides also a smoothing distribution, it will be a reference method for comparison with our NPMMH algorithm.

4.3.6 *Demonstration of the recursion efficiency*

To assess performances of the previous presented methodology, an academic example is simulated, with the SDOF equation and constant dynamic parameters over the simulation. State vector is composed of angular position and its first derivative ($x_k = [\alpha_k, \dot{\alpha}_k]^T$). Damping ratio D is assumed equal to 0, and Bayesian inference is performed for the frequency parameter f. A signal is simulated with the RK4 process on a time interval of $0.1\,s$, with a time sampling interval of $\Delta t = 0.1\,ms$. A frequency of $50\,Hz$ is used for the simulation, and the state vector is simulated with a noise covariance matrix Q_x defined as:

$$Q_x = q \begin{bmatrix} \Delta t^2/3 & \Delta t/2 \\ \Delta t/2 & 1 \end{bmatrix} \tag{4.3.15}$$

with $q = 10^{-3}$. Measurement is the acceleration, defined as:

$$z_k = -C\,(2\pi f)^2\,\alpha_k \tag{4.3.16}$$

with $C = 0.1$ and Gaussian noise corruption of variance $\sigma^2 = 1$. Angle α and the measurements are shown on figure .

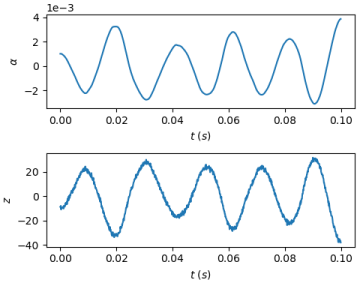

Figure 4.6: State and measurement simulated with the RK4 process.

A Kalman filter computes the exact likelihood for each new frequency value. A particle filter is then run to get an approximation of this likelihood. Of course, using a particle method in this linear situation is not necessary, but this particle example allows comparison with an exact basis. The MH algorithm uses the Kalman filter likelihood, and the PMMH the particle filter likelihood. The NPMMH is then compared with the results of this recursion.

This situation allows quantification of both accuracy of results and chain mixing. Preliminary MH and PMMH chains are first simulated to initiate the proposal covariance matrices and initial dynamic parameter. Five chains are then created to compute the PSRF score. A final chain for inference is then simulated to calculate posterior mean, compute IACT scores. For the NH simulation, no preliminary chain is necessary: initial frequency uses the global optimum resulting from the Bayesian optimization procedure. The latest is also used to compute the initial proposal covariance matrix. Five chains are also simulated for the estimation of the PSRF score.

Preliminary chains for the PSRF score calculation bring notable information. The five chains simulated with the PMMH sampling in figure 4.8 exhibit very poor convergence, compared with the two other samplers as shown in figures 4.7 and 4.9. Even in this low dimensional estimation problem, the particle filter gives high variance in the estimation of the likelihood. Figure 4.10 gathers the final simulated chain, and their related distribution over the parameter space. The distribution obtained from the PMMH is narrow, and the true value

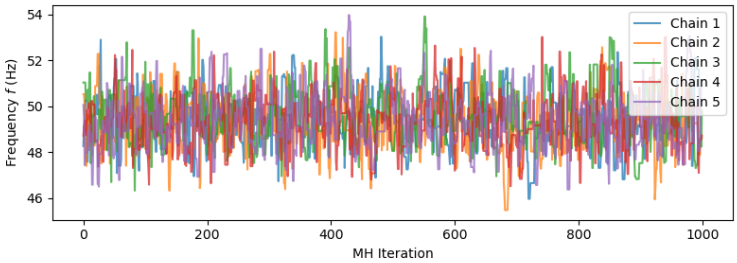

Figure 4.7: 5 chains used for the PSRF calculation of the MH sampler using Kalman filter.

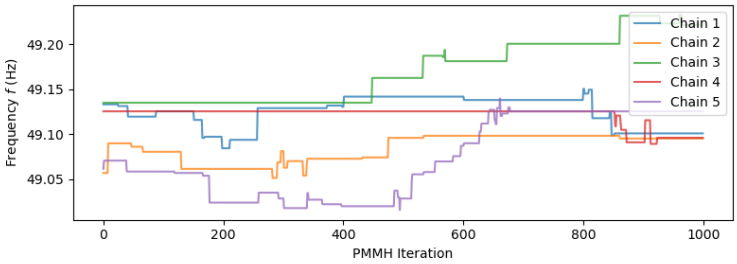

Figure 4.8: 5 chains used for the PSRF calculation of the Noisy PMMH sampler using particle filter.

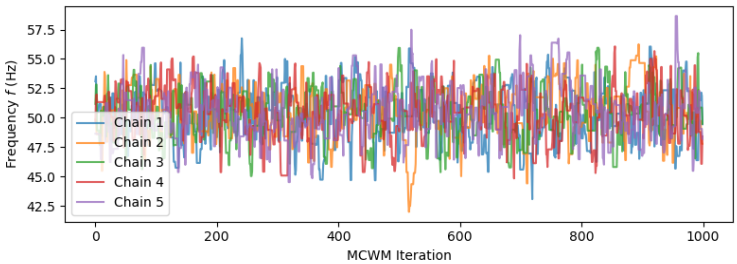

Figure 4.9: 5 chains used for the PSRF calculation of the PMMH sampler using particle filter.

of the frequency is even not included in the empirical distribution. The Noisy PMMH sampler produces distribution closed to the one obtained from the MH sampler. In addition, the mixing appears better. This visual analysis is comforted by the trend of the autocorrelation function 4.11. The Noisy PMMH exhibit the better forgetting properties, with the faster decrease of the autocorrelation.

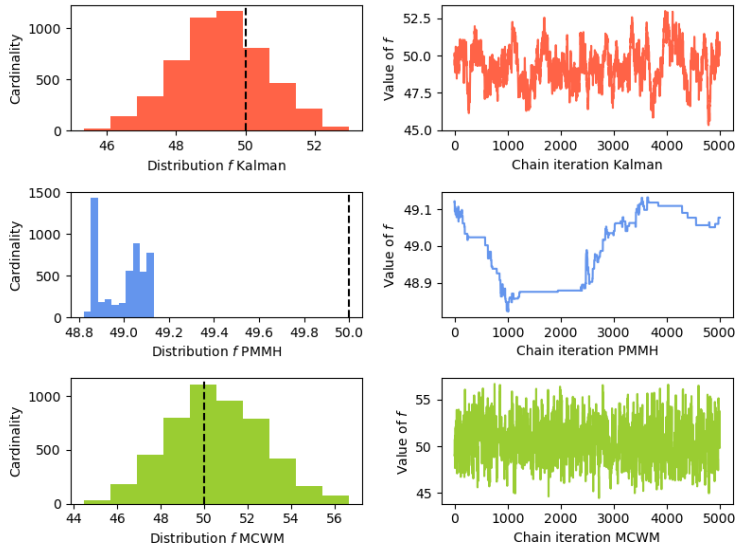

Figure 4.10: Final chains and distributions of frequency parameter f for the MH, PMMH and Noisy PMMH sampler.

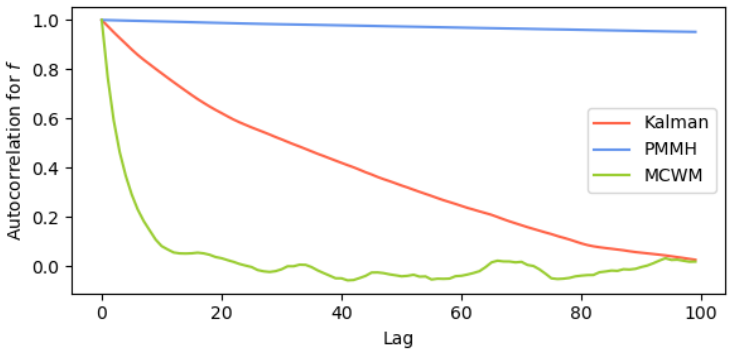

Figure 4.11: Final autocorrelations of frequency parameter f for the MH, PMMH and Noisy PMMH sampler.

Sampling method	PSRF (5 chains)	IACT	Posterior Mean	Posterior Variance
MH	1.007	76.4	49.3	1.64
PMMH	3.319	196	49.0	0.00890
Noisy PMMH	1.002	7.74	50.5	4.69

Table 4.1: Performance comparisons between MH, PMMH and NPMMH for sampling of frequency parameter f

Quantitative indicators, gathered in table 4.1 confirms these observations. However, the better mixing of the Noisy PMMH came at the cost of an increased variance. The Posterior variance calculated is indeed of the same order of magnitude, although slightly superior to the one from the MH sampler.

Of course, since each NPMMH iteration requires two evaluations of the likelihood, one for the previous parameter vector, and one for the proposed at current iteration, a prior remark would question the choice between using such algorithm, or running a traditional PMMH sampler with twice number of iteration. This is relevant in a context in which the total number of possible evaluation of the likelihood, and hence the number of particle filter runs, is limited. However, the previous results give no hesitation: even reduced by half, the NPMMH sampler is far more efficient in terms of mixing properties. Escaping a stationary phase, with locally 0 acceptance rate, is an issue hardly solved with use of a longer chain. Only lower variances could result in an improvement of the PMMH chain mixing.

Therefore, even if this short example may be solved with perhaps more performant methods, such as optimization based estimation schemes, it provides at least a limitation for the classical PMMH sampler. More complex problems, for instances higher dimensional state vectors, or non-linear systems, could results in even worse convergence issues. The proposed NPMMH methodology allows an operator to obtain a pragmatical estimation algorithm, which only requires little knowledge about the physical system, in the form of a differential equation locally valid, and which do not need special tuning. Only the number of particle should be defined, with wide estimates of the possible interval values of the model's parameters. The only cost of this convenient algorithm is a higher variance in the final posterior parameter distribution.

4.3.7 *Experiment and simulation results*

For non-constant dynamic parameter, even for a simple model, like the SDOF one, the unavailability of a close form solution makes necessary the approximation with sequential methods. For our problem, the RK4 algorithm provides an efficient resolution scheme. We model important changes of dynamic parameters: 40 % for D (from 0.05 to 0.03), and 80 % for f (from 50 Hz to 90 Hz). The attenuation is performed during 0.25 s, with a sampling interval equal to $\Delta t = 0.25\,ms$ (for a total of 1000 points). Dynamic parameters are set constant during the first 40 ms at the beginning and end of the signal, with linear variations between them, as shown is figure 4.12. Signal is corrupted with a white Gaussian noise of variance $\sigma^2 = 0.01$, producing a SNR of 28 dB.

State vector is set to $x_k = [\gamma_k, \dot{\gamma}_k]^T$, and measurement equation is :

$$z_k = L\left[-(2\pi f_k)^2\gamma_k - 2D_k(2\pi f_k)\dot{\gamma}_k\right] \tag{4.3.17}$$

Simulation is performed with inference made on the parameter vector $\left[\sigma^2, \sigma_D, \sigma_f, \sigma_{r^D}, \sigma_{r^f}\right]^T$, for a complete characterization of the measurement noise and dynamic parameter evolution behaviour. Parameter q is set to 10^{-9}. The Bayesian optimization procedure selects the optimal parameter vector, which is used then to estimate the filtering

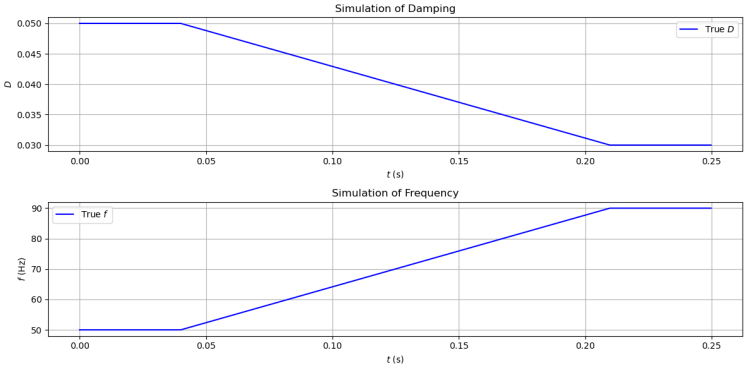

Figure 4.12: Variations of dynamic parameters for simulation purpose.

vector $x_{1:N}^{\mathrm{PF}}$ and the posterior mean vector $x_{1:N}^{\mathrm{post}}$, defined as:

$$x_{1:N}^{\mathrm{post}} = \sum_{i=1}^{N_p} w_N^{(i)} \, x_{1:N}^{(i)} \tag{4.3.18}$$

Particle filter used 1000 particles, and its runs was kept to generate 100 backward samplings with the Particle smoother. The filters used in the NPMMH recursions employed only 100 particles, but 2000 iterations of the chain were simulated, with start from the optimal vector obtained by the BO loop.

To compare our methodology with a non-stochastic method, we propose to use an approximation derived from the clean SDOF model (without noise). With notation $\alpha(t) = A(t)\cos(\phi(t))$, the analytic representation is $\alpha_a(t) = A(t)e^{i\phi(t)}$, under the constraint $A'(t) \ll \phi'(t)$. An evaluation procedure of damping and frequency from the signal (displacement only) has been proposed [Fel97] with introduction of a global damping variable $h(t)$. The SDOF equation becomes with this representation:

$$\ddot{\alpha_a}(t) + 2h(t)\dot{\alpha_a}(t) + \omega^2(t)\alpha_a(t) = 0 \tag{4.3.19}$$

Solving this equation with identification of real and imaginary parts gives (we omitted the time notation for A, ϕ and their derivatives):

$$h(t) = -\frac{\dot{A}}{A} - \frac{\ddot{\phi}}{2\dot{\phi}} \tag{4.3.20}$$

$$\omega_0^2(t) = \dot{\phi}^2 - \frac{\ddot{A}}{\dot{A}} + \frac{2\dot{A}^2}{A^2} + \frac{\dot{A}\ddot{\phi}}{A\dot{\phi}} \tag{4.3.21}$$

The instantaneous damping ratio is easily deduced as:

$$D(t) = \frac{h(t)}{\omega(t)} \tag{4.3.22}$$

The estimation of instantaneous amplitude and phase is required for calculations of damping and frequency. To provide clean vectors (with minimal level of noise), the CWT is first applied. Using wavelets based methods has been demonstrated to be performant for estimations of amplitude and phase in this context [Cur+08b]. Extraction of the ridge is made on the frequency bandwidth of interest (between $10\,Hz$

and 110 Hz). Since only the acceleration is accessible in the resonant-column experiments, a major hypothesis is made: we assumed the acceleration to be proportional to the displacement. Since non-linear behaviours are supposed inside the measurement signal, this methodology is expected to produce biased estimations. However, error will be computed in the simulation.

From the synthesized data, we compare them quantitatively with their estimated values. The *mean relative absolute error* (MRAE) is used here (in percent), defined here for each state x as:

$$\frac{1}{N}\sum_{k=1}^{N} 100 \left| \frac{\widehat{x}_k - x_k}{x_k} \right| \tag{4.3.23}$$

However, to equitably compare the performances of methods, the error must not be computed on the full-time interval. Indeed, the CWT method does not provide reliable estimation of amplitude and phase at the beginning and end of the signal, due to zero-padding. At the edges of the signal, wavelet coefficients calculated suffer from the discontinuity. The *cone of influence* (COI) is the region of the spectrum where this phenomenon can be neglected. The limit used in this article is the concept of Torrence [TC98]: for each scale s, the temporal limit of the COI is defined as the e-folding time τ^* such that the autocorrelation of the power spectrum decreases by a factor e^{-1} , for a temporal discontinuity at the signal edge. For a Morlet mother wavelet, τ^* can be analytically calculated, and is equal, for the same mother wavelet configuration as Torrence, to $\sqrt{2}s$.

To bypass this issue, we decided to estimate the MRAE inside the COI, only after the first two periods of the minimal frequency of the signal, and before the two last ones. The filtering estimations do not suffer from such limitations. The only information to be provided is the initial value of the states.

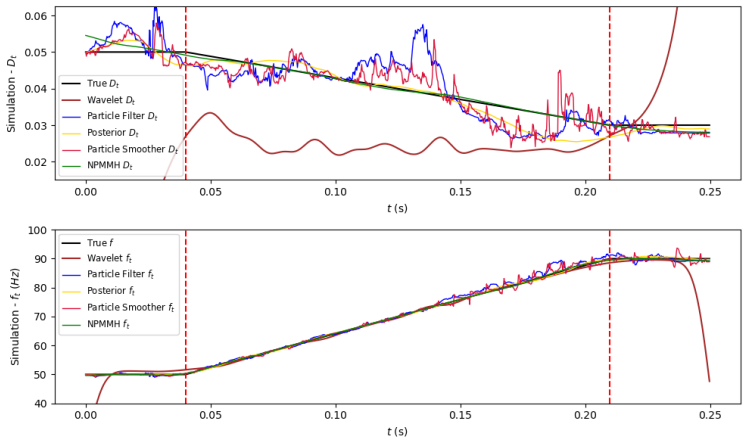

Figure 4.13: Estimations of dynamic parameters with simulated signals, for the particle filter estimate, the posterior estimate and the wavelet based method.

All estimated vectors are displayed in figure 4.13. The wavelet based method exhibits a higher bias, with highly degenerated estimations outside the COI. Filtering methods do not share this drawback.

Performances of the NPMMH are far superior to the ones obtained

Estimation scheme	MRAE for Damping D	MRAE for frequency f
Filtering mean	11.2	0.969
Posterior mean	6.50	0.328
CWT method	37.7	0.564
Particle Smoother	8.65	0.687
Noisy PMMH	0.691	0.198

Table 4.2: Mean relative absolute error comparisons (in percent) for estimations on simulated attenuation with time varying dynamic parameters

from all other methods. It achieved the best estimations for both damping and frequency. Qualitatively, estimated curves of dynamic parameter appear also smooth. Even if this criterion has little importance to achieve the best estimation performances, it ensures a clean extraction of points from the curves, if such operation is necessary. The simple filtering estimations (with PF) achieve largely better performances for damping, in comparison with the CWT methods, but slightly worst for frequency. Estimation with the posterior mean overtakes the two previous methods for both damping and frequency. The only burden of this methods is the necessity to store ancestors of particles, which is not a strong drawback for the benefit its brings. Surprisingly, PS performances are inferior to the ones of the posterior mean, although a clear improvement is observed in comparison with the RBPF. An advantage of the smoother is however the availability of the variances of the smoothed distribution, which adds a qualitative indicator of the estimation.

4.4 Resonant column experiment

4.4.1 *Experiment results*

Acceleration measurements from the resonant-column testing device exhibit a strong change in dynamic parameters at the very beginning of the SDOF part, just following the end of forced excitation by coils. The vibratory behaviour with attenuation of the signal next continues until the end of measurement. Therefore, a filtering starting right after the end of excitation would directly have to face strong non-linearities, resulting in poor estimations at these specific measurement points. Our selected strategy is then to start the estimation process at the very end of the measurement signal, where the attenuated signal has already reached a stationary regime. Changes of dynamic parameters will slowly happen, until the complex or brutal variations occurred. To get such time-inversion phenomena, a variable change $t' = -t$ is performed, transforming the SDOF equation 4.2.3 into:

Such variable change only impacts the odd derivatives, adding a multiplication by -1.

$$\ddot{\alpha}(t') - 2D\omega\dot{\alpha}(t') + \omega\alpha(t') = 0 \tag{4.4.1}$$

The transition matrices is then modified to include this change. Notably, it will only impact the terms in which the damping appears. As in the previous example, the Bayesian optimization provides an estimation of the optimal parameter. However, the main difference with this practical signal is the need to define the interval of estimation.

The end of forced excitation can be approximated with the decay of the input signal. Fitting a sinusoid with amplification with time at the end of the measurement signal is a viable option (amplification with reversed time). Increase of the mean square error between the signal and the fitted curve indicates pragmatically the change of regime of the dynamic parameters, from constant regime to time-varying evolution.

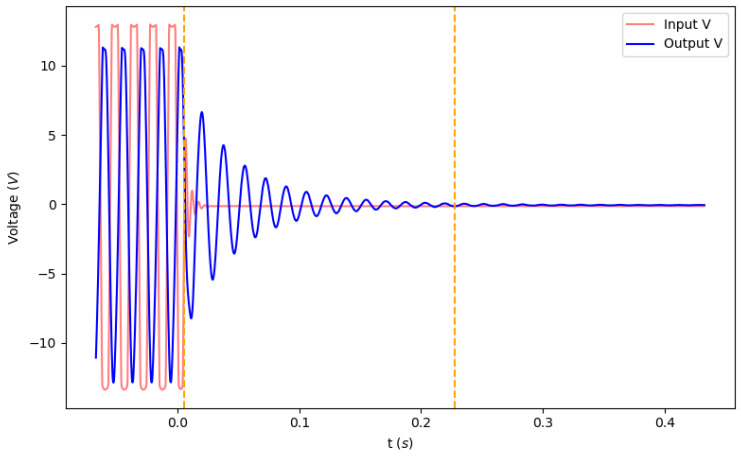

Figure 4.14: High pressure ($60\,kPa$) confinement with high voltage excitation of a soil specimen.

An experimental program was conducted on a soil sand specimen made of Ottawa sand. The non-linear dynamic behaviour was studied using different excitation levels. The experiments were performed at the non-destructive testing laboratory at the department of Civil and Environmental Engineering at the University of Waterloo. A series of measurements is therefore obtained. The transmitter used in this test protrudes $6\,mm$ into the soil specimen and has a $14\,mm$ by $1.0\,mm$ cross-section. The soil specimen has diameter of $7\,cm$ and height of $14\,cm$. Dry sand has density of $1.85\,kg/m^3$. The specimen is compacted and covered with a latex membrane to contain the sand. A $60\,kPa$ vacuum pressure is then applied at the bottom of the specimen. Response is measured by two accelerometers, positioned at a distance from the resonant-column centre equals to $6.8\,cm$. This series allows us to perform for each signal a least square fitting with curve model 4.2.10, to obtain one dynamic parameter vector, which is latter associated to the maximum input excitation level (and therefore the maximum strain of the signal).

From the obtained series, the one with the largest excitation level is kept to perform estimations. This ensure that a wide range of strain will be found in the signal. For our inspected signal, measurements with related excitation voltage are represented in figure 4.14, with borders of estimation. Measurements are thus in volts (V), and factor L equals $0.128\,Vs^2$. Parameter q is set to $1e-9$. Transition matrix used is the one of equation 4.2.26.

Estimated states for the RBPF with their $\pm\sigma$ uncertainty limits are displayed in figure 4.15. A strong narrowing of this uncertainty band is observable at the beginning of the record, which corresponds to the final estimation. At these points, variations of dynamic parameters

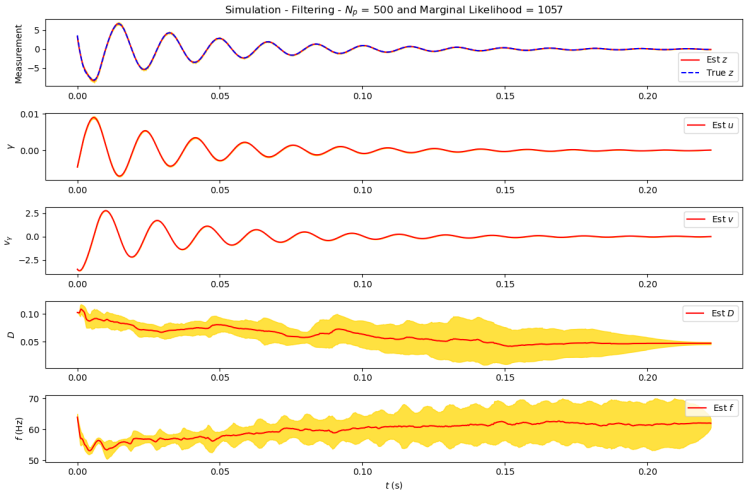

Figure 4.15: Particle filter estimations for high pressure ($60\,kPa$) confinement signal with high voltage excitation of a soil specimen. Interval of estimation is located between the orange vertical lines.

are important, reducing the possible admissible values. The advanced results coming from smoothing distribution are shown in figure 4.16. All methods exhibit similar trends and behaviours, but the NPMMH estimations still provide the smoothest curves.

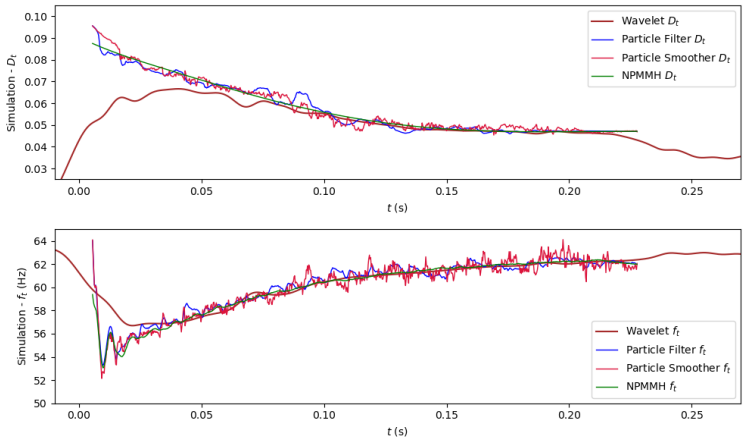

Figure 4.16: Estimations with all methods for high pressure ($60\,kPa$) confinement signal with high voltage excitation of a soil specimen.

A focus on the NPMMH results (figure 4.17) highlights the inverse variations of damping and frequency in real measurements. This phenomenon, generalization of the degradation curves, is therefore empirically demonstrated with high strain levels. The sequential filters and smoothers do not assume a prior trend or monotony. However, the estimations obtained during the previous experiments recover the variations of hyperbolic models, suggesting the validity of the hypothesis.

Final representation of the signal with its dynamic parameter variables highlight the phenomenon demonstrated with several signals: decrease of the damping with brutal increasing frequency, until a slow stabilization of their values.

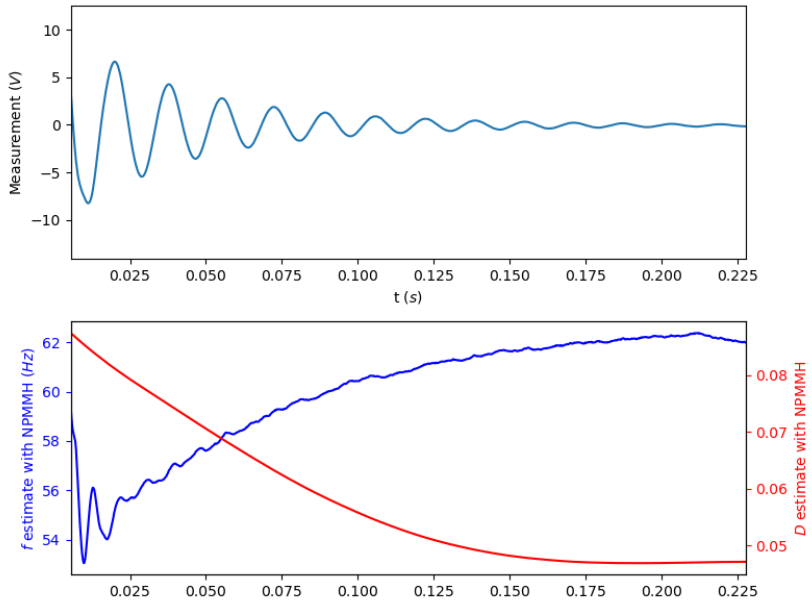

Figure 4.17: Dynamic parameter evolution for high pressure ($60\,kPa$) confinement with high voltage excitation of a soil specimen.

4.4.2 *Discussion on degradation curves*

We suggest here to use the information about damping and frequency at every local maximum or minimum of the strain, estimated during the filtering process. To compute the time of local extrema, we directly use the estimated velocity. When it crosses zeros, it should correspond to an extremum of the displacement. The related damping and frequency values are next reported.

Hyperbolic curves have been fitted to datasets obtained from the series of experiments (every measurement gives a strain and a dynamic parameter vector), from the PS estimations and from the NPMMH estimations. Parametric models used for curve fitting are hyperbolic ones. Curves and points are gathered in figure 4.18. However, after curve fitting, we observed an additional damping, due to some physical elements of the measurement chain. A rescaling of the curves is performed to have a minimum value of damping at 0.02.

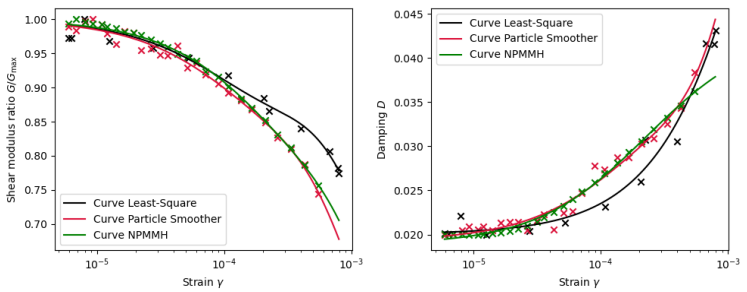

Figure 4.18: Dynamic parameter evolution for high pressure ($60\,kPa$) confinement with high voltage excitation of a soil specimen.

Trends between the experimental curves and the estimations from

smoothers are similar. However, for the shear strain modulus, a small gap is noticeable for high strains. Error with damping can be explained by the need for the least-square method to find a compromise between error at the beginning and error at the end of the signal, after attenuation. One single value has to be selected, without possible variations of damping. Therefore, underestimation of the damping is expected. One further explanation could be a high non-linear behaviour of the soil sample during high excitations. Consequently, estimations with high strains could not be compared with low strain measurements. This hypothesis requires further investigations.

Existing methods for resonant column results analysis gather the maximum strains of multiple experiments. Relationships between damping or frequency and strain are assumed to keep their validity inside one measured signal. Indeed, for high excitation signals, many resources could be saved if degradations curves are obtained directly from a few measurements.

4.5 Conclusion

This chapter introduced a probabilistic methodology for feature extraction of hidden information form accelerometer measurements. First, a state space representation is needed to allow modelling of time varying characteristics. A local available knowledge about the physical process, here the SDOF model equation, is the only requirement. A transition model for dynamic parameters of interest is created from weak assumptions about their trends. The Rao-Blackwellized particle filter is the core method to provide an estimate about the hidden information. The proposed Bayesian methodology, based on a noisy MCMC chain, allows an integration of model parameter, to marginalize results over possible choices available. Efficiency is achieved by first looking for an optimal choice of model parameters, and secondly running a noisy sampling chain. No specific tuning of the model is therefore needed, at the cost of higher time of processing. Estimation of hidden variable form acceleration is provided, with improved performances with comparisons with time-frequency analysis. Experiments conducted on a resonant-column and analysed with this methodology identify new behaviours of a soil sample subject to external perturbations.

Previous methods presented are transferable to inspection of other mediums. With the same physical equations, analysis of railways sleepers could be performed, adding new perspectives and information to frequency [Hoa17] or finite element [Cla+19] analyses, with probabilistic quantification on estimations.

Conclusion

Main contributions

This thesis addressed the issue of characterization of defects from ultrasonic sensors, based on the statistical signal processing framework. A focus has been made on applications to rail crack detection. Majority of our data came from a non-contact technique, newly applied to rail inspection by RAILENIUM, the Electro-Magnetic Acoustic Transducers (EMAT). The sensors take advantage of the properties of a specific kind of wave they generate, Rayleigh waves. They have the ability to propagate on surface, and interact with damages. Objective was to investigate the possibility to perform assessment of a rail health with mobile sensors.

My first contribution was to gather two major schemes: statistical estimation, with inclusion of powerful decompositions methods of measurement signals, and the data association field. The Matching Pursuit (MP) algorithm offers an abstract representation of ultrasonic data, which value has been recognized, either for denoising purposes or parametric description of phenomena. However, current aspiration of the railway industry for mobile inspection at high speed induces a new dimension in the representation of data: the sequential one. My work has build a framework to include evolution of ultrasonic events across measurements, thanks to the help of Random Finite Sets (RFS). Models were developed in the thesis to understand, characterize and follow ultrasonic pulses. RFS filtering and tracking tools integrate the uncertainty related to measurements, and succeeded in assimilating noised and capricious data, which is inherent to field inspection. The algorithms developed for this purpose share a fundamental property: their ability to estimate complex relationships between measurements sequentially, offering promises of online inspection technologies. All the algorithms used in this thesis have been programmed during the thesis in Python.

My second contribution incorporates a practical consideration about the latter filtering and tracking methodologies. The thesis proposes a handy solution to tune automatically them, preventing a human operator from spending huge amount of time in the comprehension of the filters, and manual selection of the parameters governing them. The proposed methodology relies on simplification assumptions of the initial δ-GLMB tracker, a complex and robust algorithm for creation of tracks and pattern from measurements. To achieve en efficient tuning, maximization of the data likelihood is achieved with a Poisson approximation of the multi-target likelihood, allowing use of a simpler algo-

rithm, the Probability Hypothesis Density (PHD) filter. A Bayesian procedure next assimilates information coming form this simplification to produce an automatic and efficient search of optimal parameters. This approach has been successfully tested on laboratory, with experimental rails portions, before a field application on the rails of Eurotunnel. These field experiments have been challenging tests, to discover limitations, but also opportunities brought by the methodology developed in this thesis.

My last contribution investigated a theoretical field. A problem coming from geotechnical sciences, the characterization of soil behaviours, opened the door of Bayesian filtering of measurements. Advanced methodologies for estimation of complex distributions, such as smoothing and Monte Carlo Markov Chain (MCMC) algorithms offered new perspectives to this specific field. This thesis created new processes to analyse soils and statistically capture the uncertainty of related information. Experiments on a specific device, the resonant-column, allowed an original application of our procedures, and led to new insights of the soil inspection problem.

Work summary

Important results ensuing from this thesis work are here gathered. The major one is the demonstration of the possibility to process a flow of ultrasonic measurements with sequential filtering techniques. Sparse decompositions of signals, joint to filters, allow extraction of statistical characteristics to assess a medium health. In our rail application, identification with EMAT sensors of the nature of ultrasonic pulses inside a signal conducted to separation of waves reflected by defects. This not only confirms the detection of cracks on the rail surface, but also brings new possibilities to characterize location of defects. Laboratory experiments with machined rails demonstrated the interest of RFS filters or trackers, to improve the detection of defects (head-checking and squat), in static and mobile inspection scenarios.

Field experiments conducted at Eurotunnel, with different intensities of head-checking, confirms the ability of the proposed methodology to identify changes in the estimated features. An information about the availability of estimations thanks to the RFS model allowed identification of troublesome measurements, offering an indicator for further analysis of the related rail area by a human operator. However, our estimations procedures highlighted the compulsory need for a proper design of the measurement device. Uncertainty about estimated features could be reduced, for instance, with inclusion of external information, like the distance between sensors and rail.

Our practical solution to solve the hard problem of model parameter estimation using RFS filters and tracker was successfully applied to our ultrasonic data. Several simulations concluded to the validity and

efficiency of the approach, and application of EMAT signals on our laboratory rails brought enough performances for interpretation of results. The similar optimization procedure was repeated with classical sequential filters for vector estimations, and prevent us from manual tuning of parameters.

The last relevant results is related to general analyse of vibratory signals. Estimation of time-varying parameters with filtering methods conducted to development of practical Bayesian estimation procedures. Experimentations will highly non-linear signals resulting from soil testing allowed the calculation of new characteristics in the domain of geotechnical sciences. Again, the methodology implied an automatic integration of parameter uncertainty in the model. The presented algorithm, contrary to previous applications with rail inspection, sacrificed the time efficiency for precision performances.

Perspectives

The first wish arising from my work is to perform high speed inspection of rails. Issues coming from the increase of speed have been identified, and the statistical methodologies previously presented offers solutions and quantification of loss of measurement quality.

The previous techniques focused on Rayleigh wave propagation. Inspection of internal defects could be possible, with use of new ultrasonic sensors, such as Phased Array transducers, which provide a complete scan of medium (see figure 5.19).

Ferromagnetic materials have been inspected, but ultrasonic devices can perform analysis and inspection of a wide range of mediums. Particularly, concrete structures could benefit from the joint use of sparse decompositions and tracking algorithms. Indeed, these environments are dispersive and prone to generate very noisy signals.

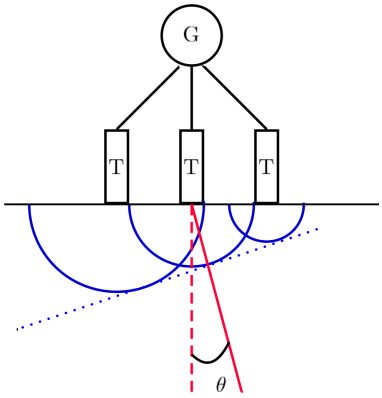

Figure 5.19: Phased array sensors, which produce a wavefront which inclination θ is controlled by a generator G, which supplies all transducers T, allowing a change of the inspected direction during measurements

Résumé en français

La pression croissante portée sur l'exploitation des infrastructures ferroviaires a rendu nécessaire le développement de nouvelles techniques d'inspection, plus rapide et plus sûre. Réduire le temps passer à déterminer l'état et la santé des rails permettra l'élaboration de plans de maintenance plus efficaces. Dans ce cadre, l'inspections dynamique des rails est nécessaire: les outils de mesure sont embarqués sur des trains spécifiques, ou des dispositifs roulant expérimentaux, effectuant des enregistrement à vitesse élevée. Des séquences de mesures sont ainsi obtenues. Plutôt que d'analyser chacun des enregistrements séparément, cette thèse se proposer d'étudier la séquence en elle même, permettant ainsi l'extraction d'informations relative à l'évolution des signaux. Plusieurs caractéristiques sont extraites grâce à des outils temps-fréquence. Leur association au travers des différentes acquisitions peut être représentée avec des modèles d'états, permettant de synthétiser l'information disponible en utilisant des densités de probabilité. Cette thèse propose une modélisation des phénomènes physiques, ainsi qu'une application sur des données de laboratoires et expérimentales.

Inspection sans contact des surfaces de rails

La présente thèse s'intéresse tout particulièrement aux défaut de surface apparaissant sur le champignon du rail, prêt de la zone de contact entre le rail et la roue. Ces défauts apparaissent graduellement avec le temps, suite aux contraintes élevées répétées imposées au matériau. Des modifications de la structure du rail sont mis en jeu, provoquant l'apparition et le développement de fissures. Les Squats sont reconnaissables par des taches sombres sur la zone de contact rail-roue. Les Head-check se développent à côté de cette zone. Ces fissures qui se propagent vers l'intérieur du champignon du rail. Les techniques de contrôle non-destructif usuelles, par capteur ultrasons piézoélectriques ou capteur par courants de Foucault ne permettent pas une bonne détection de ces défauts. Afin de les étudier, cette thèse a employé des capteurs EMAT (Electro Magnetic Accoustic Transducers). Ceux-ci possèdent la particularité de générer une onde ultra-sonore dans un matériau ferromagnétique en utilisant des dispositifs magnétiques. Aucun contact direct n'est donc nécessaire entre les capteurs et le milieu à inspecter, ni aucun couplant. De plus, les EMAT permettent de générer des ondes dites de Rayleigh, qui possèdent la propriété de se propager uniquement à la surface du milieu inspecté. Ce type d'ondes est donc intéressant pour l'études des défauts de surface, puisque leur interactions avec les ondes de Rayleigh va pouvoir être enregistrées.

Une représentation d'un signal ultra-sonore est appelé A-scan. Il traduit le déplacement du matériau au cours du temps, selon un axe spatial déterminé. L'agrégation de différents A-scan (si les capteurs bougent par exemple) forme une image de la zone inspectée appelée B-scan. Celui-ci permet une inspection visuelle de la zone mesurée. Pour un milieu comportant des bordure, on peut constater l'apparition de raies diagonales traduisant les réflexions de l'onde générée à ces endroits: réflexions à l'avant du capteur, et à l'arrière du capteur. L'onde transmise directement (traversant le milieu de l'émetteur au récepteur) donne des raies ou parcours droits sur le B-scan. Une séparation de ce B-scan permettrait donc d'identifier individuellement les parcours des ondes et leur nature (onde directe ou réflexions ou échos).

Dans le cadre des campagnes de mesures réalisées pour la thèse, deux échantillons de rails avec défauts usinées ont été inspectés. Ces mesures en laboratoires permettent de fournir des références pour les algorithmes de traitement à appliquer sur les mesures provenant des mesures EMAT. Des tournées d'inspection sur les rails en exploitations d'Eurotunnel ont ensuite été permises grâce à un chariot de mesure réalisé par RAILENIUM.

Traitement du signal statistique des signaux ultra-sonores

L'analyse temps-fréquence grâce à la transformée en ondelettes continue permet l'estimation de certaines caractéristiques, notamment de la fréquence de l'impulsion et du temps de vol. Néanmoins, dans le cas d'un signal composé de plusieurs impulsion, cette décomposition ne permet pas d'obtenir des informations multiples. La décomposition parcimonieuse d'un signal vise à trouver une représentation de ce signal dans une certaine base, appelée dictionnaire, composé d'atomes. Pour garder une propriété de parcimonie, le signal doit pouvoir être représenté comme une somme d'un petit nombre de ces atomes. Le dictionnaire est dit «surcomplet» si sa dimension est largement supérieure à celle du signal. Ceci permet d'obtenir la propriété de parcimonie. Les signaux ultra-sonores peuvent être vu comme une somme finie d'ondelettes de Gabor, composées d'une sinusoïde multiplié par une enveloppe gaussienne, et paramétrées par un vecteur θ composé de 5 scalaires: amplitude A, fréquence centrale ν_c, étalement ς^2, phase ϕ et temps de vol u (ondelette associée notée f_θ):

$$f_\theta(t) = A \exp\left(-\frac{(t-u)^2}{s}\right) \cos\left(2\pi\nu_c(t-u)+\phi\right)$$

Les signaux ultra-sonores sont donc parcimonieux dans le dictionnaires des ondelettes de Gabor. L'étape de décomposition d'un signal est un problème NP difficile. Néanmoins, une méthode de décomposition séquentielle, appelée Matching Pursuit (MP), permet d'obtenir une séquence d'ensembles d'ondelettes, chacun ce ces ensembles s'approchant du signal d'origine. Une modélisation probabiliste offre une amélioration des performances, grâce à l'emploi de l'algorithme Expectation-Maximization (EM). Cette méthodologie est employé sur nos signaux EMAT, et permet d'obtenir pour chaque acquisition un nombre restreint de vecteur de paramètres.

L'étude des variations de cette représentation permettra d'extraire des informations additionnelles des signaux EMAT, et de caractériser les défauts mesurés. L'étude séquentielles des observations et l'association des mesures est possible grâce à une représentation probabiliste par vecteurs d'états. L'information est résumé par une densité de probabilité, qui va évoluer au cours du temps. Les filtres séquentiels permettent de transmettre et propager ces densité de probabilité en les mettant à jours avec de nouvelles mesures. Les caractéristiques intéressantes sont ensuite extraites de ces densités. Dans le cas d'une observation vectorielle, les filtres de la littératures, tels que le filtre de Kalman, ou le filtre particulaire, permettent d'estimer les densités de probabilité à chaque nouvelle mesure.

Dans notre cas, chaque mesure est en réalité un ensemble Z, de cardinalité variable et inconnue. La théorie des «Finite Set Statistics» développée par Mahler a permis le développement d'une procédure de filtrage, par le filtre «Probability Hypothesis Density» (PHD), en travaillant sur des ensembles aléatoires finis (EAF). Ce filtre fournit une estimation séquentielle de la cardinalité et une estimation de l'ensemble d'état caché X, composé d'estimations des vecteurs d'états le composant. L'état des cibles est ainsi estimé au cours du temps. Néanmoins, l'association entre les états estimés des cibles n'est pas réalisé entre les instants. Un traqueur permet d'effectuer cette association. Le filtre «Delta Generalized Labeled Multi-Bernoulli» (δ-GLMB) permet de réaliser cette action, dans un cadre Bayésien. Il propage une densité de probabilité notée $\pi(\mathbf{X})$ sur des ensembles finis aléatoires labellisés ($\mathbf{X} \in \mathcal{X} \times \mathbb{L}$, avec $\mathcal{X}$ l'ensemble des états, $\mathbb{L}$ l'ensembles des labels), qui permet ainsi d'identifier les cibles au cours du temps. Ces outils statistiques vont permettre de suivre l'évolution des ondelettes à travers les acquisitions, en évaluant leur nombre et leurs caractéristiques.

Caractérisation des signaux ultra-sonores sur rails

On modélise l'évolution des paramètre des ondelettes de Gabor par des modèles dynamiques du premier ordre, et par des marches aléatoires gaussiennes. Ainsi, chaque paramètre d'une ondelettes peut varier sensiblement avec les temps, soit à cause des perturbations du dispositif de mesure, ou soit à cause d'une modification des propriétés du matériaux. On introduit des variables supplémentaires, nommés «tempos»,

pour modéliser des variations linéaires de certains paramètres, respectivement de l'amplitude et du temps de vol des impulsions (tempos associées notés p_A et p_u). Un problème conséquent lors de l'utilisation de filtres est le choix des paramètres du model et ceux régissant les filtres. Dans le cas des filtres sur ensembles finis aléatoires, le vecteur de paramètres peut contenir les paramètres de naissance et survie des cibles, ainsi que les niveaux de bruit de la mesure. L'utilisateur doit effectuer un choix de ces paramètres avant d'effecteur la procédure d'estimation, en se basant sur sa propre expérience de la mesure. Or, lorsque de nouvelles mesures apparaissent, ou lorsque les observations sont produites par des systèmes de mesure innovants (comme dans le cas des EMAT), aucun choix a priori n'est disponible. Une détermination du vecteur de paramètres à partir des mesures est néanmoins possible. Il s'agit dans ce cas de trouver le vecteur ϑ qui maximise la vraisemblance des donnée $\mathcal{L}_\vartheta(Z) = p(Z \mid \vartheta)$.

Lorque la méthode l'estimation est complexe, fournir une valeur de la vraisemblance peut être couteux, de même pour une estimation de son gradient. La maximisation est ainsi difficile. Pour résoudre ce problème, l'Optimisation Bayésienne permet de trouver un optimum global en utilisant des outils statistiques. La fonction à maximiser est approchée par une fonction probabiliste, ici par une Régression par Processus Gaussien. On cherche de manière séquentielle à maximiser un équilibre entre la valeur possible donnée par la fonction probabiliste et l'incertitude présente. La fonction probabiliste est ensuite mise à jour avec la nouvelle valeur de la vraisemblance. L'intérêt ce cette méthode est de pouvoir ainsi maximiser la vraisemblance avec un nombre limité d'évaluation, sans calcul de gradient, et en incluant la possibilité que le calcul de la vraisemblance puisse être stochastique.

Pour l'utilisation des filtres sur EAF, un calcul de la vraisemblance est possible pour le filtre PHD. Néanmoins, pour le filtre δ-GLMB, lorsque le nombre de cible ou de mesure est conséquent, des problèmes combinatoires rendent le calcul difficile. La thèse propose donc d'approximer la vraisemblance grâce au modèle de Poisson, en la calculant avec les équations de récursion du filtre PHD. L'Optimisation Bayésienne permet d'obtenir un vecteur de paramètre optimal, ensuite utilisé par le filtre δ-GLMB. Les algorithmes de filtrages ont d'abord été testés sur des mesures de laboratoires. Un essai sur un cylindre d'acier a permis de montrer les possibilités d'extraction d'ondes de réflexions de la technique présentée précédemment, ainsi que sa capacité à estimer la vitesse de propagation des ondes. Sur les rails de laboratoires avec défauts usinés, la technique permet la décomposition du B-scan. Différents parcours d'ondes sont ainsi extraits: un B-scan pour les ondes directement transmises, un pour les échos arrières, et un pour les échos avants. Les paramètres des filtres sont estimés pour chaque nouvel ensemble d'acquisitions.

Pour éprouver les performances des algorithmes développés sur des données terrain, un chariot de mesure dynamique avec capteurs EMAT a été construit par RAILENIUM. Il a été employé pour réaliser des mesures sur les voies d'Eurotunnel, de manière dynamique: les mesures sont réalisées lorsque le chariot se déplace à vitesse constante. Deux récepteurs EMAT sont montés sur le chariot. Ceci permet de réaliser des comparaisons entre les signaux enregistrés, pour une même onde générée. Un retard pour le capteur le plus éloigné, ainsi qu'une atténuation, par rapport au capteur plus proche de l'émetteur. On calcule ainsi une différence des temps de vol Δu, ainsi qu'un ratio des amplitudes r_A. Des variations de ces indicateurs indique une modification de la structure du matériau, et ainsi la présence de défauts. Un calcul direct de ces indicateurs peut être sensible aux bruit de mesure. De plus, lorsqu'un capteur est mis en défaut, ces ratios perdent leur signification. L'utilisation des EAF permet d'obtenir une quantification de l'incertitude relative à la mesure. Lorsqu'un parcours d'une ondelette n'est pas disponible, le calcul des indicateurs n'est pas réalisé. Ceci permet de montrer directement la mise en défaut du processus, plutôt que le calcul d'indicateurs aux valeurs anormales.

Dans le cas des essais à Eurotunnel, plusieurs zones sont inspectées: une avec Head-Check léger, une avec Head-Check modéré, ainsi qu'une avec Head-Check important. Les indicateurs calculés avec les filtres EAF permettent d'identifier clairement ces zones, et ainsi attester la présence de défauts.

Caractérisation local de signaux pour la détection de défauts

L'étude des phénomènes vibratoires peut être généralisée à d'autres signaux obtenu à partir de mesures sur des infrastructures ferroviaires. Les techniques précédemment décrites se sont focalisées sur les données ultrasonores. Mais le formalisme employé, notamment l'utilisation des modèles d'états et des filtres probabilistes, peut être transposé aux signaux de fréquence plus basse. En particulier, des mesures provenant d'accéléromètres correspondent à ce critère. Par exemple, des mesures de vibrations des blochets pourraient également être analysées par ces moyens. Une comparaison ultérieure de nos méthodes avec celles aujourd'hui utilisées pour l'inspection des blochets des voies d'Eurotunnel serait un axe de développement intéressant.

Dans le cadre de la thèse, n'ayant pas de mesures provenant directement d'infrastructures ferroviaires, une étude d'un dispositif fournissant des données similaires a été réalisé. Les mesures provienne de l'expérience de la Colonne Résonante, utilisée en ingénierie civile pour déterminer les propriétés dynamiques d'échantillon de terrains. Une quantité de terre, sable ou argile est confinée, et soumise à des sollicitations de torsions. Lorsque ces sollicitations cessent, on enregistre la réponse du milieu testé grâce à des accéléromètres. A partir des signaux, on estime des paramètres dynamiques synthétisant les propriétés du matériaux. Dans le cas de l'expérience de la Colonne Résonante, les phénomènes physiques mis en jeu peuvent être modélisés par un système à un degré de liberté. Les paramètres dynamiques estimés sont le Taux d'Amortissement D et la Fréquence de Résonance f (notée ω pour la fréquence en radians, avec $\omega = 2\pi f$).

Une hypothèse forte faite en ingénierie civile pour obtenir les estimations des paramètres dynamiques est leur constance au cours du temps. Néanmoins, il a été montré empiriquement que des phénomènes de non-linéarité apparaissent lors des expérimentations. L'hypothèse précédente ne permet pas de modéliser les phénomènes de réponse des échantillons des terrains. La thèse propose d'étudier ces non-linéarités grâce à une estimation du changement des paramètres dynamiques au cours du temps. A chaque instant t, une estimation du vecteur de paramètres dynamiques $u_t = (D_t, f_t)^T$ est réalisée. On fait l'hypothèse que le vecteur u_t varie lentement au cours du temps. Localement, le système à un degré de liberté est donc valable. Contrairement aux chapitres précédents, où l'estimation d'une densité de probabilité filtrée $p(u_t \mid z_{1:t})$ pouvait être réalisée, la présente estimation vise à estimer directement la densité de probabilité jointe $p(u_{1:T} \mid z_{1:T})$. Cette procédure, appelée lissage, permet d'inclure également les observations futures dans l'estimation. L'ensemble des mesures sont prises en compte pour réaliser l'estimation des vecteurs d'états à l'instant t, contrairement à l'estimation par filtre, qui n'utilise que les mesures obtenues jusqu'à l'instant t. Un filtre particulaire Rao-Blackwellisé (filtre particulaire incluant un calcul analytique d'une partie des densités de probabilité) peut être utilisé lorsqu'un sous-ensemble du vecteur d'état est soumis à des évolutions linéaires sous bruit Gaussien. Lorsque les paramètres de ce filtre sont déterminés, dans notre application grâce l'approche par optimisation bayésienne, on peut ensuite appliqué un lisseur particulaire adapté au filtre particulaire Rao-Blackwellisé. La thèse propose également une approche entièrement Bayésienne, qui intègre l'incertitude relative aux paramètres du filtre. Le calcul de la vraisemblance étant stochastique, un algorithme par «Monte Carlo Markov Chain» (MCMC) bruitée est proposé, pour obtenir des échantillons du posterior de manière efficace, à partir de données fournies par des accéléromètres. Des simulations permettent de monter la supériorité de ces deux méthodes par rapport à la procédure de filtrage, pour des variations lentes des paramètres dynamiques.

En faisant varier le niveau des sollicitations, et en répétant l'expérience, on obtient, des courbes de dégradations, représentant l'évolution des paramètres dynamiques en fonction du niveau des sollicitations. La répétition des expérimentations est une opération couteuse en temps. La méthode d'analyse proposée dans cette thèse permet d'extraire, à partir d'un seul signal obtenu avec une sollicitation forte, les paramètres dynamiques pour différents sollicitations changeant au cours du temps. Des courbes de dégradations sont ainsi extraites à partir d'un unique signal. La méthodologie a été testé sur un échantillon de sable d'Ottawa. Une comparaison avec la procédure complète (par de multiples sollicitations), à permis de montrer la consistance des résultats obtenus grâce aux procédures de lissage, que ce soit le lisseur particulaire Rao-Blackwellisé, ou

la procédure par MCMC bruitée. L'approche proposée permet ainsi un gain en temps significatif, tout en permettant d'estimer les paramètres dynamiques en intégrant l'incertitude résidant dans les paramètres du problème. Pour les signaux ultra sonores issues de capteurs EMAT, ceci permet notamment d'identifier les chemins des impulsions, et de les extraire des représentations en B-scan.

Ouverture

La présente thèse a montré l'intérêt de la modélisation probabiliste par modèles d'états sur l'utilisation des signaux vibratoires et ultra-sonores dans un contexte de contrôle non destructif. La propagation de densité des probabilité, effectuée de manière séquentielle au cours du temps, permet d'identifier certaines caractéristiques des ondes mesurée, liées à leurs propriétés physiques. Ceci afin d'améliorer la détection de défauts de surface sur le rail. Des procédure automatisées ont été développées pour palier à l'épineux problème du réglages des algorithmes séquentiels, ne nécessitant pas de connaissance avancée de l'opérateur. Les méthodes ont été éprouvées sur des simulations, données de laboratoires et expérimentales. Les gains obtenus en utilisant ces techniques couvrent un large spectre: possibilité d'estimer des quantités inaccessibles sans associations entre les acquisitions, quantifications des incertitudes des prédictions, et automatisation de l'analyse des signaux.

Cette thèse s'est intéressé à des types des signaux ultra-sonores bien particuliers, obtenus par des dispositifs se déplaçant à faible vitesse. Pour valider l'approche de manière plus industrielle, et éprouver les performances des méthodes développées, des campagnes d'inspections sur de plus grandes longueurs de rail en embarquant le dispositif de mesure sur des trains d'inspection seront nécessaires. Ceci sera permis par l'installation des capteurs dans des structures mobiles appropriées. D'autres types de défauts pourraient également être étudiées, notamment ceux internes aux rails. Ceci nécessitera l'utilisation de capteurs différents. Néanmoins, les outils développés ne sont pas restreint aux seules données EMAT, et leur généralisation à d'autres capteurs pourra être étudiée. Enfin, l'analyse de milieux aux propriétés physiques et mécaniques différentes pourrait être envisagé, comme par exemple les bétons. Les représentations des signaux seront à adapter, mais les processus des filtrages séquentiels pourrait apporter des informations supplémentaires, de façon similaire à ce qui a été réalisé dans cette thèse, sur les signaux ultra-sonores sur rails.

Bibliography

[02] *UIC 712 Défauts de rails*. 4th. 2002.

[69] *UIC 861-3, Profiles Unifiés de Rails à 60 kg. Types UIC 60 et 60 E*. 1969.

[ADH10] C. Andrieu, A. Doucet, and R. Holenstein. *Particle Markov chain Monte Carlo methods*. In: Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72.3 (2010), pp. 269–342.

[AR09] C. Andrieu and G. Roberts. *The Pseudo-Marginal Approach for Efficient Monte Carlo Computations*. In: The Annals of statistics 37.2 (2009), pp. 697–725.

[AT08] C. Andrieu and J. Thoms. *A tutorial on adaptive MCMC*. In: Stat Comput 18.4 (2008), pp. 343–373.

[Bea+12] M. Beard, B.-T. Vo, B.-N. Vo, and S. Arulampalam. *Gaussian mixture PHD and CPHD filtering with partially uniform target birth*. In: 2012 15th International Conference on Information Fusion. IEEE, 2012, pp. 535–541.

[Bea03] M.A. Beaumont. *Estimation of Population Growth or Decline in Genetically Monitored Populations*. In: Genetics 164.3 (2003), pp. 1139–1160.

[Ben04] M. Bentoumi. *Outils pour la détection et la classification : Application au diagnostic de défauts de surface de rail*. PhD dissertation. Université Henri Poincaré - Nancy 1, 2004.

[Ber+11] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. *Algorithms for Hyper-Parameter Optimization*. In: Proceedings of the 24th International Conference on Neural Information Processing Systems. NIPS'11. Red Hook, NY, USA: Curran Associates Inc., 2011, pp. 2546–2554.

[BG98] S.P. Brooks and A. Gelman. *General Methods for Monitoring Convergence of Iterative Simulations*. In: Journal of Computational and Graphical Statistics 7.4 (1998), pp. 434–455.

[BL95] Y. Bar-Shalom and X.R. Li. *Multitarget-multisensor tracking: principles and techniques*. 3rd ed. Yaakov Bar-Shalom, 1995.

[Bro+11] S. Brooks, A. Gelman, G. Jones, and X.L. Meng. *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, 2011.

[But00] J.C Butcher. *Numerical methods for ordinary differential equations in the 20th century*. In: Journal of Computational and Applied Mathematics 125.1 (2000), pp. 1–29.

[CB07] D.E. Clark and J. Bell. *Multi-target state estimation and track continuity for the particle PHD filter*. In: IEEE transactions on aerospace and electronic systems 43.4 (2007), pp. 1441–1453.

[CH10] J. Choi and J. Hong. *Characterization of wavelet coefficients for ultrasonic signals*. In: Journal of Applied Physics 107.11 (2010), pp. 114909–114909–6.

[Cha+11] S. Challa, M.R. Morelande, D. Mušicki, and R.J. Evans. *Fundamentals of Object Tracking*. Cambridge University Press, 2011.

[Che+12] X. Chen, R. Tharmarasa, M. Pelletier, and T. Kirubarajan. *Integrated Clutter Estimation and Target Tracking using Poisson Point Processes*. In: IEEE transactions on aerospace and electronic systems 48.2 (2012), pp. 1210–1235.

[Che18] L. Chen. *From labels to tracks: it's complicated*. In: Signal Processing, Sensor/Information Fusion, and Target Recognition XXVII. Ed. by Ivan Kadar. Vol. 10646. International Society for Optics and Photonics. SPIE, 2018, pp. 8–13.

[Cla+19] Benjamin Claudet, Tien Hoang, Denis Duhamel, Gilles Foret, Jean-Luc Pochet, and Francis Sabatier. *Wave finite element method for computing the dynamic response of railway transition zones subjected to moving loads*. In: 7th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering Methods in Structural Dynamics and Earthquake Engineering. 2019, pp. 4538–4547.

[CS05] G. Cardoso and J. Saniie. *Ultrasonic data compression via parameter estimation*. In: IEEE Trans. Ultrason., Ferroelect., Freq. Contr. 52.2 (2005), pp. 313–325.

[Cur+08a] R.O. Curadelli, J.D. Riera, D. Ambrosini, and M.G. Amani. *Damage detection by means of structural damping identification*. In: Engineering Structures 30.12 (2008), pp. 3497–3504.

[Cur+08b] R.O. Curadelli, J.D. Riera, D. Ambrosini, and M.G. Amani. *Damage detection by means of structural damping identification*. In: Engineering Structures 30.12 (2008), pp. 3497–3504.

[DGA00] A. Doucet, S. Godsill, and C. Andrieu. *On sequential Monte Carlo sampling methods for Bayesian filtering*. In: Statistics and computing 10.3 (2000), pp. 197–208.

[DKT16] A. Dey, J. Kurz, and L. Tenczynski. *Detection and evaluation of rail defects with non-destructive testing methods*. In: 19th World Conference on Non-Destructive Testing. 2016, pp. 1–9.

[DLR77] A. P. Dempster, N. M. Laird, and D. B. Rubin. *Maximum Likelihood from Incomplete Data Via the EM Algorithm*. In: Journal of the Royal Statistical Society: Series B (Methodological) 39.1 (1977), pp. 1–22.

[Dou+15] A. Doucet, M.K. Pitt, G. Deligiannidis, and R. Kohn. *Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator*. In: Biometrika 102.2 (2015), pp. 295–313.

[DS01a] R. Demirli and J. Saniie. *Model-based estimation of ultrasonic echoes. Part I: Analysis and algorithms*. In: IEEE Trans. Ultrason., Ferroelect., Freq. Contr. 48.3 (2001), pp. 787–802.

[DS01b] R. Demirli and J. Saniie. *Model-based estimation of ultrasonic echoes. Part II: Nondestructive evaluation applications*. In: IEEE Trans. Ultrason., Ferroelect., Freq. Contr. 48.3 (2001), pp. 803–811.

[Fel97] M. Feldman. *Non-linear free vibration identification via the Hilbert transforms*. In: Journal of Sound and Vibration 208.3 (1997), pp. 475–489.

[FH94] J.A. Fessler and A.O. Hero. *Space-alternating generalized expectation-maximization algorithm*. In: IEEE Trans. Signal Process. 42.10 (1994), pp. 2664–2677.

[FM20] J. Facciorusso and C. Madiai. *On Cohesive Soil Damping Estimation by Free Vibration Method in Resonant Column Test*. In: Geotechnical Testing Journal (2020).

[FS11] T. Flury and N. Shephard. *Bayesian inference based only on simulated likelihood: Particle filter analysis of dynamic economic models*. In: Econometric Theory 27.5 (2011), pp. 933–956.

[Gar+18] J. García-Gómez, R. Gil-Pita, M. Rosa-Zurera, A. Romero-Camacho, J.A. Jiménez-Garrido, and V. García-Benavides. *Smart Sound Processing for Defect Sizing in Pipelines Using EMAT Actuator Based Multi-Frequency Lamb Waves*. In: Sensors 18.3 (2018), p. 802.

[GDW04] S.J. Godsill, A. Doucet, and M. West. *Monte Carlo Smoothing for Nonlinear Time Series*. en. In: Journal of the American Statistical Association 99.465 (2004), pp. 156–168.

[GGV11] J. García-Martín, J. Gómez-Gil, and E. Vázquez-Sánchez. *Non-Destructive Techniques Based on Eddy Current Testing*. In: Sensors 11.3 (2011), pp. 2525–2565.

[Gil13] J. Gilles. *Empirical Wavelet Transform*. In: IEEE Transactions on Signal Processing 61.16 (2013), pp. 3999–4010.

[GMN97] I.R. Goodman, R.P. Mahler, and H.T. Nguyen. *Mathematics of Data Fusion*. USA: Kluwer Academic Publishers, 1997.

[Gom18] M.E. Gomes-borges. *Real-Time Sensor Management Strategies for Multi-Object Tracking*. Theses. Ecole Centrale de Lille, 2018.

[GR92] A. Gelman and D.B. Rubin. *Inference from Iterative Simulation Using Multiple Sequences*. In: Statistical Science 7.4 (1992), pp. 457–472.

[HL10] J. Houssineau and D. Laneuville. *PHD filter with diffuse spatial prior on the birth process with applications to GM-PHD filter*. In: 2010 13th International Conference on Information Fusion. Edinburgh: IEEE, 2010, pp. 1–8.

[Hoa17] Tien Hoang. *Comportement mécanique et dégradation d'une voie ferrée en tunnel*. PhD dissertation. Paris Est, 2017.

[Hua+09] S. Huang, W. Zhao, Y. Zhang, and S. Wang. *Study on the lift-off effect of EMAT*. In: Sensors and Actuators A: Physical 153.2 (2009), pp. 218–221.

[HVV15] H.G. Hoang, B.-T. Vo, and B.-N. Vo. *A fast implementation of the generalized labeled multi-Bernoulli filter with joint prediction and update*. In: 2015 18th International Conference on Information Fusion (Fusion). ISIF, 2015, pp. 999–1006.

[HW08] N.E. Huang and Z. Wu. *A review on Hilbert-Huang transform: Method and its applications to geophysical studies*. In: Reviews of geophysics 46.2 (2008).

[HWZ12] M.R. Hoseini, X. Wang, and M.J. Zuo. *Estimating ultrasonic time of flight using envelope and quasi maximum likelihood method for damage detection and assessment*. In: Measurement 45.8 (2012), pp. 2072–2080.

[IHF16] J. Isla-Garcia, B. Herdovics, and Cegla F. *SHM with EMATs*. In: 8th European Workshop on Structural Health Monitoring (EWSHM 2016), July 5-8, 2016 in Bilbao, Spain (EWSHM 2016). 8. 2016.

[Jia+07] X. Jian, S. Dixon, N. Guo, and R. Edwards. *Rayleigh wave interaction with surface-breaking cracks*. In: Journal of Applied Physics 101 (2007).

[JSW98] D. Jones, M. Schonlau, and W. Welch. *Efficient Global Optimization of Expensive Black-Box Functions*. In: Journal of Global Optimization 13 (1998), pp. 455–492.

[Kal60] R.E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. In: Journal of Basic Engineering 82.1 (1960), pp. 35–45.

[Kit87] G. Kitagawa. *Non-Gaussian State-Space Modeling of Nonstationary Time Series*. In: Journal of the American Statistical Association 82.400 (1987), pp. 1032–1041.

[Kit96] G. Kitagawa. *Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models*. In: Journal of Computational and Graphical Statistics 5.1 (1996), pp. 1–25.

[Kub+11] M. Kubinyi, O. Kreibich, J. Neuzil, and R. Smid. *EMAT noise suppression using information fusion in stationary wavelet packets*. In: IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control 58 (2011), pp. 1027–36.

[KWM08] M. Kerr, A. Wilson, and S. Marich. *The Epidemiology of Squats and Related Defects*. In: Conference on Railway Engineering (2008 : Perth, W.A.) Railway Technical Society of Australasia : Engineers Australia, 2008, pp. 83–96.

[LHL10] F. Lian, c. Han, and W. Liu. *Estimating Unknown Clutter Intensity for PHD Filter*. In: IEEE transactions on aerospace and electronic systems 46.4 (2010), pp. 2066–2078.

[Lin13] F. Lindsten. *Backward Simulation Methods for Monte Carlo Statistical Inference*. In: FNT in Machine Learning 6.1 (2013), pp. 1–143.

[LLC10] F. Liang, C. Liu, and R. Carroll. *Advanced Markov Chain Monte Carlo Methods: Learning from Past Samples*. Chichester, UK: John Wiley & Sons, Ltd, 2010.

[LLY09] C. Lin, P. Liu, and P. Yeh. *Application of empirical mode decomposition in the impact-echo test*. In: NDT & E International 42.7 (2009), pp. 589–598.

[Lu+16] Z. Lu, C. Yang, D. Qin, and Y. Luo. *Estimating the parameters of ultrasonic echo signal in the Gabor transform domain and its resolution analysis*. In: Signal Processing 120 (2016), pp. 607–619.

[LVN16] S. Lin, B.-T. Vo, and S.E. Nordholm. *Measurement driven birth model for the generalized labeled multi-Bernoulli filter*. In: 2016 International Conference on Control, Automation and Information Sciences (ICCAIS). Ansan, South Korea: IEEE, 2016, pp. 94–99.

[LW01] J. Liu and M. West. *Combined Parameter and State Estimation in Simulation-Based Filtering*. In: Sequential Monte Carlo Methods in Practice. Ed. by A. Doucet, N. de Freitas, and N. Gordon. New York, NY: Springer New York, 2001, pp. 197–223.

[MA10] I. Murray and R.P. Adams. *Slice sampling covariance hyperparameters of latent Gaussian models*. In: 23 (2010). Ed. by J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta.

[Mah03] R. Mahler. *Multitarget bayes filtering via first-order multitarget moments*. In: IEEE transactions on aerospace and electronic systems 39.4 (2003), pp. 1152–1178.

[Mah07a] R. Mahler. *PHD filters of higher order in target number*. In: IEEE transactions on aerospace and electronic systems 43.4 (2007), pp. 1523–1543.

[Mah07b] R. Mahler. *Statistical Multisource-Multitarget Information Fusion*. USA: Artech House, Inc., 2007.

[Mah12] R. Mahler. *A comparison of "clutter-agnostic" PHD filters*. In: Signal Processing, Sensor Fusion, and Target Recognition XXI. Ed. by Ivan Kadar. Vol. 8392. International Society for Optics and Photonics. SPIE, 2012, pp. 285–296.

[Mah13a] R. Mahler. *"Statistics 102" for Multisource-Multitarget Detection and Tracking*. In: IEEE journal of selected topics in signal processing 7.3 (2013), pp. 376–389.

[Mah13b] R. Mahler. *Divergence detectors for multitarget tracking algorithms*. In: Signal Processing, Sensor Fusion, and Target Recognition XXII. Ed. by Ivan Kadar. Vol. 8745. International Society for Optics and Photonics. SPIE, 2013, pp. 157–167.

[Mal08] S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. 3rd. USA: Academic Press, Inc., 2008.

[McK+14] T. McKinley, J. Ross, R. Deardon, and A. Cook. *Simulation-based Bayesian inference for epidemic models*. In: Computational Statistics & Data Analysis 71 (2014), pp. 434–447.

[MK08] G.J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. 2008, p. 400.

[MPP03] V. Moshkovich, A. Passi, and G. Passi. *Recent Advances in the Ultrasonic Inspection Recording and Reporting - Instrumentation: Part 1*. In: e-Journal of Nondestructive Testing (NDT) ISSN 1435-4934 (NDT.net Journal) 9.9 (2003).

[MVV11] R. Mahler, B.-T. Vo, and B.-N. Vo. *CPHD Filtering With Unknown Clutter Rate and Detection Profile*. In: IEEE transactions on signal processing 59.8 (2011), pp. 3497–3513.

[MZ93] S.G. Mallat and Z. Zhang. *Matching pursuits with time-frequency dictionaries*. In: IEEE Trans. Signal Process. 41.12 (1993), pp. 3397–3415.

[Nda16] B. Ndao. *Génération et détection sans contact des ondes de Rayleigh par méthodes ultrasons-laser et EMAT en mode statique et dynamique : application à la détection de défauts surfaciques dans le champignon du rail*. Dissertation in French. Université Polytechnique Hauts-de-France. PhD dissertation. Valenciennes, 2016.

[Ni+10] C. Ni, Y. Shi, Z. Shen, J. Lu, and X. Ni. *An analysis of angled surface-breaking crack detection by dual-laser source generated ultrasound*. In: NDT & E International 43.6 (2010), pp. 470–475.

[NPY11] I. Nevat, G. Peters, and J. Yuan. *Channel Tracking in Relay Systems via Particle MCMC*. In: 2011 IEEE Vehicular Technology Conference (VTC Fall). 2011, pp. 1–5.

[NW06] J. Nocedal and S.J. Wright. *Numerical Optimization*. 2nd edition. Springer Series in Operations Research and Financial Engineering. New York: Springer, 2006.

[OF98] J.C O'Neill and P Flandrin. *Chirp hunting*. In: Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis (Cat. No.98TH8380). IEEE, 1998, pp. 425–428.

[PCV09] K. Panta, D.E. Clark, and B.-N. Vo. *Data Association and Track Management for the Gaussian Mixture Probability Hypothesis Density Filter*. In: IEEE transactions on aerospace and electronic systems 45.3 (2009), pp. 1003–1016.

[Pet+11] G. Peters, M. Briers, P. Shevchenko, and A. Doucet. *Calibration and Filtering for Multi Factor Commodity Models with Seasonality: Incorporating Panel Data from Futures Contracts*. In: Methodology And Computing In Applied Probability 15 (2011).

[PRD08] M. Papaelias, C. Roberts, and C. Davis. *A review on non-destructive evaluation of rails: State-of-the-art and future development*. In: Proceedings of The Institution of Mechanical Engineers Part F-journal of Rail and Rapid Transit - PROC INST MECH ENG F-J RAIL R 222 (2008), pp. 367–384.

[Pun17] Y. Punchihewa. *Efficient generalized labeled multi-bernoulli filter for jump Markov system*. In: 2017 International Conference on Control, Automation and Information Sciences (ICCAIS). Chiang Mai: IEEE, 2017, pp. 221–226.

[Qiu+14] X.B. Qiu, L.L. Liu, C.L. Li, J.L. Wei, Y.F. Wu, and X.C. Cui. *Defect Classification by Pulsed Eddy-Current Technique Based on Power Spectral Density Analysis Combined With Wavelet Transform*. In: IEEE Transactions on Magnetics 50.9 (2014), pp. 1–8.

[Raj+18] J. Rajamäki, M. Vippola, A. Nurmikolu, and T. Viitala. *Limitations of eddy current inspection in railway rail evaluation*. In: Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 232.1 (2018), pp. 121–129.

[RC04] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. New York, NY: Springer New York, 2004.

[RCV10] B. Ristic, D. Clark, and B.-N. Vo. *Improved SMC implementation of the PHD filter*. In: 2010 13th International Conference on Information Fusion. Edinburgh: IEEE, 2010, pp. 1–8.

[Reu+13a] S. Reuter, B.-T. Vo, B. Wilking, D. Meissner, and K. Dietmayer. *Divergence detectors for the delta-generalized labeled multi-Bernoulli filter*. In: 2013 Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF). IEEE, 2013, pp. 1–6.

[Reu+13b] S. Reuter, B. Wilking, J. Wiest, M. Munz, and K. Dietmayer. *Real-Time Multi-Object Tracking using Random Finite Sets*. In: IEEE transactions on aerospace and electronic systems 49.4 (2013), pp. 2666–2678.

[Reu+14] S. Reuter, B. Vo, B. Vo, and K. Dietmayer. *The Labeled Multi-Bernoulli Filter*. In: IEEE Transactions on Signal Processing 62.12 (2014), pp. 3246–3260.

[Ric45] S.O. Rice. *Mathematical Analysis of Random Noise*. In: Bell System Technical Journal 24.1 (1945), pp. 46–156.

[Ris+12] B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo. *Adaptive Target Birth Intensity for PHD and CPHD Filters*. In: IEEE transactions on aerospace and electronic systems 48.2 (2012), pp. 1656–1668.

[RTS65] H.E. Rauch, F. Tung, and C.T. Striebel. *Maximum likelihood estimates of linear dynamic systems*. en. In: AIAA Journal 3.8 (1965), pp. 1445–1450.

[RW06] C.E. Rasmussen and C.K.I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. Cambridge, Mass: MIT Press, 2006.

[Sär13] S. Särkkä. *Bayesian Filtering and Smoothing*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2013.

[SBG12] S. Särkkä, P. Bunch, and S.J. Godsill. *A Backward-Simulation Based Rao-Blackwellized Particle Smoother for Conditionally Linear Gaussian Models*. In: IFAC Proceedings Volumes 45.16 (2012), pp. 506–511.

[Sch+16] I. Schlangen, J. Franco, J. Houssineau, W.T.E. Pitkeathly, D. Clark, I. Smal, and C. Rickman. *Marker-Less Stage Drift Correction in Super-Resolution Microscopy Using the Single-Cluster PHD Filter*. In: IEEE journal of selected topics in signal processing 10.1 (2016), pp. 193–202.

[Sha+16] B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, and N. de Freitas. *Taking the Human Out of the Loop: A Review of Bayesian Optimization*. In: Proc. IEEE 104.1 (2016), pp. 148–175.

[She+15] C. Sherlock, A. Thiery, G. Roberts, and J. Rosenthal. *On the efficiency of pseudo-marginal random walk Metropolis algorithms*. In: The Annals of statistics 43.1 (2015), pp. 238–275.

[SLA12] J. Snoek, H. Larochelle, and R.P. Adams. *Practical Bayesian Optimization of Machine Learning Algorithms*. In: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2. NIPS'12. Red Hook, NY, USA: Curran Associates Inc., 2012, pp. 2951–2959.

[SNV17] S. Santa-aho, A. Nurmikolu, and M. Vippola. *Automated Ultrasound-based Inspection of Rails: Review*. In: International Journal of Railway 10 (2017), pp. 21–29.

[Son+11] Z. Song, T. Yamada, H. Shitara, and Y. Takemura. *Detection of Damage and Crack in Railhead by Using Eddy Current Testing*. In: Journal of Electromagnetic Analysis and Applications 3 (2011), pp. 546–550.

[STF17] A. Sophian, G. Tian, and M. Fan. *Pulsed Eddy Current Non-destructive Testing and Evaluation: A Review*. In: Chinese Journal of Mechanical Engineering 30 (2017), pp. 500–514.

[SVL07] S. Särkkä, A. Vehtari, and J. Lampinen. *Rao-Blackwellized particle filter for multiple target tracking*. In: Information Fusion 8.1 (2007). Special Issue on the Seventh International Conference on Information Fusion-Part II, pp. 2–15.

[SVV08] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. *A Consistent Metric for Performance Evaluation of Multi-Object Filters*. In: IEEE Transactions on Signal Processing 56.8 (2008), pp. 3447–3457.

[TC98] C. Torrence and G. Compo. *A Practical Guide to Wavelet Analysis*. In: Bulletin of the American Meteorological Society 79.1 (1998), pp. 61–78.

[VD91] M. Vucetic and R. Dobry. *Effect of Soil Plasticity on Cyclic Response*. In: Journal of Geotechnical Engineering 117 (1991), pp. 89–107.

[Vih12] M. Vihola. *Robust adaptive Metropolis algorithm with coerced acceptance rate*. In: Stat Comput 22.5 (2012), pp. 997–1008.

[VM06] B.-N. Vo and W.-K. Ma. *The Gaussian Mixture Probability Hypothesis Density Filter*. In: IEEE transactions on signal processing 54.11 (2006), pp. 4091–4104.

[Vo+13] B.-T. Vo, B.-N. Vo, R. Hoseinnezhad, and R. Mahler. *Robust Multi-Bernoulli Filtering*. In: IEEE journal of selected topics in signal processing 7.3 (2013), pp. 399–409.

[VSD05] B.-N. Vo, S. Singh, and A. Doucet. *Sequential Monte Carlo methods for multitarget filtering with random finite sets*. In: IEEE Transactions on Aerospace and Electronic Systems 41.4 (2005), pp. 1224–1245.

[VV13] B.-T. Vo and B.-N. Vo. *Labeled Random Finite Sets and Multi-Object Conjugate Priors*. In: IEEE transactions on signal processing 61.13 (2013), pp. 3460–3475.

[VVC07] B.-T. Vo, B.-N. Vo, and A. Cantoni. *Analytic Implementations of the Cardinalized Probability Hypothesis Density Filter*. In: IEEE transactions on signal processing 55.7 (2007), pp. 3553–3567.

[VVE11] A.T. Vu, B.-N. Vo, and R. Evans. *Particle Markov Chain Monte Carlo for Bayesian multi-target tracking*. In: 14th International Conference on Information Fusion. IEEE, 2011, pp. 1–8.

[VVH17] B.-N. Vo, B.-T. Vo, and H.G. Hoang. *An Efficient Implementation of the Generalized Labeled Multi-Bernoulli Filter*. In: IEEE Trans. Signal Process. 65.8 (2017), pp. 1975–1987.

[VVP14] B.-N. Vo, B.-T. Vo, and D. Phung. *Labeled Random Finite Sets and the Bayes Multi-Target Tracking Filter*. In: IEEE transactions on signal processing 62.24 (2014), pp. 6554–6567.

[War16] A.N.I. Wardana. *A comparative study of EMD, EWT and VMD for detecting the oscillation in control loop*. In: 2016 International Seminar on Application for Technology of Information and Communication (ISemantic). 2016, pp. 58–63.

[WH09] Z. Wu and N.E. Huang. *Ensemble Empirical Mode Decomposition: A noise-assisted data analysis method*. In: Adv. Adapt. Data Anal. 01.01 (2009), pp. 1–41.

[XGY09] B. Xu, V. Giurgiutiu, and L. Yu. *Lamb waves decomposition and mode identification using matching pursuit method*. In: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2009. Ed. by Masayoshi Tomizuka. Vol. 7292. International Society for Optics and Photonics. SPIE, 2009, pp. 161–172.

[Yan+16] J. Yang, L. Yang, Y. Yuan, and H. Ge. *Probability hypothesis density filter with adaptive parameter estimation for tracking multiple maneuvering targets*. In: Chinese Journal of Aeronautics 29.6 (2016), pp. 1740–1748.

[YJH17] W. Yi, M. Jiang, and R. Hoseinnezhad. *The Multiple Model Vo–Vo Filter*. In: IEEE Transactions on Aerospace and Electronic Systems 53.2 (2017), pp. 1045–1054.

Development of new methods for Non Destructive Testing for Railway Infrastructure. Use of statistical signal processing.

Maintenance of railway infrastructures has nowadays to anticipate the degradation of equipment to avoid any damaging incidents. Detecting a defect on the rail at its earliest stage is necessary to ensure efficient and optimal maintenance. Electro-Magnetic Acoustic Transducers (EMAT) are a contactless ultrasonic inspection technique for non-destructive assessment of the condition of ferromagnetic materials. Their application to rails allows identification of surface defects when measurements are obtained from immobile sensors. However, railway infrastructure operations are urging to reduce the inspection time of rails. Measurement tools must thus be carried into motorized structures. Mobile inspections cause degradations of ultrasonic signals. This thesis proposes tools and methodologies from statistical signal processing to improve the detection and characterization of surface defects on the rail, from ultrasonic signals obtained by mobile sensors. First, statistical indicators on EMAT signals to identify rail defects and representations of measurements in state spaces are introduced. They provide information on the evolution of ultrasound signals during mobile measurements, using filtering algorithms, which enable quantification of uncertainty about the made estimates. A methodology for automatically adjusting the parameters of these filtering algorithms to adapt to received data is then presented. The methodology is based on Bayesian modelling, requiring little knowledge of algorithms from a human operator. Finally, the more general case of vibrational acceleration signals is investigated. Identifying the evolution of characteristics within these signals allows a better understanding of the physical phenomena governing these signals. Bayesian estimation procedures incorporate uncertainty about the prior knowledge of the problem, and the results provided by estimation methods.

Keywords: Multi-Object Tracking, Ultrasonic Testing, Rail Surface Defect Detection, Vibratory Signal Analysis, Bayesian Statistical Estimation

Développement de Nouvelles Méthodes pour l'Évaluation Non-Destructive des Infrastructures Ferroviaires. Utilisation de traitements statistiques des signaux.

La maintenance des infrastructures ferroviaires doit désormais anticiper la dégradation du matériel pour éviter tout incident dommageable. Détecter un défaut sur le rail à son stade le plus précoce est nécessaire pour garantir une maintenance efficace et optimale. Les Electro-Magnetic Acoustic Transducers (EMAT) sont une technique d'inspection ultrasonore sans-contact pour l'évaluation non destructive de l'état des matériaux ferromagnétiques. Leur application aux rails permet l'identification des défauts de surface lors d'une prise de mesure immobile sur le rail. Néanmoins, l'exploitation du réseau poussant à la diminution du temps d'inspection, les outils de mesures doivent être embarqués dans des structures motorisées. L'inspection mobile cause une dégradation des signaux ultrasonores. Cette thèse propose des outils et méthodologies du traitement statistique du signal pour améliorer la détection et la caractérisation des défauts de surface du rail, à partir de signaux ultrasonores obtenues par des capteurs mobiles. Nous proposons des indicateurs statistiques sur les signaux EMAT permettant une identification des défauts du rail. Des représentations des mesures dans des espaces d'états apportent des informations sur l'évolution des signaux durant la mesure mobile, grâce à des algorithmes de filtrage, qui fournissent des mesures d'incertitudes sur les estimations réalisées. Nous apportons une méthodologie permettant de régler automatiquement les paramètres de ces algorithmes de filtrage pour s'adapter aux données reçues. Elle repose sur des modélisations bayésiennes, nécessitant peu de connaissances sur les algorithmes de la part d'un opérateur humain. Enfin, nous nous sommes intéressés au cas plus général des signaux vibratoires d'accélération. L'identification de l'évolution des caractéristiques au sein de ces signaux permet une meilleure compréhension des phénomènes physiques régissant la mesure. Des procédures bayésiennes intègrent l'incertitude sur les connaissances a priori du problème, et les résultats fournis par les méthodes d'estimations.

Mots clés : Tracking Multi-Objects, Inspection Ultrasonore, Détection de Défauts de Surface du Rail, Analyse de Signaux Vibratoires, Estimation Statistique Bayésienne