Diagnosis of Soft Faults in Complex Wired Networks

Nour Taki

université PARIS-SACLAY

Diagnostic des Défauts Non-Francs dans les Réseaux Filaires Complexes

Diagnosis of Soft Faults in Complex Wired Networks

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°580: Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat : Traitement du signal et des images
Graduate School : Informatique et science du numérique, Référent : Faculté des sciences d'Orsay

Thèse préparée dans la Laboratoire des signaux et systèmes (Université Paris-Saclay, CNRS, CentraleSupélec) et l'Institut LIST (Université Paris-Saclay, CEA), sous la direction de Claude DELPHA, Maître de Conférences, HDR, les co-encadrements de Demba DIALLO, Professeur des universités, Wafa BEN HASSEN, Docteur-Ingénieur R&D, et Nicolas RAVOT, Docteur-Ingénieur R&D

Thèse soutenue à Paris-Saclay, le 28 Février 2022, par

Nour TAKI

Composition du jury

Maan El Badaoui El Najjar Professeur des universités, Université de Lille	Président
Pascal Maussion Professeur des universités, Université de Toulouse	Rapporteur
Virginie Degardin Professeure des universités, Université de Lille	Rapporteur
Lionel Pichon Directeur de Recherche C.N.R.S., Université Paris Saclay	Examinateur
Yassine Amirat Enseignant chercheur, HDR, Institut Supérieur d'Electronique et du Numérique de Brest	Examinateur
Claude Delpha Maître de Conférences, HDR, Université Paris Saclay	Directeur de Thèse
Demba Diallo Professeur des universités, Université Paris Saclay	Co-Encadrant
Wafa Ben Hassen Docteur-Ingénieur R&D au CEA	Co-Encadrante
Nicolas Ravot Docteur-Ingénieur R&D au CEA	Co-Encadrant

THESE DE DOCTORAT
NNT : 2022UPASG004

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Diagnostic des Défauts Non-Francs dans les Réseaux Filaires Complexes

Mots clés: Diagnostic filaire complexe; Défaut non-franc; Réflectométrie dans le domaine temporel; Réflectométrie distribuée; Analyse en composantes principales; Sélection de fréquence maximale; Sélection de capteurs.

Résumé: Les câbles électriques sont utilisés dans tous les secteurs pour transférer de l'énergie ou de l'information. Pendant le fonctionnement, les câbles peuvent être sujets à des défauts francs (circuit ouvert ou court-circuit) ou des défauts non-francs (endommagement de l'isolant, pincement, etc.) dus à une mauvaise utilisation, aux conditions environnementales ou au vieillissement. Ces défauts doivent être détectés à leur stade le plus précoce pour éviter une interruption de la fonction ou des conséquences plus graves. Parmi les méthodes de diagnostic des réseaux filaires qui ont été étudiées dans la littérature, la réflectométrie électrique a été considérée la plus efficace surtout dans le cas d'un défaut franc. Cependant, cette méthode s'avère moins fiable en présence d'un défaut non-franc caractérisé, généralement, par une signature de faible amplitude sur le réflectogramme qui dépend non seulement de la variation de l'impédance caractéristique du câble au niveau du défaut mais également de la configuration du signal de test telle que sa bande passante. En effet, l'augmentation de la fréquence maximale du signal de test améliore la résolution "spatiale" de l'information des défauts non-francs. Cependant, elle accentue, en même temps, les phénomènes d'atténuation et de dispersion du signal de test rendant ainsi la détection de ces défauts moins fiable, et surtout dans le cas des réseaux filaires complexes où la réflectométrie pourrait souffrir de problèmes d'ambiguïté liée à la localisation des défauts. Dans ce cadre, la réflectométrie distribuée où plusieurs capteurs sont installés aux extrémités du réseau sous test est appliquée entraînant l'apparition d'autres problématiques telles que le partage des ressources, la fusion de capteurs pour la prise de décision, la consommation d'énergie, etc.

Dans ce contexte, cette thèse propose de développer deux approches : la première permet de choisir la meilleure fréquence maximale à appliquer au signal de test pour la détection des défauts non-francs. La seconde approche a pour objectif de choisir les capteurs les plus pertinents pour leur diagnostic dans les réseaux filaires complexes. Pour cela, une combinaison entre les données basées sur la réflectométrie et l'algorithme d'analyse en composantes principales (PCA) est utilisée. Le modèle de la PCA est développé pour détecter les défauts non-francs existants. Associés à une analyse statistique basée sur Hotelling's T^2 et Squared Prediction Error (SPE), les paramètres requis sont identifiés. Une étude expérimentale est réalisée, et une analyse de leurs performances en environnement bruité est effectuée.

Title: Diagnosis of Soft Faults in Complex Wired Networks

Keywords: Complex wire diagnosis; Soft fault; Time domain reflectometry; Distributed reflectometry; Principal component analysis; Maximal frequency selection; Sensors selection.

Abstract: Electrical cables are used in all sectors to transfer energy or information. During operation, the cables may be subject to hard faults (open circuit, short circuit) or soft faults (isolation damage, pinching, etc.) due to misuse, environmental conditions, or aging. These faults must be detected at their earliest stage to avoid interruption of the function or more serious consequences. Even though several electric and non-electric wire diagnosis methods have been studied and developed throughout the last few decades, reflectometry-based techniques have provided effective results with hard faults. However, they have been shown to be less reliable whenever soft faults are addressed. Indeed, soft faults are characterized by a small impedance variation, resulting in a low amplitude signature on the corresponding reflectograms. Accordingly, the detection of these faults depends strongly on the test signal configuration, such as its bandwidth. Although the increase of the maximal frequency of the test signal enhances the soft fault's "spatial" resolution, its performance is limited by signal attenuation and dispersion. Moreover, although reflectometry offers good results in point-to-point topology networks, it suffers from ambiguity related to fault location in more complex wired networks (Multi-branched). As a solution, distributed reflectometry method, where sensors are implemented in the extremities of the network under test, is used. However, several issues are enforced, from the computing complexities and sensors fusion problems to the energy consumption.

In this context, this Ph.D. dissertation proposes to develop two approaches: the first selects the best maximal frequency for soft fault detection, and the second selects the most relevant sensors to monitor and diagnose those faults in multi-branched wired networks. The proposed solution is based on a combination between reflectometry and Principal Component Analysis (PCA). The PCA model coupled with statistical analysis based on Hotelling's T^2 and Squared Prediction Error (SPE) is used to detect the soft faults and select the required parameters. Experimental validation is carried out, and performance analysis in the presence of noise is investigated.

Acknowledgment

This thesis took place between the Reliability and Sensor Integration Laboratory (LFIC) of CEA LIST and the Signal and Systems laboratory (L2S) at CentraleSupelec. I would therefore like to thank the CEA for giving me the opportunity to carry out this work through the supervision of Dr. Wafa Ben Hassen and Dr. Nicolas Ravot, who supervised me throughout these three years. They were always available, listened to my many questions, and helped me navigate the progress of my work. The many discussions we had and their advice have a lot to do with the result of this work. This thesis owes them a lot. For all this, thank you.

I want to express my deep gratitude to Prof. Claude Delpha, my thesis director. Thank you for your devoted time to guide me throughout these three years. Thank you for your invaluable advice, ongoing support, and endless patience. I had a great pleasure and honor to work with such an outstanding researcher as you. Thank you for the help and support, both professionally and on a human level, which I needed in the most challenging times.

My sincere thanks to my thesis co-director, Prof. Demba Diallo, for his support and kindness. Thank you very much for your help, valuable comments, ongoing support, and for your modesty.

My thanks also go to the jury members who have accepted to evaluate the work achieved in this thesis. Starting with Mr. Maan El Badaoui El Najjar, Professor at Université de Lille, for doing me the honor of chairing the jury. I am very grateful to Mr. Pascal Maussion, Professor at Université de Toulouse, and Mrs. Virginie Degardin, Professor at Université de Lille, who agreed to review this thesis and the time they devoted to do that. I would like to sincerely thank Mr. Lionel Pichon, CNRS Director at Université Paris Saclay, and Mr. Yassine Amirat, Associate Professor at the Institut Supérieur d'Electronique et du Numérique de Brest, for their participation as examiners.

I extend my thanks to all members of LFIC and L2S laboratories. A huge thank you to Mr. Nicolas Gregis for the precious help you gave me, especially during the last months of the thesis! Thank you for your precious advice and comments.

All this little dream would have never been possible without my family's love, trust, and support. I am so grateful to my dearest mom Yosra, my dear dad Hassan, and to my dearest sisters Fatima and Sara, for your immense love and encouragement, for your continuous support and patience. Thank you for being by my side during my ups and downs. Thank you for giving me so much joy and warmth in life. Thank you for always being there for me. Thank you for teaching me to never give up.

Contents

Acronyms

P_d	Probability of Detection. 62, 89, 156
P_{FA}	Probability of False Alarm. 62, 89, 156
SPE	Squared Prediction Error. 56, 76, 105, 154
ABS	Anti-lock Braking System. 7
AC	Air Conditioner. 7 Alternating current. 11
AFSA	Air Force Safety Agency. 1, 12, 151
ANN	Artificial Neural Networks. 48
AWGN	Additive White Gaussian Noise. 91
BFS	Breadth-First Search. 53
BTDR	Binary Time Domain Reflectometry. 29
CAD	Computer-Aided Design. 30
CAGR	Compound Annual Growth Rate. 6
CAN	Controller Area Network. 9, 105
CPV	Cumulative Percent Variance. 55
CTDR	Chaos Time Domain Reflectometry. 29
CTFDR	Cluster Time Frequency Domain Reflectometry. 39
CUT	Cable Under Test. 73
DC	Direct current. 11
DSP	Digital Signal Processor. 30
ECU	Electronic Control Units. 7
EM	Electromagnetic. 24
EMC	Electromagnetic Compatibility. 37
FAA	Federal Aviation Agency. 67
FDD	Fault Detection and Diagnosis. 54
FDR	Frequency Domain Reflectometry. 3, 6, 27
FDTD	Finite-Difference Time-Domain. 24
FEP	Fluorinated Ethylene Propylene. 64
FMCW	Frequency Modulated Continuous Wave. 27
FPGA	Field-Programmable Gate Array. 30
FTP	Foiled Twisted Pair. 9
GA	Genetic Algorithm. 48, 50
IT	Information Technology. 6
JTFDR	Joint Time Frequency Domain Reflectometry. 38

kg	kilograms. 7
km	kilometers. 7
MCTDR	Multi-Carrier Time Domain Reflectometry. 29
MLP-NN	Multi-Layer Perceptron Neural Networks. 51
MSE	Mean Square Error. 52
NAA	Nearest Neighbor Algorithm. 53
NASA	National Aeronautics and Space Administration. 67
NAVAIR	Naval Systems Air Command. 67
NUT	Network Under Test. 35, 49
OMTDR	Orthogonal Multi-tone Time Domain Reflectometry. 29
PC	Principal Components. 55
PCA	Principal Component Analysis. 62, 105
PDFDR	Phase Detection Frequency Domain Reflectometry. 27
PTFE	Polytetrafluoroethylene. 64
PWVT	Pseudo Wigner Ville Transform. 39
RF	Radio Frequency. 9
RMSE	Root Mean Square Error. 74
SACM	Self-Adaptive Correlation Method. 39, 40
SD	Speedy Delivery. 45
SMSW	Signature Magnification by Selective Windowing. 39, 40
SNR	Signal to Noise Ratio. 91
SSTDR	Spread Spectrum Time Domain Reflectometry. 29
STDR	Sequence Time Domain Reflectometry. 29
STP	Shielded Twisted Pair. 9
SVD	Singular Value Decomposition. 55
SWR	Standing Wave Reflectometry. 27
TDR	Time Domain Reflectometry. 3, 6, 27, 36, 62, 105
TFC	Time-Frequency Correlation. 39
TV	Television. 7
UTP	Unshielded Twisted Pair. 9
VNA	Vector Network Analyzer. 57
WVT	Wigner-Ville transform. 38

List of Symbols

ΔG_d	Fault conductance variation $(Siemens/m)$
ΔZ_d	Fault impedance variation (Ω)
α	Attenuation constant $(Neper/m)$
β	Phase constant $(radians/m)$
ΔC	Variation of Capacitance C $(Farad/m)$
ΔG	Variation of Conductance G $(Siemens/m)$
ΔL	Variation of Inductance L $(Henry/m)$
ΔR	Variation of Resistance R (Ω/m)
ΔZ_c	Impedance variation (Ω)
δ	Dirac pulse
δ	Resolution of the soft fault
Δ_v	The variation with respect to the $RLCG$ parameters of the cable under test
ϵ	Threshold of the probability of false alarm
ϵ_0	Vacuum permittivity (F/m)
ϵ_d	Threshold of the probability of detection
ϵ_r	Relative permittivity (F/m)
Γ	Reflection coefficient
γ	Propagation constant
Γ_E	Reflection coefficient at the input of the cable
Γ_l	Reflection coefficient at the end of the cable
Λ	Diagonal eigenvalues matrix
λ_j	Eigenvalues
μ	Permeability (H/m)
μ_0	Vacuum permeability (H/m)
ν_j	Noise vector

ω	Angular frequency (rad/s)
σ	Conductivity (S/m)
θ_f	Fault overture or angular cutaways (°)
$\mathbf{p_j}$	Orthonormal eigenvectors
$\mathbf{R}_{\theta°\mathbf{f}}$	Fault signature data vector with angular cutaways θ° and frequency f
$\mathbf{R_{Sj}}$	faulty TDR response corresponding the sensor **Sj**
$\mathbf{x_j^*}$	Variables of X^*
$\mathbf{x_j}$	Variables of X
$\mathbf{R_f}$	Column vector of the matrix X^* corresponds to the cable healthy TDR response at the frequency f
B_T	Bandwidth of the test signal
C	Capacitance $(Farad/m)$
c_α	Critical value of the normal distribution at α significance level
C_d	Capacitance of the fault $(Farad/m)$
CM	Covariance matrix of X_0
D	Internal diameter of the outer conductor
d	External diameter if the inner conductor
E_f	Fault energy at frequency f (J/Hz)
E_s	Energy spectral density (J/Hz)
F	Fault case ($l_f = 5mm$,$\theta_f = 180$°
f_c	Cut-off frequency
f_n	Selected frequency in the presence of noise
$F_{k,n-k,\alpha}$	Fisher–Snedecor distribution critical value
f_{max}	Maximal frequency
Fj	Fault case j
G	Conductance $(Siemens/m)$
$H(f)$	Cable transfer function
$h(t)$	Channel impulse response
I_m	Identity matrix

k	Number of the retained PC
L	Inductance ($Henry/m$)
l	Length of the model (m)
L_d	Inductance of the fault ($Henry/m$)
l_f	Fault length (m)
m	Variables number
n	Observations number
P	PCA loading matrix
P_1, P_2	Fault interfaces
p_ν	Noise power
p_s	Signal power
P_{nd}	Probability of detection calculated at frequency f_n
Q_α,T^2_α	Confidence limits of the Q andT^2 tests, respectively
R	Resistance (Ω/m)
R_d	Resistance of the fault (Ω/m)
S	Scattering parameters
$s(t)$	Reflected signal
T	Principal component score matrix
U_i	Gaussian pulse
v_p	Propagation velocity (m/s)
V_{mes}	Measured TDR response
V_{sim}	$RLCG$-based model TDR response
X	New measurements matrix
x, y, z	Axis
X^*	Original fault-free, noise-free, data matrix
X_0	Normalized data matrix
X_ν	Noisy data matrix
x_f	Fault position

Z_0 Input impedance (Ω)

Z_c characteristic impedance of a transmission line (Ω)

Z_d Fault impedance (Ω)

Z_g Generator impedance (Ω)

Z_l Load impedance (Ω)

Sj Sensor number j

List of Figures

List of Tables

General Introduction

Electrical cables are used in all industrial sectors to transfer energy or information. In the last few decades, there has been increasing electrification of several functions in industrial systems. This increase in the usage of electrical components was joined with an increase in the demand for electrical cables, e.g., in transportation systems, industrial machinery, power plants, infrastructures, in addition to most of nowadays' human facilities and utilities. Thus, the length and the complexity of wired networks have also grown. For example, the combined length of electrical wires in a modern car is more than 4 km compared to a few hundred meters 30 years ago, higher than 40 km in modern fighter aircraft, nearly 200 km in high-speed trains, and 400 km in recent civil planes. In a country with over 40000 km of railways, almost 1 million km of electrical cables are used for the infrastructures. Therefore, electrical wires play a critical role in networks that are becoming fundamental subsystems whose operation is crucial.

The cable in a network will at some point show signs of weakness due to either external causes, such as chemical contamination, mechanical vibration, moisture penetration, etc., or internal causes such as defective manufacturing, local heating, etc. Under these conditions, more or less severe faults can appear, from a simple crack in the sheath to a cable cut. Among the most frequent wiring faults are short and open circuits, usually referred to as hard faults, characterized by the fact that the functionality is lost. On the other side, any minor alteration that affects a cable is classified as a soft fault, namely insulation damage, cracks, etc. The latter can be of very different kinds and are more difficult to detect. At first glance, they seem harmless and without significant consequences for the system. However, cable aging, mechanical stresses, and hostile environments can evolve soft faults into hard ones.

There are electrical wires dedicated to safety and control operations, so any shortening in their performance due to a fault might be extremely costly, both in terms of human lives and economics. The National Transportation Safety Board (NTSB) investigation revealed that the Boeing 747 TWA Flight 800 disaster in 1996 and the Swissair MD-11 disaster in 1998 were both caused by electrical wiring faults, which resulted in hundreds of deaths. Many other incidents were noted, which did not result in catastrophic accidents, but were attributed to wiring failures such as Boeing 757 of AA (2008) and Airbus340 of VA (2009) [16]. Based on the data collected by the Air Force Safety Agency (AFSA) [17], cables' faults are the main cause of aircraft accidents, with 29%.

Thus, to ensure the reliable use of cables, it is necessary to detect faults that might jeopardize the whole system. Different methods have been developed to improve the reliability of wired networks to detect and locate certain types of faults in cables. Among these methods, we can distinguish the classical visual inspection, X-ray, capacitive and inductive measurements [18,19], and reflectometry techniques that are widely used and easily embedded. Even though several electric and non-electric wire diagnosis methods have been studied and developed throughout the last few decades [20, 21], reflectometry-based techniques are still in the center stage of research and industrial applications. They have been introduced since the middle of the twentieth century. Their general concept relies on the propagation of an electromagnetic waveform in

the wired network to be tested, followed by the analysis of the reflected signals to detect the presence, position, and nature of an impedance discontinuity possibly caused by a fault [21].

Depending on the analysis domain of the reflected wave, reflectometry methods can be categorized into two main families: Time Domain Reflectometry (TDR) [22] and Frequency Domain Reflectometry (FDR) [23]. They have provided effective results with hard faults due to their high reflection coefficients, but they have shown poorer performance whenever soft faults are addressed [24, 25]. As a matter of fact, soft faults that are usually characterized by small reflection coefficients, produce weak echoes compared to those caused by junctions within a network for example. In addition, the energy of the test signal may be significantly attenuated due to the presence of cable inhomogeneity, junctions, coupling, splices, etc., making the detection of soft faults (i.e., chafing, bending radius, pinching) more complex. Moreover, the soft fault detection is disturbed by environmental conditions such as vibration, high temperature, crosstalk, noise disturbance, etc. Additionally, the signature of such fault can be masked in complex networks because of multiple echoes in the reflected signal resulting from the presence of junctions and impedance discontinuities in the network.

As a solution, further development is needed to make the reflectometry method sensitive enough to detect and locate soft faults efficiently. In this context, several post-processing methods have been proposed [26–29]. However, these methods are prone to test signal attenuation and dispersion phenomena [30]. Indeed, the phenomena of attenuation and dispersion significantly reduce the location accuracy when the propagation distance is important [9]. Hence, the choice of the test signal bandwidth is critical and affects the diagnosis performance. Additionally, the maximal frequency of the reflectometry signal is a critical parameter in detecting and locating a fault in a cable. Indeed, the higher the frequency, the better the reflectogram's resolution and the localization accuracy of small faults. However, in the case of fault detection on long cables, increasing the signal frequency is not recommended as it introduces dispersion and increases signal attenuation [31, 32]. In practice, the expert configures and calibrates the Vector Network Analyzer (VNA) at a given frequency and records the healthy cable measurement. Measurements at the same frequency are then done on a faulty cable. Analysis of the measurements is established at this frequency on the computer. If the fault is not detected, this operation must be repeated. Therefore, there is a loss of information and time in addition to the subjectivity of the decision-making. Hence, more accurate and not-user dependent technique should be proposed to cope with this problem.

The problems associated with diagnosis, described previously, are aggravated in the case of a multi-branched network [33]. In such networks, using a single sensor may be no longer possible to monitor the whole network due to signal attenuation and connection complexity. As a solution, distributed reflectometry is used to overcome ambiguity problems and maximize the diagnosis coverage. It consists of performing reflectometry measurements at different extremities of the Network Under Test (NUT). However, the injection of multiple signals leads to computational complexities and sensors fusion problems [11]. Moreover, energy consumption is a major drawback regarding environmental constraints. The study on the reduction of the sensors' number in complex networks and its impact on the diagnosis quality is provided in [34]. However, it has revealed other challenges related to bandwidth allocation, communication protocol, and noise interference mitigation. Thus, in [35], the cable life profile is included (such as, environmental stress (temperature, vibration, moisture, etc.), cable type, cable age, channel

noise, diagnosis method, etc.), allowing to reduce the diagnosis cost by avoiding the use of too many sensors in the network. With this solution, the reliability of the sensors in emission and reception is considered in the obtained statistics. This reliability differs from a sensor to another and impacts the fault location [34]. In fact, the reflectometer's reliability in terms of injection or reception impacts the fault characterization. For example, in the case of an unreliable reflectometer, an erroneous test signal may be injected. Then, a false interpretation of the reflectogram is done involving unnecessary intervention. Hence, solutions for solving these problems are hot topics in wire diagnosis.

This thesis proposes to address these two main problems in soft fault wire diagnosis using reflectometry. Hence, we propose a particular way to combine statistical analysis with reflectometry. In this study, the principal component analysis (PCA) is considered the statistical tool to improve reflectometry's soft fault detection problem.
PCA is a multivariate data-driven statistical modeling technique [36]. It uses information redundancy in a high-dimensional correlated input space to project the original data set into a lower-dimensional subspace defined by the principal components (PCs).
Although the combination of reflectometry and PCA in the literature was proposed for an efficient reduction of the reflectometry data space [37], the application of PCA on reflectometry-based data has not been used, so far, for fault diagnosis in wired networks. For fault detection, the PCA model of the system is developed in our work, based on healthy operating system data, and then used to check new measurement data.

This work intends to improve the soft fault diagnosis in multi-branched networks using this PCA-based reflectometry method. Two main research questions are tackled. The first is related to the best frequency selection (chapter 3) to be used in the TDR construction, and the second one is related to the sensor selection (chapter 4) in multi-branched wiring networks using the distributed reflectometry. For both, the performances and limitations of the proposal are studied. The influence of disturbance noise is considered. The results are derived using simulated data and validated using an experimental bench.

On this basis, the organization of this dissertation into four chapters is as follows:

- Chapter 1 presents the context of this study, which is the on-board diagnostic of complex wired networks for the detection and localization of soft faults. To do this, we first present the different types of cables and their fields of application. Secondly, we introduce the typologies of faults, such as hard or soft faults. To better understand the principle of reflectometry, we recall the concepts of line theory. Finally, the reflectometry limitations and drawbacks are discussed considering the two main families proposed in the literature: Time Domain Reflectometry (TDR) and Frequency Domain Reflectometry (FDR). Highlighting some main limitations, we then discuss the main trends in the literature and introduce this work proposal to cope with some well-identified problems.

- Chapter 2 details the problematic of this work. It starts with presenting the soft fault detection and localization problems in wired networks using reflectometry methods. Then it discusses several post-processing methods proposed in the literature and their limitations concerning the test signal attenuation and dispersion phenomena, which are discussed in detail, and the frequency selection problem is introduced. Next, the challenges associated

with diagnosis in the case of a multi-branched network related to distributed reflectometry are explained, and the need to reduce the number of the sensors is established. Finally, the PCA for Fault Detection and Diagnosis (FDD) is introduced.

- Chapter 3 introduces a new approach for frequency selection in the case of soft fault diagnosis. The proposed method permits to configure and calibrate the VNA at different frequencies. It performs measurements on different frequencies for the healthy case. After which, the PCA model is established. It performs the new measurements at different frequencies. If a dissimilarity is detected between the model and the new data, the contribution of each variable (i.e., frequencies) to this variation is calculated. The algorithm allows then to choose the most relevant frequency to monitor the soft fault. The relevance of this choice is validated by studying the performance of noisy signals. The advantages of the proposal are thus time-saving and enabling automatic computerized decision-making.

- Chapter 4 introduces the new approach for selecting the relevant sensors to monitor and diagnose soft faults in multi-branched wired networks. It combines TDR distributed reflectometry with PCA. Indeed, for a given NUT, a distributed reflectometry approach is considered where the sensors perform their reflectometry measurements. These collected data are used to establish a PCA model coupled with statistical tests to evaluate new measurements status. Whenever a fault is detected, the relevant sensors for monitoring and diagnosis are identified with high accuracy. Based on these results, the sensor's number could be reduced, and the non-selected ones could be inactivated temporarily, reducing energy consumption, computing complexities, and sensor fusion problems. As for the previous chapter, the performance and limitations of the proposed technique are studied in a noisy environment and validated experimentally.

The manuscript ends with a section devoted to drawing the main conclusions and proposing several perspectives.

1 - Wire Diagnosis by Reflectometry

Contents

1.1 . Introduction

In the era of wireless networks, the use of electrical cables' remains inevitable in almost all modern systems. They play a primary role in energy and signal transportation, where wiring networks have become essential subsystems whose proper functioning is of critical importance. After installation, electrical cables age and may show signs of weakness, leading to the appearance of faults. They can also be exposed to aggressive environmental factors (humidity, heating, corrosion, etc.) or aggressive operating conditions (mechanical attacks). In this case, they lose their function and may cause serious flaws. Therefore, it is necessary to establish a diagnosis system to detect and locate faults in the cable. In this context, reflectometry is a promising method to meet this demand.

This chapter aims to present the problems encountered in electric cables and the different diagnosis methods used to detect and locate faults. After a brief introduction to the growing need for electrical cables in today's modern systems, the different types of cables and their domain of application is first introduced. Then, the various faults encountered that can pose significant safety issues with severe economic impacts are presented. According to their severity, we can distinguish two types of faults: hard and soft. Next, several methods designed to detect and locate cable faults are introduced where their advantages and limitations are discussed.

In the literature, lots of work dedicated to fault detection are based on the use of reflectometry [22, 23, 38, 39]. This technique uses the same principle as Radar. Its main idea consists of injecting a signal into the cable under study and letting it flow through it with respect to the physics propagation laws. When an impedance discontinuity occurs, part of the signal is back-propagated, and the analysis of this reflected signal can be used to characterize the fault. In our work, we consider this principle combined with statistical signal processing to propose an efficient fault detection and diagnosis solution.

In the following, the reflectometry technique is presented. Its limitations and drawbacks are discussed considering the two main families proposed in the literature: TDR and FDR. Highlighting some main limitations, we then discuss the main trends in the literature and introduce this work proposal to cope with some well-identified problems.

1.2 . Electrical Cables and their Applications

With the evolution of electrical systems, the use of electrical cables has become mandatory. Their existence is essential for energy and information transmission, ensuring the normal behavior and optimal system's performance.

According to [1], the global cable market size in 2018 is estimated to be 164.94 billion of US Dollars. It is expected to grow at a Compound Annual Growth Rate (CAGR) of 4.9% over the forecast period $(2019 - 2025)$. Figure 1.1 shows the global cables market share, by end-use, in 2018. High demand for these cables by end-users (such as aerospace & defense [22, 40], building & construction, oil & gas, energy & power [40], Information Technology (IT) & telecommunications and industrial machinery [41]) is expected to be a leading factor in this

market's growth over the forecast period.

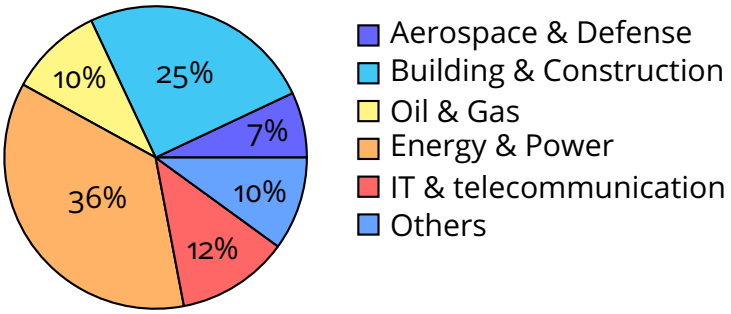


Figure 1.1 – Global wires and cables market share, by end-use, 2018 [1]

Huge attention on the widespread use of electrical wires had been noticed with the increasing complexity of modern systems such as automotive and avionics industries. For example, leisure and entertainment facilities have been added in today's vehicles (such as Air Conditioner (AC), radio, Television (TV), and navigation systems), which require wiring networks. Furthermore, a new concept has emerged in the automotive industry; "X-by-wire" technology. The idea is to replace the main mechanical and hydraulic control systems (such as motors and Anti-lock Braking System (ABS)) by electrical or electro-mechanical systems. This has the effect to considerably increase the number of on-board systems to be interconnected. This need for on-board electronics has evolved considerably in recent years, from 2% to 30% rising a car production cost between 1920 and 2000 and have reached 35%, or even 40% in 2015 [11]. This trend has strengthened, leading to increased in complexity and mass (from 1% to 5% of the weight of a vehicle) as well as the volume of electrical harnesses. In thirty years, the length of the cables carried on an automobile has more than tenfold, passing from nearly 200 to more than 4000 meters [17]. Today, it is common to have in a vehicle nearly 4 kilometers (km) of cables combined, weighing more than 80 kilograms (kg). At the same time, the number of connections has increased from 200 to over 2000. While the advent of "X-by-Wire" technology is a major revolution in the automotive industry, issues of cost, wiring, connectivity, reliability and control have arisen.

A typical cable network in a modern car shows bundles of long cables and their connectors, referred as the automotive wire harness (Figure 1.2). About 35% of the entire automobile infrastructure is related to electrical systems [42]. As the number of Electronic Control Units (ECU) and sensors distributed throughout the vehicle continues to grow, manufacturers have resorted to multiplexing in order to reduce the amount of cables. This technology consists of exchanging several information among the different components with a single physical medium (i.e. a bus) instead of passing each information through a dedicated wire. For this, we need to define with accuracy: the transmission medium (wires, twisted pair, coaxial cable, optical fiber, etc.), the types of information (power signals, control signals, high frequency signals, digital signals , etc.) and the communication protocol.

The Automotive industry is not the only industry in which wires' length has increased significantly. Figure 1.3 presents the cumulative lengths of cable in different applications, from 40 km in a civil helicopter and to $5,000$ km in a nuclear power plant. This number can reach

$40,000$ km in the railway infrastructure of wide countries. In new infrastructures, it is expected to be $1,000,000$ km long [9]. This increase has resulted in complexity making the maintenance operation very complex [22].

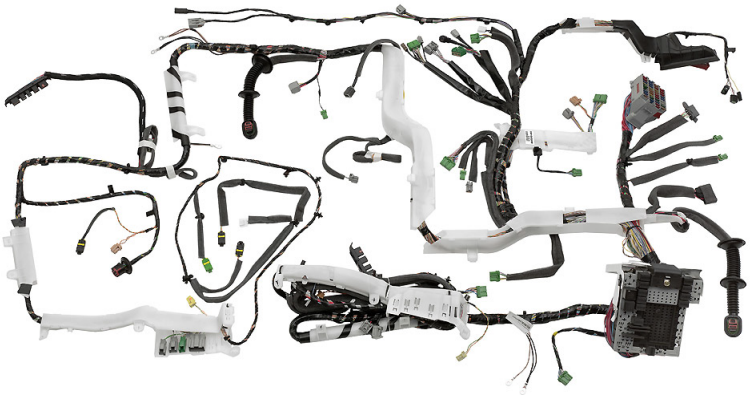


Figure 1.2 – A complete network of electrical cables of a typical modern car [2]

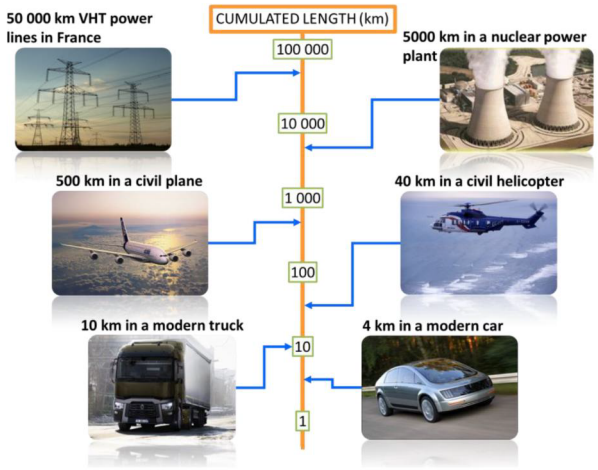


Figure 1.3 – Cumulative cable lengths in various applications [3]

In each of the above applications, the types of cable depend on the nature of the signal to be transmitted and the environment in which the system operates. The transmitted signal can be analog or digital with low or high power and propagating at low, medium, or high frequency [43]. Environment, on the other hand, varies widely. The cable could be in the air as for power transmission networks, in the sea or buried for underwater and underground connections [44].

To illustrate this latter purpose, Figure 1.4 presents an example of different types of electrical cables used in high-speed train. Medium and high-voltage cables are used for energy transmission, coaxial cables for high-frequency transmission systems (radio, radar, and data), low and high-temperature power cables to provide energy and deal with extreme temperatures and weight constraints. Moreover, two main types of cables are used for computer networks: coaxial cables [45] and twisted pair cables [46].

Based on this example, we can identify the three most common cables: coaxial, twisted pair, and power cables. Their main specificities are described in the following subsections.

1.2.1 . Coaxial Cable

Coaxial cables, invented in 1930 by Herman Affel [47], are perhaps the most common, basic, and easy to understand cables of common transmission line designs. It uses an inner conductor made up of a single copper strand or several twisted strands to conduct electrical power. An insulating dielectric material surrounds this core to avoid contact with the braided metal shield (operating as a ground). This shielding protects data against electromagnetic disturbance and ensures high-speed transmission over long distances [48]. Finally, the cable is often protected from the external environment by an insulating outer sheath or jacket, as shown in Figure 1.5. The term "coaxial" is derived from the fact that the core and the shield share the same geometric axis.

The advantage of the coaxial design is that the electric and magnetic fields are confined in the dielectric, and there is almost no leakage outside the shielding layer. Signals inside the cable are protected from electromagnetic interference with outside electric and magnetic fields. This feature makes coaxial cable an ideal choice for carrying high-power signals that are not allowed to radiate or couple to adjacent structures and weak signals that cannot tolerate environment interference [48].

The applications for this cable are mainly Radio Frequency (RF), microwave transmission, video cabling, CATV distribution, computer, and instrumentation data connections [49]. Generally, this cable's use is extended to any application domains where the signal must undergo minimal attenuation and distortion, such as telecommunications, aerospace, and military [49].

1.2.2 . Twisted Pair Cable

Introduced in 1881 by Alexander Graham Bell [50], twisted-pair cable, is made up of two twisted strands of copper (a forward and a return conductor of a single circuit). This twisting aims to maintain the distance between the wires and reduce cross-talk (or electromagnetic interference) between neighboring pairs due to nearby sources [51].

Twisted pair is used as a transmission medium for on-board local networks, such as the FlexRay network, new communication protocols for "X-by-wire" applications, and the Controller Area Network (CAN) used in automobiles.

Depending on the shielding, twisted pairs could be divided into several categories such as Unshielded Twisted Pair (UTP), Shielded Twisted Pair (STP), Foiled Twisted Pair (FTP) (Figure 1.6). UTP is often used in Ethernet networking and has no shield. On the contrary, for STP typically used in Token Ring networks, each twisted pair is surrounded by a conductive shielding layer giving more immunity against electromagnetic interference. Finally, for FTP, a common shielding sheath encloses all the twisted pairs. The latter cable is used in telephone and computer networks applications.

Figure 1.4 – A map showing different cable types with their function in the railway industry [2]

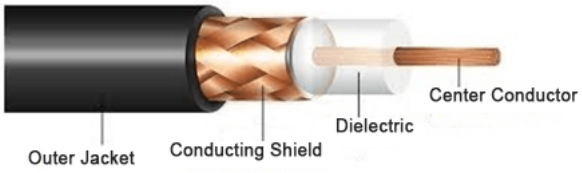


Figure 1.5 – Front cut of the coaxial cable showing its different layers [4]

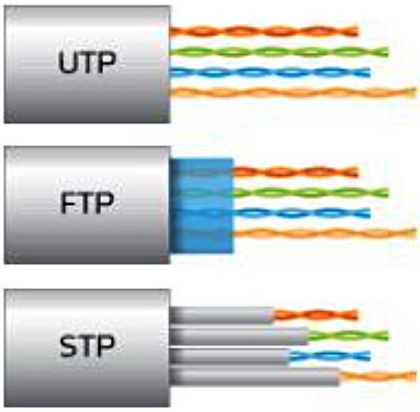


Figure 1.6 – Twisted-pair cables examples [5]

1.2.3 . Power Cable

The first power distribution system has been developed by Thomas Edison in 1882. These cables are frequently used at high voltage to transmit electrical power, whether in the Alternating current (AC) or Direct current (DC) mode. One of their characteristics is that the insulation should withstand high voltage stress.

Figure 1.7 displays a cross-section cut view of a high voltage cable, with its main layers. An inner conductor is rounded by an insulator. Then the insulator is covered by an external semiconductor with a metallic screen and finally with and outer protective sheath. During cable laying for a given application, the inner conductor should tolerate pulling stresses. The metallic screen provides mechanical protection for the cable during installation. The latter is insulated from the surrounding medium and protected from corrosion through the outer protective sheath made from a polymer material [2].

It is important to note that, in this work, coaxial cables are used through their numerical model to validate the proposed methodology but also physically in the conducted experimental measurements for proof of concept. Then the comparison could then be made between simulated data and measured ones. In fact, the selection of this kind of cable was motivated by the fact that they are very commonly used in a wide range of applications.

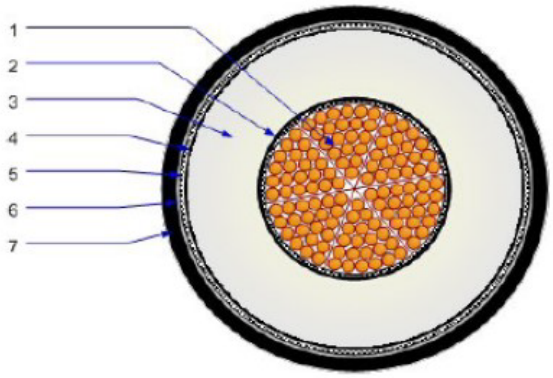


Figure 1.7 – A cross-section cut of a high-voltage power cable: (1): Conductor, (2): internal semi-conductor, (3): insulation, (4): external semi-conductor, (5): metallic screen and (6): outer protective sheath [6]

As shown from the above examples, the extensive use of wiring networks in today's systems has become crucial. Whatever the application domain, cables are victims of the environments they operate in and often face aggressive conditions (thermal, mechanical, electrical, and environmental). Hence, one day or another, they will show signs of weakness or aging, resulting in the appearance of anomalies, usually referred to as wiring faults. These anomalies may cause system dysfunctions and induce severe consequences, especially if these cables are part of a critical system where safety is an issue.

Therefore, in the next section, the main types of faults affecting a cable will be described, as well as the main reason for their appearance.

1.3 . Faults in Electrical Cables

For many years, cables have been considered as elements that could be installed and operates for the system's life [52]. However, this changes rapidly as signs of cable failure appear. These failures may cause more or less severe anomalies, such as loss of electrical signal, information distortion, power surge, system malfunction, smoke, fire, etc. According to the fields of application, the consequences of the faults or failures can be catastrophic or even fatal. For example, studies conducted in the avionics industry indicated that the crashes of both Boeing 747 of TWA 800 (1996) and the MD-11 of Swissair 111 (1998) were related to the faulty wiring systems [22]. Many other incidents were noted, which did not result in catastrophic accidents, but were attributed to wiring failures such as Boeing 757 of AA (2008) and Airbus340 of VA (2009) [16]. Based on the data collected by the AFSA [17], cables' faults are the main cause of aircraft accidents, with 29% followed by connectors' faults with 14%.

The maintenance of cables also comes with a significant cost. The US Navy, for example, spends 1.8 million person-hours per year troubleshooting and repairing aircraft wiring systems. Approximately 1077 missions aborts and $147,674$ non-mission-capable hours per year are lost due to cable incidents [53]. The delivery of Airbus' flagship aircraft A380 had been delayed due

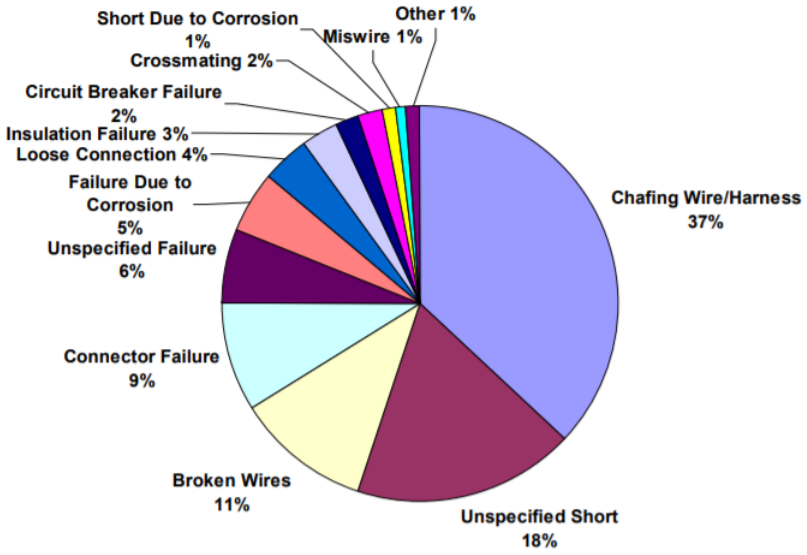


Figure 1.8 – A pie chart showing detailed types of faults in aeronautical cables from the data collected between 1980 and 1999 by the US NAVY [7]

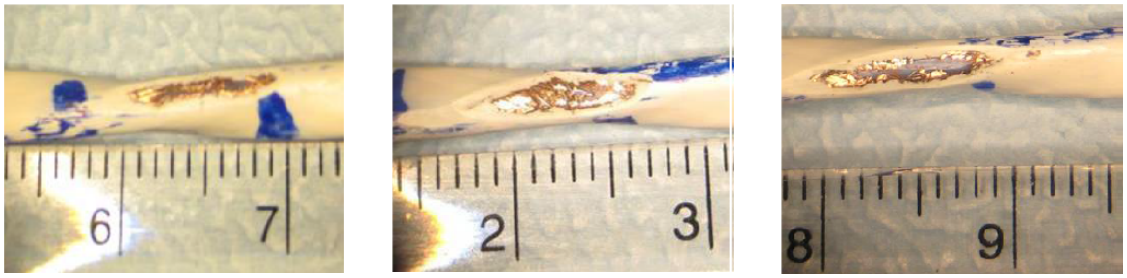
to wiring problems, resulting in $\$6$ billion extra cost [54].

In this context, assessment of the most prominent faults in electrical wiring had been made using programs of the United States Coast Guard, NAVAIR[1], NASA[2], and the FAA[3]. Even though each of the mentioned agencies has its own maintenance practices, data collection methods, and operating environments, they all experience similar cable problems. Figure 1.8 represents the data collected by the US NAVY concerning wiring faults discovered during maintenance operations between 1980 and 1999 [7]. It is noted that wire chafing represents roughly 37%, followed by short circuits with 18%, and open circuits with 11%. Figure 1.9 shows an example of an aircraft cable with progressive chaffing that might lead to catastrophic casualties if not early detected.

Generally, the leading causes of cable degradation and faults may be due to external or internal factors [55, 56]. External factors include chemical contamination, mechanical aggression (e.g., vibration), an incorrect application, maintenance operations (e.g., human interventions), corrosion due to humidity, and oxidation due to chemicals and environmental effects (e.g., temperature). Internal factors include insulation aging [57, 58], undetected manufacturing faults, and local heating (Figure 1.10).

These factors are responsible for modifications in the cable's parameters, resulting in signal propagation perturbation: the cable is considered faulty. According to their severity assessment, faults in cables are generally classified into two main families: hard faults and soft faults (Figure 1.11).

[1]Naval Air Systems Command
[2]National Aeronautics and Space Administration
[3]Federal Aviation Agency



(a) (b) (c)

Figure 1.9 – A progressive shield-chafing fault in an aircraft cable where (a) chafing 6Thousands (K) cycles beyond initial shield exposure, (b) chafing after 8K cycles, (c) chafing after 10K cycles. Ruler ticks indicate millimeters [7]

(a) (b) (c)

Figure 1.10 – Faults of different origins: (a) mechanical, (b) thermal, and (c) chemical [8]

(a) (b)

Figure 1.11 – Example of faults: (a) open circuit (hard fault), (b) insulation damage (soft fault) [8]

1.3.1 . Hard Faults

Among the most commonly encountered faults one can mention short circuits and open circuits, which are the source of many fires or signal losses [18, 59]. A short circuit results from the low resistance connection of two or more conductors due to insulation damage. On the other hand, due to the cable's mechanical damage or violent movement, one or more conductors rupture will lead to an open circuit. They are known as "hard faults". These faults cause considerable variation in the cable characteristic impedance and appear as a complete interruption of the flow of energy or information. Thus, it profoundly affects the performance of the system. Hard faults give rise to tragic accidents following sudden system malfunction.

1.3.2 . Soft Faults

Soft faults are characterized by a small impedance variation of the cable, making them difficult to detect. They do not prevent energy or information circulation, but the quality of its transmission will depend on the state of degradation of the cable. Consequently, these faults do not always result in a severe breakdown. In general, any fault that is not considered hard shall be defined as a soft one; these include crack of the sheath, insulation damage, frays, etc. However, the origins of the mentioned degradations, such as aging, mechanical stresses, and hostile environment, can make the faults progressive, leading to severe accidents. If they are not detected within a short time, they can evolve into hard faults.

Proper and efficient wire diagnosis methods that detect faults as quickly as possible, especially before they become critical, can guarantee "faithful" management of the system's proper functioning. Thus, the objective is to detect and locate faults with spatial precision to avoid unnecessary costly interventions. To address this challenge, industry and academia have started to study and develop smart fault detection, diagnosis, and prevention systems in electric cables. Then these studies help to better define condition based maintenance and then preserve in the maximum the remaining useful life of the cable.

It shall be important to point out here that the two main families of wiring faults, hard and soft faults, are basically permanent faults, that will appear continuously once they occur. With this in mind, there exist wiring faults that occur commonly in aircrafts, and referred to as intermittent arc faults (about 37%). These are the most frustrating, mysterious and extremely difficult faults to detect and locate because they can appear in a few milliseconds due to vibrations for example and then disappear [60].

In our work, we propose to mainly focus on the detection and diagnosis of soft faults.

1.4 . Cable Fault Detection and Localisation by Reflectometry

There is a need for new and practical wire diagnosis methods ensuring the reliable use of cables and enabling the detection and localization of faults. There are several wire-based diagnosis methods that have been developed.

These include visual inspection, X-ray, infrared thermal imaging, capacitance and inductive measurements, etc [20, 21]. Although visual inspection is the most intuitive technique that relies entirely on the primitive visual intervention of human sensibility, the increasing complexity of wired networks made it non-efficient. Since a large portion of the cable network are positioned in hard-to-reach places, generally hidden by massive structures like electric panels, components, or other cables [61], only 25% of the existing faults in an aircraft could be detected by this method [7]. The X-ray inspection method depends upon heavy equipment usage, needs direct cable access, and human intervention for post-acquisition data analysis leading to non-practicability in many cases. Infrared thermal imaging is not able to characterize all the fault types [18]. Capacitive and inductive methods are based on the cable's capacitance and inductance measurement. Even though they are robust and easy when used to diagnose a point-to-point cable [19], they are not suitable in many cases: complex network analysis, cable

	Visual inspection	X-Rays	Capacitive and inductive methods	Frequency domain reflectometry	Time domain reflectometry
Long cable (i.e., >30m)	☹	☹	●	☺	☺
Buried cable	☹	☹	●	☺	☺
Hard fault	☺	☺	☺	☺	☺
Soft fault	●	☺	☹	●	●
Intermittent fault	☹	☹	☹	●	●
Online Diagnosis	☹	●	☹	☺	☺
Complex networks	☹	☹	☹	☹	●

Table 1.1 – Comparison of diagnosis methods: ☺: the method detects the fault. ●: the method detects the fault under conditions. ☹: the method does not detect the fault [14]

in operation, soft faults.

Table 1.1 [14] presents the main advantages and disadvantages of these methods. Despite their limitations, these methods remain inexpensive and straightforward and can provide the location of hard faults. They are ideal for integration into handheld test equipment and are an easy-to-use alternative to manual search methods.

Although several electric and non-electric wire diagnosis methods have been developed, reflectometry-based techniques are the most promising. Reflectometry methods in the time and frequency domains make it possible to detect and locate hard faults due to their significant signatures, which may be detectable by measurement devices. However, these current methods are generally not effective for the detection of soft faults where there is only a slight variation in the properties of the cable [24, 25]. The conclusion drawn from the table is that reflectometry in its different methodologies seems the most efficient method to detect soft faults for most of the given combinations. This technology and its capabilities are detailed in the following.

1.4.1 . Principle of Reflectometry

Reflectometry is a high-frequency method commonly used for monitoring the condition of cables. It can provide information for the detection, location and characterization of electrical faults. It uses the principle of a radar (Figure 1.12) based on the propagation of electromagnetic waves within a system or medium to be diagnosed (in our case, a cable): a probe signal is injected at one end of the system to diagnose, this signal flows under the propagation laws of the current medium, and a part of its energy is reflected back to the injection point whenever an impedance discontinuity (junction or fault) is encountered (Figure 1.13). Analyzing the

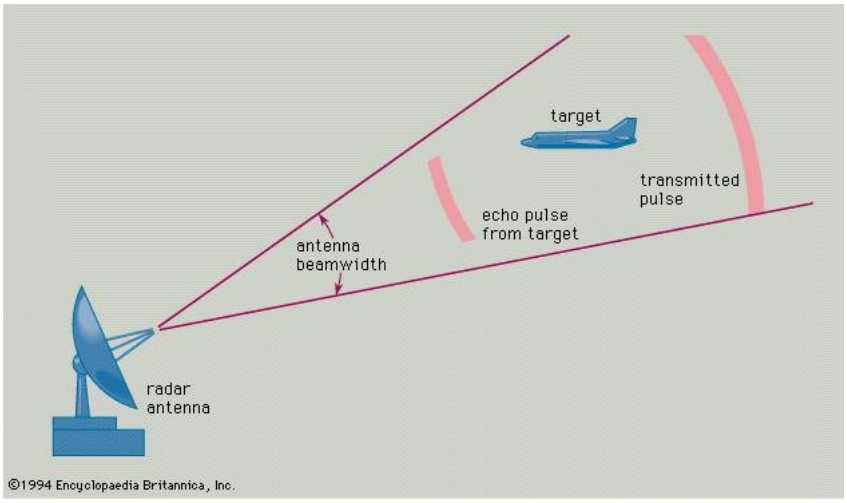
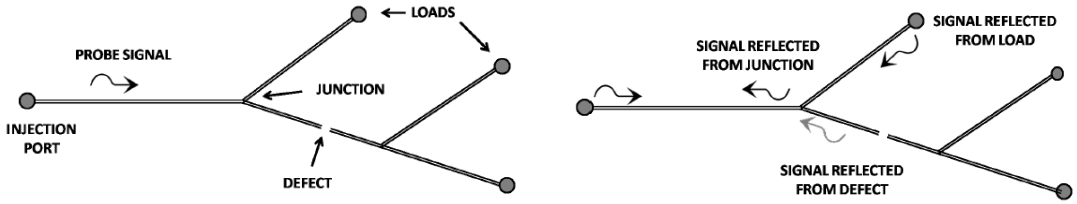


Figure 1.12 – Radar systems using reflectometry principle for locating targets



(a) (b)

Figure 1.13 – Principle of reflectometry [9]

reflected signal allows to have information about the presence, the position, and the type of discontinuity [21]. This knowledge is then very useful for maintenance operators.

Let us consider a transmission line, of length l, supplied at one end by a voltage generator of impedance Z_g and at the other end, closed by a load impedance Z_l as shown in Figure 1.14. At low frequency, the wavelength λ is greater than that of line l and the electromagnetic wave is almost constant at any point of the line, whatever the load impedance Z_l. Hence, it is unnecessary to introduce the propagation notion. On the other hand, at high frequency, since the wavelength λ is shorter than that of line l, the amplitude of the wave is no longer constant throughout the line. In this case, the propagation phenomenon must be taken into account. Reflectometry is inspired by this phenomenon to extract information on the condition of the transmission line.

In the case of a simple model of a transmission line, the channel response is approximated by a succession of Dirac pulses. Each pulse represents the echo of the incident signal returned by an impedance discontinuity. Each pulse is delayed by a propagation time τ_i, which corresponds to the time required to make the round trip between the injection point and the discontinuity i, and weighted by an attenuation coefficient α_i. The channel impulse response $h(t)$ is given

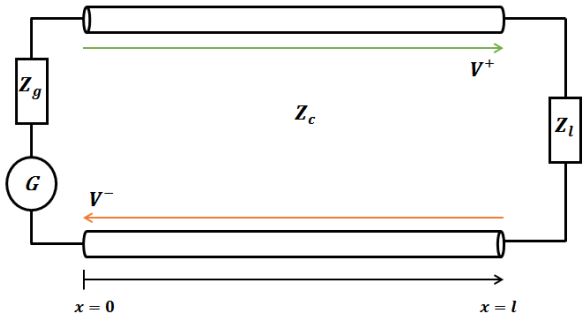


Figure 1.14 – Model of a transmission line of length l and characteristics impedance Z_c

by:

$$h(t) = \sum_i \alpha_i \delta(t - \tau_i) \tag{1.1}$$

The injection of a signal $e(t)$ at the input of the line with a channel response $h(t)$ gives rise to a reflected signal $s(t)$ expressed as follows:

$$s(t) = e(t) * h(t) = \sum_i \alpha_i e(t) * \delta(t - \tau_i) = \sum_i \alpha_i e(t - \tau_i) \tag{1.2}$$

The operator $(*)$ represents the inner product. The amplitude of the incident signal $e(t)$ is therefore attenuated by a factor α_i and delayed by τ_i at each discontinuity i. Analysis following echoes can detect, locate and characterize the nature of the discontinuity.

As reflectometry methods are based on the propagation of electrical signals in a transmission line and in order to better understand how it is applied to wire diagnosis, a preliminary reminder of the theory of transmission lines is necessary to understand the phenomenon of propagation of an electromagnetic wave in a transmission line. For this, the use of a physical model of the wired network and the numerical simulation of reflectometry are very useful.
Many authors [7, 9, 62, 63] rely on a 4-parameter model ($RLGC$) to describe the propagation of a wave (with associated voltage V and current I) inside a cable, from which come the well-known telegrapher's equations [64]. The $RLCG$ model of a transmission line is the subject of the next section.

1.4.1.1 RLCG Model of a Transmission Line

At high frequency, a transmission line can be modeled by the "$RLCG$ model" [65], which divides a transmission line of length l into several consecutive elementary segments of length dx. Each of them is represented by the equivalent electric model of Figure 1.15. It consists of the following parameters: resistance (R), inductance (L), capacitance (C) and conductance (G), often given as per unit length values. The $RLCG$ elements are referred to as the primary parameters of a propagation line and are defined linearly as follows:

- Linear resistance R expressed in (Ω/m), which depends on the electrical resistivity and the section of the transmission line.

- Linear inductance L expressed in $(Henry/m)$, which breaks down into two parts: an interior inductance and an exterior inductance. The inner inductance is due to the magnetic field inside the conductors and the external inductance is due to the magnetic field between conductors. The linear inductance depends on the diameter of the conductors, the distance between the two conductors and the permeability of the materials.
- Linear capacitance C expressed in $(Farad/m)$, which depends on the permittivity of the insulation placed between the two conductors.
- Linear conductance G expressed in $(Siemens/m)$, which represents a leakage current flowing between two conductors when the insulation separating these two conductors is not perfect.

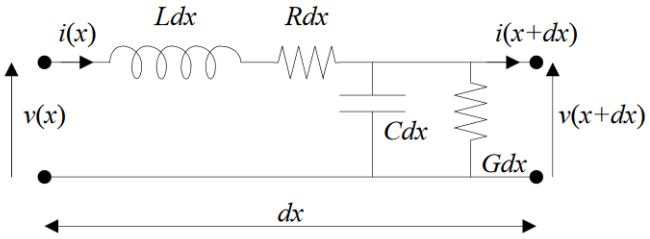


Figure 1.15 – Diagram of the $RLCG$ model equivalent to a segment of a transmission line

Reflectometry is inspired by the phenomenon of reflection to ensure the diagnosis of cables. In what follows, the process to find the reflection coefficient $\Gamma(x)$ is described.
The application of Kirchhoff's laws provides the following equations:

$$v\left(x+dx,t\right)-v\left(x,t\right)=-Rdxi(x,t)-Ldx\frac{\partial i(x,t)}{\partial t} \tag{1.3}$$

$$i\left(x+dx,t\right)-i\left(x,t\right)=-Gdxv(x,t)-Cdx\frac{\partial v(x,t)}{\partial t} \tag{1.4}$$

where v and i represent the voltage and the current, respectively, at the instant t in the segment of the line of length dx. By applying the partial derivative, we obtain the two differential equations, known as the *Telegraphists equations* [66], as follows:

$$\frac{\partial v(x,t)}{\partial x}=-Ri(x,t)-L\frac{\partial i(x,t)}{\partial t} \tag{1.5}$$

$$\frac{\partial i(x,t)}{\partial x}=-Gv(x,t)-C\frac{\partial v(x,t)}{\partial t} \tag{1.6}$$

We propose to solve (1.5) and (1.6) in harmonic regime, by considering that the line is subjected to a sine wave of pulsation $\omega=2\pi f(rad/s)$ where f is the frequency [67]. In this case, the current and voltage waves are given, respectively, by:

$$v(x,\omega,t)=V(x,\omega)e^{j\omega t} \tag{1.7}$$

$$i(x,\omega,t)=I(x,\omega)e^{j\omega t} \tag{1.8}$$

where $V(x,\omega)$ and $I(x,\omega)$ represent the complex amplitudes associated with the voltage $v(x,t)$ and the current $i(x,t)$ respectively. By replacing $v(x,t)$ and $i(x,t)$ in (1.5) and (1.6) by their respective complex expressions, we get:

$$\frac{\partial V(x,t)}{\partial x} = -\left(R + j\omega L\right) I(x,t) \tag{1.9}$$

$$\frac{\partial I(x,t)}{\partial x} = -\left(G + j\omega C\right) V(x,t) \tag{1.10}$$

By deriving (1.10) with respect to x, we can write:

$$\frac{\partial^2 I(x,t)}{\partial x^2} = -\left(G + j\omega C\right) \frac{\partial V(x,t)}{\partial x} \tag{1.11}$$

By injecting (1.9) into (1.11), we find (1.12):

$$\frac{\partial^2 I(x,t)}{\partial x^2} = \left(G + j\omega C\right)\left(R + j\omega L\right) I(x,t) \tag{1.12}$$

In the same way, we can obtain (1.13):

$$\frac{\partial^2 V(x,t)}{\partial x^2} = \left(G + j\omega C\right)\left(R + j\omega L\right) V(x,t) \tag{1.13}$$

We thus obtain the following two propagation equations:

$$\frac{\partial^2 I(x,t)}{\partial x^2} - \gamma^2 I(x,t) = 0 \tag{1.14}$$

$$\frac{\partial^2 V(x,t)}{\partial x^2} - \gamma^2 V(x,t) = 0 \tag{1.15}$$

Where γ, the propagation constant, is defined as follows:

$$\gamma = \sqrt{\left(G + j\omega C\right)\left(R + j\omega L\right)} = \alpha + j\beta \tag{1.16}$$

with α, the attenuation constant $(Neper/m)$ and β, the phase constant $(radians/m)$. In (1.16), the first term α represents the attenuation of the amplitude of the wave during its propagation while the second term $j\beta$ represents the rotation of the phase of this wave. Solving these equations in the case of a harmonic regime, gives the following expressions:

$$V(x,t) = V^{+}e^{-\gamma x} + V^{-}e^{\gamma x} \tag{1.17}$$

$$I(x,t) = I^{+}e^{-\gamma x} + I^{-}e^{\gamma x} \tag{1.18}$$

Equation (1.17) shows that there are two waves: a traveling wave $V^{+}e^{-\gamma x}$, which propagates towards the load impedance Z_l (in the positive direction x) and a reflected wave $V^{-}e^{\gamma x}$, which propagates towards the generator (in the negative direction of x) after a reflection at the level of the load impedance Z_l as shown in Figure 1.15. The voltage $V(x,t)$ and current $I(x,t)$ waves are connected by a characteristic impedance Z_c expressed as follows:

$$Z_c = V(x,t)/I(x,t) = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{1.19}$$

In the case of a lossless transmission line, $R = 0$ and $G = 0$. With these approximations, the propagation constant becomes:

$$\gamma = j\omega\sqrt{LC} = j\beta \tag{1.20}$$

Note that the attenuation constant (α) is zero. Regarding the characteristic impedance, its expression becomes:

$$Z_c = \sqrt{\frac{L}{C}} \tag{1.21}$$

The presence of a reflected wave in (1.17) can be explained by the presence of a discontinuity in the characteristics of the line or if the load impedance is different from the characteristic impedance at the end of the line ($Z_c \neq Z_l$). In addition, the variation in the intrinsic characteristics of the line following a fault can also lead to the appearance of reflections. The reflection coefficient $\Gamma(x)$ is defined by the ratio of a reflected wave to an incident wave as follows:

$$\Gamma(x) = \frac{V_{reflected}}{V_{incident}} = \frac{V^- e^{\gamma x}}{V^+ e^{-\gamma x}} = \frac{Z_l - Z_c}{Z_l + Z_c} \tag{1.22}$$

Reflectometry is inspired by this phenomenon of reflection to ensure the diagnosis of cables. In the case of a hard fault (open circuit or short circuit), it is very easy to determine the nature of the discontinuity according to the value of the load impedance Z_l.

- If $|Z_l| = 0\Omega$, then the incident wave encounters a short circuit and is reflected with a coefficient $\Gamma(l) = -1$. The reflected wave has the same amplitude as the incident one but with an opposite sign.
- If $|Z_l| = \infty$, then the incident wave encounters an open circuit and is reflected with a coefficient $\Gamma(l) = 1$. The reflected wave has the same amplitude and the same sign as the incident one.
- If $|Z_l| = |Z_c|$, then the incident wave will never be reflected ($\Gamma(l) = 0$). In this case, it is considered that the end of the line is *adapted*.

Reflectometry makes it possible not only to determine the nature of the discontinuity but also to locate it according to the following relation:

$$d = \frac{\tau v_p}{2} \tag{1.23}$$

where τ is the time required to travel the line (round-trip); v_p, the speed of propagation in the line. In the case of a lossless line, it is given as follows:

$$\beta = \omega\sqrt{LC} \Rightarrow v_p = \frac{\omega}{\beta} = \frac{1}{\sqrt{LC}} \tag{1.24}$$

As the knowledge of L and C is not guaranteed, v_p is often given by the cable manufacturer. R2RLCG method [10] is an iterative method to go back to the $RLCG$ parameters of a uniform cable section from a Time Domain Reflectometry (TDR) or Frequency Domain Reflectometry (FDR) measurement. There are methods described by some articles [68–71] to extract $RLCG$ from S parameters measurements but mainly on a thinner frequency band than R2RLCG method. This one was developed for its main advantages:

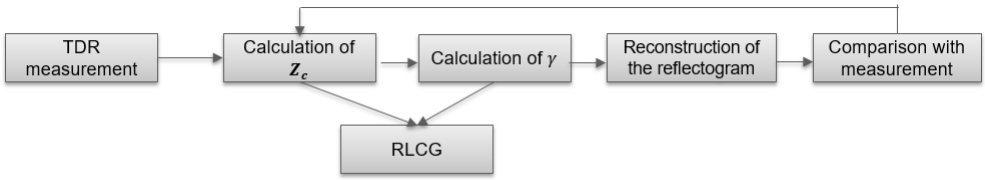


Figure 1.16 – Presentation of the method R2RLCG [10]

- Measure at one point of the cable
- Rapidity of calculus

It is based on the following prerequisites:

- Knowledge of the cable length noted l
- Knowledge of the input impedance noted Z_0
- Absence of strong reflections on the cable
- Knowledge of the load at the end of the line noted Z_l

This method is also based on another assumption as to the shape of the desired $RLCG$ parameters. Conventionally, for standard cables, it is possible to show a dependence in the square root of the frequency for R and a linear dependence of G with respect to the frequency. L and C are almost constant over the reflectometry working frequency band. This choice has consequences on the resulting parameters. The attenuation has a frequency dependence which is the sum of a term proportional to the root of the frequency and a term proportional to the frequency [10]. The phase factor is taken as proportional to the frequency, which is often the case for the frequency band considered. Likewise, a consequence of these choices is that the real part of the characteristic impedance is almost constant over the entire frequency band. This method will provide the $RLCG$ parameters over the entire frequency band and, consequently, the attenuation, the speed, and the complex characteristic impedance with high precision.

The detailed operation of this method is given in Figure 1.16. From the TDR measurement, the first estimation of the characteristic impedance of the cable Z_c is estimated based on the following equation:

$$Z_{c1} = Z_0 \frac{1+\Gamma_{E1}}{1-\Gamma_{E1}} \tag{1.25}$$

where $Z_0 = 50\Omega$ and Γ_{E1} is the given reflection coefficient at the input of the cable. The calculation of Z_{c1} and given Z_l permits to calculate an estimation of the reflection coefficient at the end of the cable Γ_l. Consequently, it is possible to make a first estimation of the propagation constant γ_1, such that:

$$\gamma_1 = \alpha + j\beta \tag{1.26}$$

where α is the cable attenuation and β is the phase constant. It is certain that these approximations do not permit to obtain the exact results. Whenever the first estimate of γ is

calculated, a corrective phase based on the assumed form of alpha(α) and beta (β) is applied. As a reminder, it is assumed that [10]:

$$\alpha(f) = a_0\sqrt{f} + a_1 f \tag{1.27}$$

$$\beta(f) = b_0 f \tag{1.28}$$

By regressing the entire frequency band using the least squares criterion, a_0, a_1, and b_0 are extracted. Once the parameters α, β, and Z_{c1} are obtained, a first estimate of the $RLCG$ parameters is calculated. As shown in Figure 1.16, these parameters are used to make a reconstruction of the reflectogram using the transfer function of the cable expressed as follows:

$$H(f) = \Gamma_E + \frac{(1-{\Gamma_E}^2)\Gamma_l \mathrm{e}^{-2\gamma l}}{1+\Gamma_E\Gamma_l \mathrm{e}^{-2\gamma l}} \tag{1.29}$$

The reconstructed reflectogram is then compared to the measurement. The parameters are also used to make a new estimation of the characteristic impedance Z_c, and a repetition of the previous calculation with the new values is made (Usually, the calculus is done for 40 times to cover a good range of the convergence, but in fact when there is a convergence, it is done in two or three iterations). The algorithm uses several iterations set by the user. At the output, the parameters are kept, which made it possible to obtain the best likelihood coefficient between the reconstructed and the measured reflectogram. Figure 1.17 shows the evolution of the correlation factor when applying the R2RLCG method on a coaxial cable type RG59 of 5 meters length. It shows that the correlation coefficient is quite good with one iteration but it becomes better with some more (even though the benefit is small) and starts regressing after too much operation [10].

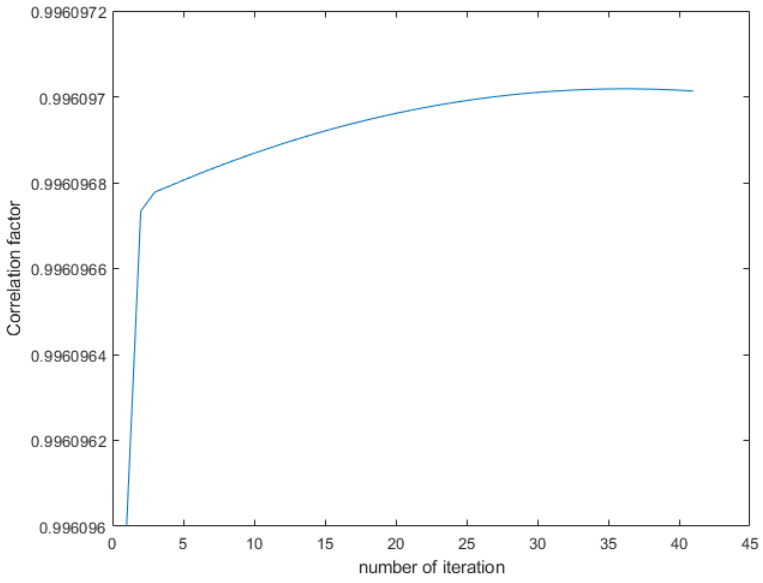


Figure 1.17 – Evolution of the correlation factor for R2RLCG on an example of coaxial cable

In the case of a soft fault, a very low reflection is resulted compared to the case of a hard fault. Indeed, only a part of the wave is reflected to the injection point while the other

part continues to propagate in the line. According to [12], impedance discontinuities greater than 10% are relatively easy to identify and locate just by looking at the response or using relatively simple algorithms to detect the fault's response automatically. Impedance differences below 10% become progressively more difficult to identify, as their response is much smaller. Eventually, the peaks from the reflection are smaller than the measurement error and cannot be detected.

The $RLCG$ model makes it possible to determine the primary $\{R,\ L,\ C$ and $G\}$ and secondary $\{Z_c, \gamma\}$ parameters of a transmission line. To better understand the operating mode of reflectometry, a frequency model is used in this work to simulate the presence of a hard or soft fault in a simple topology or a complex one (Y network, star network, etc.).

There are several methods that can be used to model wire networks [72–74]. Among these methods we have identified two particularly interesting methods: the ABCD matrix in the frequency domain and the finite difference method in the time domain (FDTD).

1.4.1.2 Time Domain Analysis

Time-domain simulation allows to follow the reflectometry signals as they travel along the wires of the network. In this work, we have employed a commercial microwave simulator: CST (Computer Simulations Technology) Microwave Studio for the 3D Electromagnetic (EM) model of a coaxial cable. This software performs a high fidelity simulation of Maxwell's equations using the Finite-Difference Time-Domain (FDTD) method [75]. The simulator allows for detailed 3D geometric modeling enabling accurate removal of shapes of shielding. It also enables arbitrary wave-shapes to be induced and the corresponding reflections and transmissions to be measured. FDTD method provides a quick result, as it only requires a 1D mesh of the wire [76]. FDTD simulation can help understand the propagation of signals along the wire [75] or in a bundle [77]. FDTD techniques work in the time domain, which makes them suitable for transient analysis. In particular, the simulation of TDR equipment is a natural implementation in an FDTD simulation.

1.4.1.3 Simulation of a Soft Fault by the Chain Matrix

Several methods have been developed to model wired networks. The most common method consists of using the matrix of parameters S, called the distribution matrix or "Scattering matrix" [78, 79]. Although the distribution matrix is efficient for the simulation of a hard fault in the network, the formalism of the ABCD matrix [80, 81] is preferable in the case of a branched network. Indeed, the ABCD matrix makes it possible to model a branched network by a simple cascade of ABCD matrices corresponding to the different branches which constitute it. The ABCD formalism makes it possible to take into account soft faults. It allows to simulate the local variations of one of the parameters R, L, C or G. To take into account three-dimensional radiation phenomena (couplings, cross-talk, etc.), "full-wave" approaches such as finite differences [82] or finite volumes [83] can also be considered. In this work, the simulation of a soft fault by the ABCD matrix is used.

The ABCD matrix is only defined in the case of a two port network as shown in Figure 1.18. It is used to connect the inputs of a circuit to the outputs as follows:

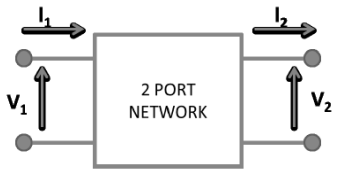


Figure 1.18 – Notations for the ABCD matrix formalism

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix} \tag{1.30}$$

V_1 and I_1 represent the voltage and current at port 1, respectively. For port 2, V_2 and I_2 also represent the voltage and the current. For a transmission line, of length l, characteristic impedance Z_c and propagation constant γ, the matrix ABCD is defined as:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh \gamma l & Z_c \sinh \gamma l \\ \left(\frac{1}{Z_c}\right) \sinh \gamma l & \cosh \gamma l \end{bmatrix} \tag{1.31}$$

The value of the ABCD matrix comes from the fact that one can reproduce complex structures with only basic circuits. In the case of a short-circuit at the end of the line (i.e. $V_2 = 0$), its equivalent impedance is obtained by:

$$Z_{sc} = \frac{B}{D} \tag{1.32}$$

In the case of an open circuit at the end of the line (i.e. $I_2 = 0$), its equivalent impedance is obtained by:

$$Z_{oc} = \frac{A}{C} \tag{1.33}$$

Using the ABCD matrix formalism in the frequency domain, we can derive a simple model for a soft fault. We assume that a soft fault, at the distance l_1 and over the length L_f of a point-to-point line of length $l \gg L_f$. To position a soft fault on a homogeneous line of length $l = l_1 + L_f + l_2$, it suffices to cascade 3 sections: a faultless section of length l_1, the fault of length L_f and another faultless section of length l_2. The two faultless sections are defined by primary parameters R, L, C and G or secondary parameters Z_c and v_p. The soft fault is defined as a local variation of these parameters. The matrices of these sections are written according to (1.31). They are then multiplied to obtain the matrix ABCD equivalent to the line:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{l_1} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{L_f} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{l_2} \tag{1.34}$$

Equation (1.31) for the faulty portion could be simplified because of the assumption $\mid \gamma_d L_f \mid \ll 1$ to:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}_{L_f} = \begin{bmatrix} 1 & Z_d \gamma_d L_f \\ \frac{\gamma_d L_f}{Z_d} & 1 \end{bmatrix} \tag{1.35}$$

Let $\epsilon = \frac{Z_d}{Z_c} - 1$, the amplitude of the soft fault signature is $\gamma_d L_f \epsilon$, which is very small and it explains why such a fault is difficult to detect using reflectometry, as verified experimentally in [84].

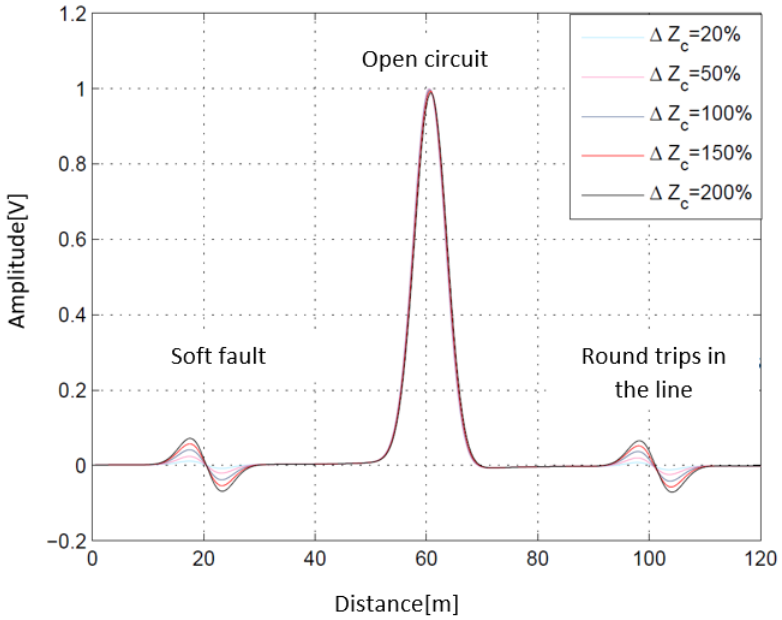


Figure 1.19 – Reflectogram of a transmission line with a soft fault and an open circuit [11]

Figure 1.19 shows a lossless transmission line of length $60m$ and characteristic impedance $Z_c = 100\Omega$. The line has a soft fault at a distance $19.5m$ from the injection point and an open-circuit (hard fault) at its end. The amplitude of the soft fault is very small compared to that of the hard fault, which makes its detection and localization difficult under certain conditions. The amplitude of the soft fault increases with the increase in the variation of the characteristic impedance of the cable.

In the case of a more complex network such as a Y-shaped network, the ABCD matrix is of great interest for the detection of a soft fault. In fact, the ABCD matrix makes it possible to take into account changes in the line by cascading the matrices of the different sections and thus facilitates the calculation of the reflection coefficient. For example, we consider a Y network made up of three branches B_1, B_2 and B_3 of lengths $l_1 = 50m$, $l_2 = 65m$ and $l_3 = 110m$. A soft fault of $50cm$ length situated at a distance $70m$ from the injection point is present on branch B_2. Figure 1.20 shows the corresponding reflectogram. The first peak corresponds to the first junction. Then, there is the weak amplitude peak, which corresponds to a soft fault. Then, we observe the two peaks due to the open circuits at the end of the lines B_2 and B_3. The other peaks are round trips of the incident signal in the network until its attenuation. The simulation window in this case is set at $350m$.

In practice, wired networks are more complex than a simple generic form (transmission line, Y network and star network). To make the interpretations more complex, we can extend the network in Y into a more complex network with n ($n > 2$) lines. In this case, reflectometry gives rise to complex signals that are very difficult to analyze and the reflectogram analysis becomes more and more complex, especially in the case of the presence of a fault in one or more lines of the network. Reflectrometry therefore requires certain computational power and a good mastery to draw useful conclusions. For this, several reflectometry methods have been proposed in order

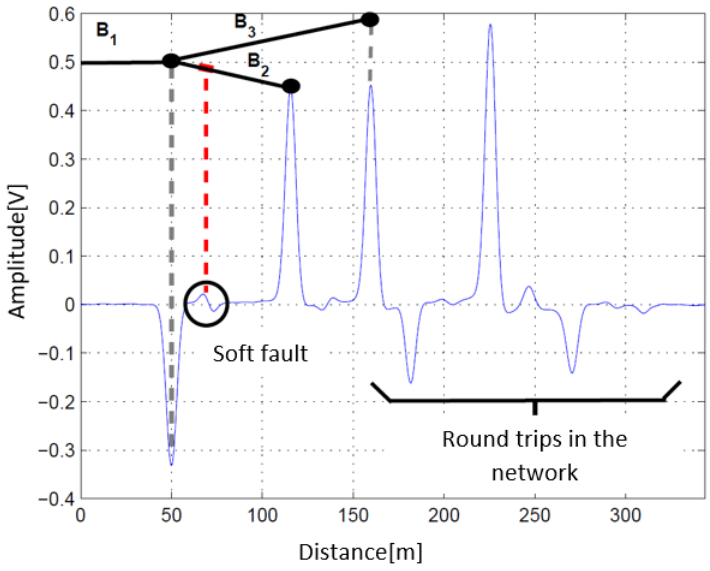


Figure 1.20 – Variation of the impedance of the soft fault [11]

to overcome the difficulties encountered by reflectometry in certain circumstances. Some of these methods are presented in the next section.

1.4.2 . Introduction of the Reflectometry Methods

The difference between reflectometry methods lies in the nature of the injected signal and the reflected one. Two main families are distinguished: FDR and TDR.

1.4.2.1 Reflectometry based on the Frequency-Domain Analysis

FDR injects a set of sine waves whose frequency varies linearly over time [23]. This signal is known as the "chirp" signal $c(t)$. The following equations define this signal:

$$c(t) = A\cos(\theta(t) + \phi) \tag{1.36}$$

$$\theta(t) = \int_0^t \pi f(u)du \tag{1.37}$$

$$f(t) = F_{min} + \frac{F_{max} - F_{min}}{t_{max}}\ t, \quad 0 \leq t \leq t_{max} \tag{1.38}$$

This signal is characterized by its amplitude A, its phase ϕ, and its frequency $\theta\,(t)$ [59] as shown in Figure 1.21.

To locate the impedance discontinuity, the analysis of one of the chirp signal characteristics (amplitude, frequency, and phase) is made, resulting in three variants of the FDR method: Standing Wave Reflectometry (SWR), Frequency Modulated Continuous Wave (FMCW), and Phase Detection Frequency Domain Reflectometry (PDFDR). SWR measures maxima and nulls in the standing wave caused by the superposition of the injected and the reflected signals

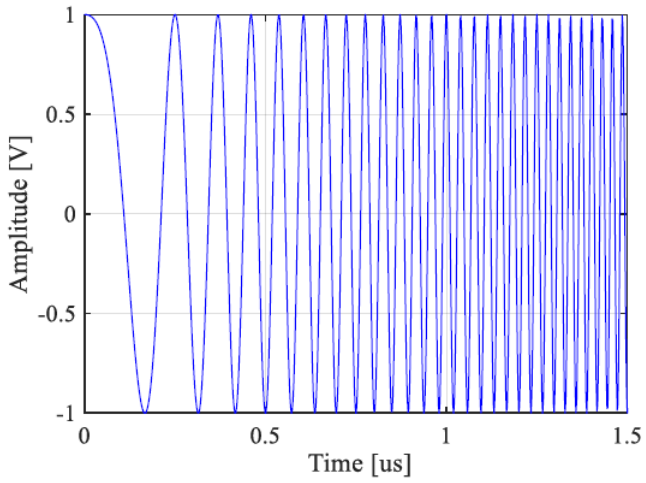


Figure 1.21 – Chirp signal for a linear frequency FDR implementation

to provide location information of the impedance discontinuity [85]. FMCW measures the frequency shift ΔF between the incident and the reflected signals delayed in time by Δt by varying the sine wave quickly, in a ramp function. With the knowledge of the wave propagation velocity v_p, the estimation of the fault location r is made as in (1.39) and (1.40) where the bandwidth of the injected signal is δF and T is the signal's period [85]. PDFDR [86] measures the phase shift instead of the frequency one. Reflectometry methods in-depth study is provided in [33].

$$\Delta F = \frac{\Delta t}{T}\delta F \tag{1.39}$$

$$r = \frac{v_p T \Delta F}{2\delta F} \tag{1.40}$$

The reflectometry methods in the frequency domain have the advantage of allowing a spectral analysis, and of exploiting the frequency behavior of the network for which the processing such as calibration and compensation operations (encountered in VNA) are immediate. In addition, they can be applied on cables under operation [33]. However, as FDR analyzes the standing wave caused by the injected and reflected signals' superposition, it is more difficult to understand and interpret [59] and it requires advanced processing techniques. For simple point-to-point cables, this analysis is relatively easy. However, it becomes too complicated for complex networks due to the difficulty of analyzing the interaction of a significant number of shifted waves [9]. In addition, since the final analysis of the fault is done in the time domain, complex processing of frequency-time transitions is required.

1.4.2.2 Reflectometry based on the Time-Domain Analysis

TDR method is the most well known reflectometry method used for fault detection and location [22]. It is based on the injection of a specific probe signal in the form of a Gaussian pulse or voltage step. The measured signal at the injection point mainly consists of multiple copies of this signal delayed in time. The delay is the round trip time between the injection point and

the discontinuity encountered for each copy. This composite signal is called a "reflectogram". Accordingly, the obtained reflectogram analysis can locate the discontinuity, provided that the propagation speed is known. Besides, the shape and the amplitude of this signal helps to identify the nature of the discontinuity. When the emitted signal is a Gaussian function centered at zero, it is written in the following form:

$$g(t) = A \exp\left(-\frac{1}{2}\left(\frac{t}{\sigma}\right)^2\right) \tag{1.41}$$

where A, the amplitude of the Gaussian and σ, the spreading factor. It defines the width at half height of the Gaussian pulse equal to $2\sqrt{2\ln(2)\sigma}$. Note that the simulations of the reflectometry response that were done in section 1.4.1.3 used a Gaussian pulse. This explains the presence of Gaussian form peaks in the reflectogram.

The propagation of the TDR pulse at a high frequency (up to 200 MHz, sometimes higher than $1GHz$) in a cable can undergo two phenomena: dispersion and attenuation [30]. Signal attenuates mainly due to the cable resistance. It is observed by the reduction in the amplitude of the signal during propagation. Consequently, it limits the diagnosis performance in terms of cable length. Dispersion results in impulse deformation and spreading. It is explained by the variation of the propagation constant, therefore of the propagation velocity, which depends on the frequency: for a test signal with extended frequency support (Gaussian or rectangular pulse, for example), its low-frequency components propagate slower than the high-frequency ones. This leads to the location accuracy reduction, especially in long-distance cables [9]. Dispersion and attenuation phenomenons will be explained in details in chapter 2.

Furthermore, the energy of the pulse signal of standard TDR is limited by its duration T. So, to deal with the pulse attenuation problem as it propagates, one needs to increase the signal's energy. Notably, this can be accomplished by using Pulse compression [87]. It is done by injecting a pseudo-random binary signal [88], then calculating the correlation between the injected and the measured signal to get the reflectogram. Linking TDR and spread spectrum techniques, known in the digital communication field, results in new reflectometry methods using the pseudo-random binary sequences [89] as a testing signal. These new methods are called Sequence Time Domain Reflectometry (STDR) [38] and Spread Spectrum Time Domain Reflectometry (SSTDR) [39, 61]. These two methods have disadvantages related to spectrum flexibility lacking, precision loss, blind zone presence, and they are limited to hard faults [90].

For a cable under operation, the basic time-domain reflectometry method is not recommended, as the power of the injected signal can disrupt or damage electronic systems connected to the cable [91]. So, to deal with this problem several diagnostic methods which are derived from TDR have been developed, such as STDR, SSTDR, Multi-Carrier Time Domain Reflectometry (MCTDR) [92], Orthogonal Multi-tone Time Domain Reflectometry (OMTDR) [11], Chaos Time Domain Reflectometry (CTDR) [93, 94] and Binary Time Domain Reflectometry (BTDR) [95, 96]. These methods are presented in detail in [90]. The MCTDR suffers from the presence of side-lobes in the case of cancellation of certain sub-carriers. In addition, OMTDR suffers from the presence of a blind zone around each peak with post-processing and filtering issues. On the other hand, the side lobes constitute the CTDR/BTDR methods' limits and can become troublesome in the detection and localization of hard and soft faults. Future work on the robustness of the methods to noise and the reduction of interference noise between the

sensors is necessary to overcome these limitations and make the detection of soft faults more efficient.

TDR methods are very suitable for diagnosing complex topology networks. They allow constructing a reflectogram where each peak is associated with a discontinuity present on the network [9]. The considerable development and improvement of electronic components during the last decade have simplified the implementation of TDR-based methods. These methods are very well suited to detect and locate hard and soft faults. In addition, the construction of a reflectogram in the time domain makes it possible to easily locate and identify the type of faults, where each echo is associated with a fault present in the cable. In addition, time-domain measurements can be much faster than frequency methods. Besides, over the last decade, a massive improvement of electronic components (Field-Programmable Gate Array (FPGA) and Digital Signal Processor (DSP), high bandwidth converters, etc.) has been observed. This made it simpler to implement TDR based methods. For this reason, methods derived from Time Domain Reflectometry are generally preferred for cable diagnostics, especially when on-board. Consequently, in this work, we will focus on the time domain reflectometry methods.

1.4.3 . Performances of Reflectometry

The performance of a diagnosis method is evaluated by its ability to detect and locate faults accurately. Once the fault is detected, a maintenance operator must access the faulty cable to repair or replace it. This kind of intervention can be quite tricky, especially in complex systems, such as avionics, where thousands of kilometers of cables are hidden under heavy structures. In addition, it requires the deployment of multiple resources (time, money, personnel, equipment, etc.). Therefore, to avoid unnecessary intervention, it is essential to have accurate information about the fault location. This underscores the fact that the reflectometry-based troubleshooting tool must be linked to or used in conjunction with a Computer-Aided Design (CAD) model of the wiring harness inside the system because wires may take complicated paths.

The reflectometry method has been proven to effectively allow the detection and location of faults in wired networks, even in complex topologies (Y network, CAN bus, etc.). Nonetheless, they face certain constraints inherent to the method itself or the physics. For instance: measurement noise, location ambiguity, blind zone, propagation inhomogeneity, interference problems [12, 14, 17] etc. The problems of soft fault detection and localization in wiring networks using reflectometry is elaborated in chapter 2.

Most TDR methods described above face location accuracy problem that manifests with increasing cable length [30]. Indeed, long cables suffer from attenuation and dispersion undergone by the test signal during its propagation, significantly reducing the location accuracy [9]. These two phenomena will be detailed in chapter 2. In this context, the choice of the test signal frequency presents a real challenge and is of high importance. In fact, reflectogram peaks must be very sharp to provide an accurate estimation of their location. However, sharp peaks require high-frequency usage (in standard TDR) or high data rate signals (for STDR based methods). In our work, we propose to address the frequency selection problem and present in chapter 3 the proposed methodology and the related results.

Another limitation is due to the fact that faults close to a connector or junctions are hard

to locate accurately. Furthermore, if the TDR system is not matched to the cable's intrinsic impedance, a part of the injected signal energy is reflected to the injection point without even going inside the cable. This leads to the appearance of a peak at the beginning of the reflectogram, which can mask the presence of another fault close to the injection point (blind zone). Hence, it is essential to either refine the methods by a specific choice of the test signal or make measurements at different network extremities. This is referred to as "distributed reflectometry" [17]. This method is explained in details in chapter 2. It solves the location ambiguity and the blind zone problems. Nevertheless, it reveals many constraints related to diagnosis cost, interference problems and sensor fusion problems, etc. In this context, chapter 4 will deal with the sensor selection for distributed reflectometry-based soft fault detection where the most relevant sensors for monitoring and diagnosing soft faults occurred in the network are identified with high accuracy.

What's more, is that the signal knowledge of the propagation speed v_p can influence the location accuracy. A TDR system measures the round trip time τ and then calculates the distance to the fault using equation (1.23). Any uncertainty on the value of v_p results in accuracy reduction on the fault location. It has been shown earlier that v_p depends on the frequency. Thus, if the test signal bandwidth is quite large, v_p is not possible to be defined accurately. Besides, v_p depends on the condition of the cable. For example, the aging of the cable can cause the propagation speed to vary [97]. Other parameters can influence the accuracy of locating a fault. This will be further elaborated in chapter 2.

1.5 . Conclusion

This chapter presents an overview of the different types of electrical cables that have been widely used in almost all modern systems. In many employed networks today, their cumulative length may exceed hundreds of kilometers, increasing networks' complexity. Moreover, cables' exposure to various aggressive stresses results in faults that make them non-operational. Two major families of faults have been distinguished: hard faults (open or short circuits) and soft faults. Soft faults feature minor changes, such as insulation wearing. In this case, signals can still propagate along the cable. However, a partial wearing may evolve into more severe degradation and eventually becomes a hard fault. This chapter underlined some faults that could have tragic consequences when cables are part of critical systems for which safety is an issue. Therefore, as an early-warning approach to ensure critical infrastructures' safe operation, it is essential to develop robust methods for identifying and locating faults.

In this context, two main groups for fault detection and location methods have been highlighted: non-reflectometry methods and reflectometry-based ones. This chapter has described the propagation phenomena in transmission lines to better understand how reflectometry works and the concrete implementation of fault detection (the RLCG Model of a Transmission Line). It has been seen that reflectometry-based methods are the most suitable for this job and can be distinguished according to the reflected signals' analysis domain. Precisely, we have seen time-domain (TDR) and frequency domain (FDR) reflectometry methods.

Ultimately, reflectometry-based-methods performance evaluation has been elaborated. Al-

though these methods show great performance, they witness several limitations. They almost shared an inability to provide efficient results with the increased network complexity. Indeed, the difficulty of analyzing the reflectometry response increases with the network's complexity involving the presence of one or several junctions, discontinuities, mismatched loads, etc. This leads to the degradation of the localization precision and consequently the quality of the diagnosis. In the case of soft faults, this problem will be more serious, and the fault may be hidden in various echoes of the signal.

Based on this chapter presentation, we have underlined that the main issues we will focus on concern soft fault diagnosis using time-domain reflectometry. For that, the next chapter will be devoted to presenting soft fault detection and localization problems in wiring networks using reflectometry methods. Then a throughout study about the existing methods in the literature for their detection and their limitations is presented. After which, the problematic of this thesis work and the raised research questions is elaborated.

2 - Soft Fault Diagnosis in Multi-branched Network Typologies: Methods and Limitations

Contents

2.1 . Introduction

Wire diagnosis addresses the problem of electrical fault detection, localization, and characterization. Various techniques are used to resolve this issue. However, the lights are spotted widely on the Time Domain Reflectometry (TDR) discussed in chapter 1, which uses a fast rise time pulse and deals with most industry's wiring issues.

However, the energy of the test signal may be significantly attenuated due to the presence of cable in-homogeneity, junctions, coupling, splices, etc. This attenuation makes the detection of electrical faults complex. This complexity increases in the presence of soft faults (i.e., chafing, bending radius, pinching). Indeed, soft faults are characterized by a small impedance variation leading to a low amplitude signature on the reflectogram. Moreover, fault detection is made even more difficult in the presence of nuisances as vibration, high temperature, crosstalk, etc. In this chapter, section 2 presents soft fault detection and localization problems in wiring networks using reflectometry methods. As a solution, further development is needed to make the reflectometry method sensitive enough to detect and locate soft faults. In this context, several post-processing methods have been proposed and will be discussed. However, these methods are prone to test signal attenuation and dispersion phenomena. Hence, these two phenomena will also be discussed in detail.

The problems associated with diagnosis, described previously, are aggravated in the case of a multi-branched network. In such networks, using a single sensor may no longer be relevant to cover the whole network due to signal attenuation and connection complexity. As a solution, distributed reflectometry method is used to remove ambiguities and improve the localization. It consists in performing reflectometry measurements at different extremities of the Network Under Test (NUT). However, the injection of multiple signals leads to computational complexities and sensor fusion problems. Moreover, energy consumption is a major drawback regarding environmental constraints. This subject is addressed in section 4.

To overcome the constraints in sections 3 and 4 and enhance the detection and location of soft faults in multi-branched topologies, section 4 presents the enhanced methods that have been used. These tools combined with distributed reflectometry allow automatic detection, location, and characterization of several faults for different types and topologies of wired networks. However, each of them has its limitations.

Therefore, in section 6, the principal component analysis (PCA) is considered the statistical tool to improve reflectometry's soft fault detection problem. Although the combination of reflectometry and PCA in the literature was proposed for an efficient reduction of the reflectometry data space [37], the application of PCA on reflectometry-based data has not been used, so far, for fault detection and diagnosis in wired networks.

Finally, section 7 presents the problematic of this thesis and the main research questions. We are then paving the way to the methods proposed in chapters 3 and 4.

2.2 . Reflectometry Acquisition System

When a cable encounters operational problems, the maintenance operator must locate the abnormality and repair it. For this, all he has to do is to connect the diagnosis system to the faulty cable. Therefore, it can indicate the location and nature of the possible fault. This type of diagnosis is called "off-board diagnosis" and it is of low cost. However, the main challenge is to provide helpful information for maintenance operators. This is not always guaranteed because the reflectogram complexity depends not only on the network's topology but also on the fault characteristics. Indeed, the measured signal can be very complex in the case of multi-branched networks due to the interconnections and the numerous round trips. The complexity is exacerbated when there are multiple faults. This may lead the maintenance operator to misinterpret the reflectogram, thereby reducing the performance of the diagnosis.

On-board diagnosis involves integrating diagnosis functions into the local environment of the wired network. It, however, reveals severe limitations in the cost, scale, and complexity of diagnosis. In addition, ease of integration and processing complexity are two main criteria to be considered when designing an on-board diagnosis system. Not to forget, compliance with standards related to the target system on which the wired network operates is a significant issue.

TDR is a suitable diagnosis solution for an on-board application. On the one hand, the required analog electronics are simple, and the measurement is fast. On the other hand, the final fault analysis is done in the time domain. The architecture of the TDR diagnosis system is mainly composed of three modules: the generation of the test signal, the signal acquisition at a sampling frequency F_s respecting Shannon's theorem ($F_s \geq 2F_{max}$), where F_s is the sampling rate and F_{max} is the maximum-frequency component of the analog signal to be sampled, and the processing module to build the reflectogram, which is analyzed for the monitoring.

In Figure 2.1, the fault detection process can be summarized as follows: the signal generator emits the test signal (e.g., a step-like pulse) that is sent through a coupler in the transmission line under test where **R** is the load. The slightest discontinuity encountered by the incident wave creates a reflected wave and whenever the reflected wave returns to the injection point, the oscilloscope will record its amplitude variation over time.

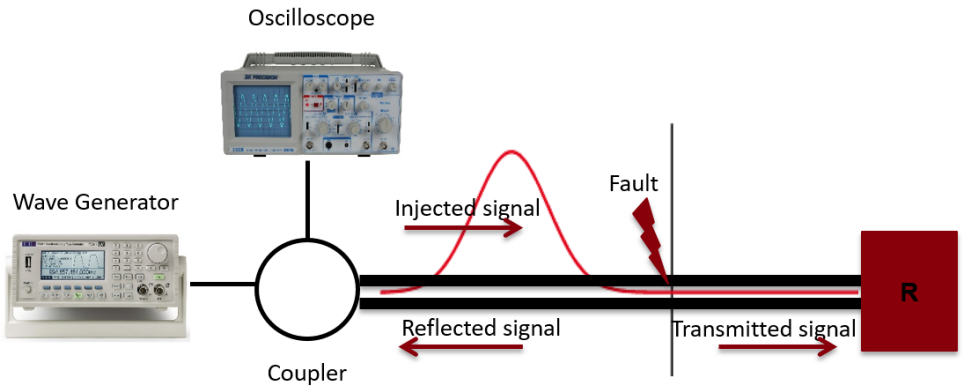


Figure 2.1 – Basic architecture of a TDR application

The processing phase consists of various operations carried out on the reflected signals to estimate the impulse response of the network $h(t)$ defined in (1.1) and extract the fault

information from it. It also includes signal processing to improve the quality of the measurements to facilitate the reflectogram analysis. For example, in the presence of noise, digital filtering could be applied to eliminate artifacts from acquisition devices, Electromagnetic Compatibility (EMC), crosstalk, loss of information, etc. Deconvolution can also be applied to retrieve the network's transfer function, which includes information on the detection and location of the fault. Although deconvolution improves localization accuracy, it is complex in multi-branched networks and can provide unstable results in noisy network [21]. Finally, the extraction of weak signals, such as those caused by soft faults, requires appropriate processing [98].

At present, the TDR and FDR reflectometry methods enable hard faults to be detected relatively well. However, several challenges are imposed in the presence of soft faults, as underlined in [12]. Why such a difficulty? A soft fault results in a very low characteristic impedance variation and therefore a very low reflection coefficient. Therefore, a signal reflecting on such degradation will be of very low amplitude. It can therefore be easily drowned in noise or masked by its proximity to another pulse of greater amplitude. Moreover, depending on whether the cable is stationary or vibrating, the reflection on the fault will be more or less detectable.

In the following, the problems of soft fault detection and localization using reflectometry methods will be discussed.

2.3 . Problems of Soft Fault Detection and Localization in Wiring Networks using Reflectometry

The data in section 2.2 are displayed on a control screen for interpretation, and the accuracy of the obtained results is linked to several factors such as the frequency of the incident signal where F_s plays an essential role in the fault location accuracy:

$$\Delta_x = \frac{\nu_p}{2F_s} \tag{2.1}$$

Where Δ_x is the variation of the position around the current fault location, and v_p is the speed of propagation in the cable. For example, for a propagation speed equal to $2 * 10^8 m/s$, the sampling frequency of $1GHz$ can give an accuracy of almost $10cm$. Other factors are the precision of time measurement, signal sampling and digitization, and the measurement or propagation noise.

In addition to electrical noise in the system, other sources of nuisances can limit practical reflectometry applications. Reflections are caused by impedance variation in the system such as mechanical vibration, temperature, moisture (e.g., being partially immersed in water), connecting and reconnecting wires, moving them around, etc. These factors can change the local impedance of cables [12]. These changes will cause additional, often time-varying reflections that increase perturbations in the reflectograms.

It is observed that noises due to vibrations or wire movement is an important factor in the capacity of soft fault detection based on reflectometry. Figure 2.2 gives a direct comparison of the peak values for different forms of discontinuities. For example, the amplitude of the peaks corresponding to frays are lower than the noise level. Hence, they may be masked by the noise,

making them undetectable. Only the peaks due to hard faults such as open or short circuits are significant, and can be definitively located using TDR.

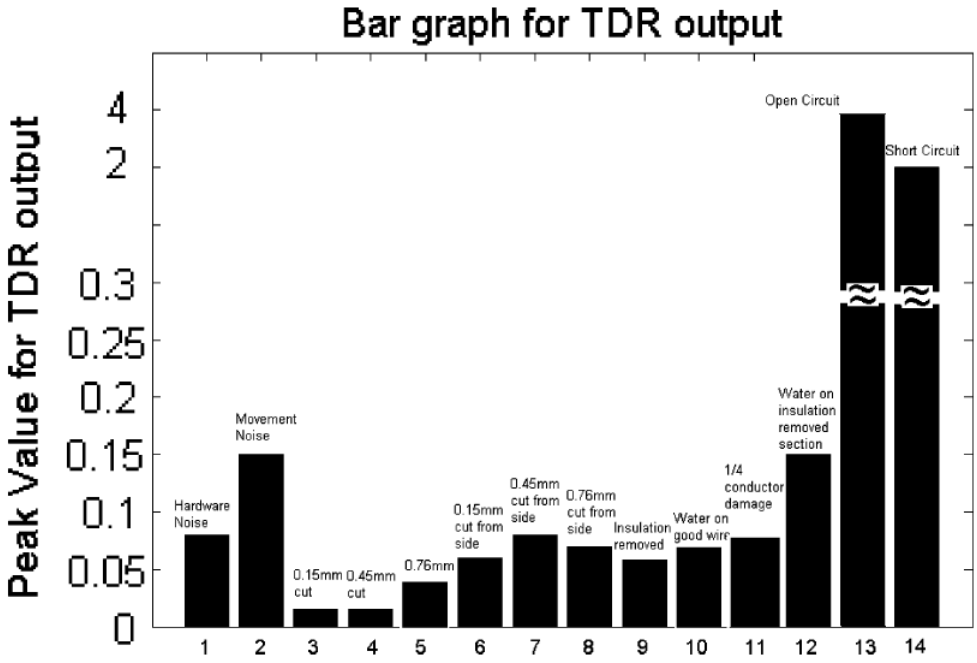


Figure 2.2 – Peak values for different frays. (1) Hardware noise; (2) movement noise; (3) 0.15 mm cut from top; (4) 0.45 mm cut from top; (5) 0.76 mm cut from top; (6) 0.15 mm cut from side; (7) 0.45 mm cut from side; (8) 0.75 mm cut from side; (9) insulation removed from single side; (10) water on good wire; (11) 1/4 conductor damaged; (12) water on cable with insulation removed; (13) open circuit; (14) short circuit [12]

According to [12], impedance discontinuities greater than 10% are relatively easy to identify and locate just by looking at the response or using relatively simple algorithms. Impedance differences below 10% become progressively more difficult to identify, as their response is much smaller. Eventually, the peaks from the reflection are smaller than the measurement error and cannot be detected. Reflections for damaged insulation on electrical cables are virtually invisible from the original reflectometry signature. Locating these types of faults requires the use of baselines and more advanced signal processing tools.

2.3.1 . Post processing Methods

Several signal processing methods have been proposed for improving the detection and location of soft faults. It is shown that the use of signal processing tools, such as the application of a time-frequency cross-correlation function using the Wigner-Ville transform (WVT). It is referred to as Joint Time Frequency Domain Reflectrometry (JTFDR) [26]. According to the Gabor-Heisenberg inequality, a signal cannot be perfectly localized both in time and frequency. Thus, in TDR as in FDR, a compromise must be found between temporal and frequency

resolution [98]. The JTFDR has been developed to allow more flexibility and to adapt to the characteristics of the cable to be diagnosed [99], [100]. This is a three-step process:

Step 1: Injection of a chirp signal in frequency, defined in such a way as to minimize the distortions during its propagation in the cable.

Step 2: Measurement of the reflectogram and calculation of its Wigner Ville transform.

Step 3: Application of a normalized time-frequency inter-correlation function, the peaks of which correspond to the discontinuities of the cable.

Authors in [101] and [102] have shown that it allows to detect soft faults in a coaxial cable. However, limitations related to the complexity and the false alarms caused by cross-terms generated by the quadratic form of the Wigner Ville transform are encountered. In order to reduce or cancel the cross-terms, authors in [27] focused on the false alarm problem by improving the time-frequency method. The method is called the Cluster Time Frequency Domain Reflectometry (CTFDR). It is made up of 3 major steps:

Step 1: TDR measurements on all the wires of the bundle.

Step 2: Application of the Time-Frequency Correlation (TFC) on the TDR results.

Step 3: Clustering of the TFC results.

Its purpose is to take advantage of the additional information given by near end crosstalk signals in order to ease soft fault detection. As the pulses reflected on such faults are of very low amplitudes, they may not be visible enough in TDR results. To emphasize them, another tool is used: the TFC, which is a normalized time frequency cross correlation. The TFC is computed after applying the Pseudo Wigner Ville Transform (PWVT) on TDR results. The PWVT was preferred to the WVT to mitigate the cross-term that can mask the real peaks or generate false positives. The last step is to apply a clustering method on the data obtained for all the lines. Its goal is to take benefit of all the available information, including near end crosstalk signals. By doing so, the risk of "missing" the presence of a fault in the bundle is decreased.

Despite the effectiveness of these algorithms, they are complex and complicated in terms of calculation time, especially for test signals with more than 1000 samples (more than 7 minutes for a computer with 15.9 GB of memory). Hence, the Self-Adaptive Correlation Method (SACM) [28] is used to amplify the signatures of soft faults and make them more easily detectable. It amplifies any signature correlated with the injection pulse. SACM has made an important gain on the amplitude of the soft fault signature. However, the gain decreases with noise when the robustness of this method was tested by adding White Gaussian Noise to the reflectogram. It has been concluded that this gain is sensitive to noise presence on the reflectogram. A Signature Magnification by Selective Windowing (SMSW) method is proposed in [29] to select the critical zone based on a predetermined window. The performance of the SMSW however depends strongly on the window width. Although those methods seem promising to locate soft faults, they are prone to false alarms since they are able to confuse the signature of

soft fault with other inhomogeneities in the cable. As a solution, an innovative approach based on advanced post-processing data fusion methods is introduced.

In [13], a fusion approach of several post-processing results is proposed where a probabilistic model is developed and used to detect the soft fault. The proposed approach described in Figure 2.3 includes several steps. After the reflectograms construction, a difference between

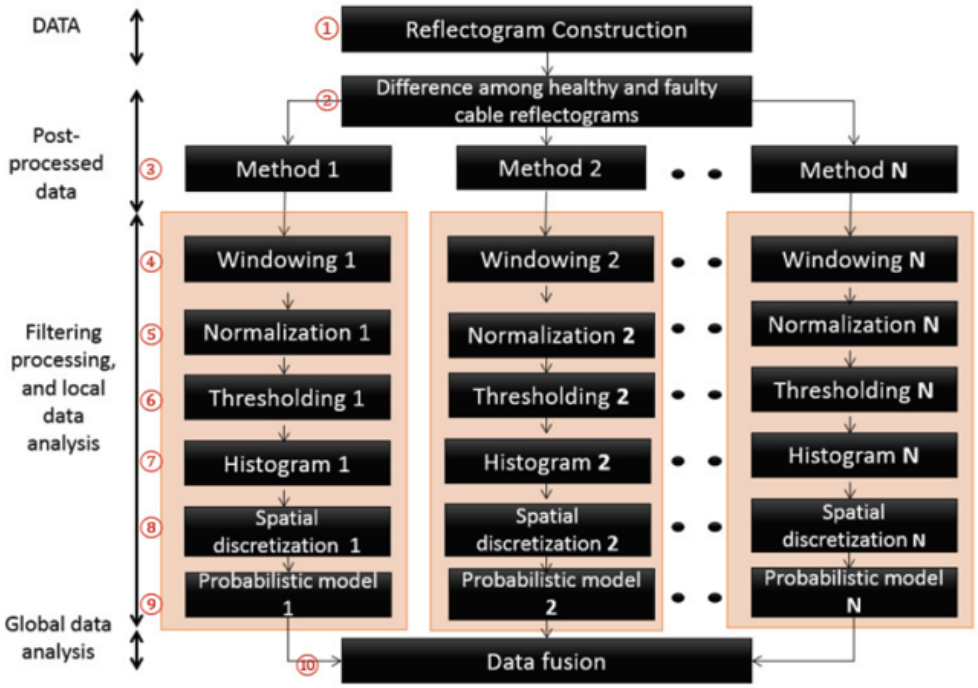


Figure 2.3 – Data fusion algorithm [13]

healthy and faulty cable reflectograms is performed to eliminate inhomogeneities related to cable manufacturing, installation, etc. Here, several peaks are present in the reflectogram, leading to diagnosis ambiguity. Hence, a post-processing method is called in step 3, such as the SACM, the SMSW, etc. After that, windowing is performed on the post-processed reflectogram to eliminate the peaks related to the impedance mismatch present at the cable extremities. Since the post-processing methods are heterogeneous, their results are adapted to make them consistent for further data fusion. To do so, a normalization step is performed with respect to the optimum values of the remaining samples. In step 6, a dynamic threshold is updated. Then, step 7 converts each reflectogram obtained at step 5 into a signal where the amplitude of each sample represents the percentage of satisfied thresholds. In fact, it is noticed that each post-processing method may introduce a slight delay. As a solution, a spatial discretization is called in step 8, where the cable length is divided into short sections based on predetermined spatial intervals, and the amplitude of the samples of the same interval are then summed. In step 9, the signal is converted into a measure of the probability of fault presence for each section. The probabilities close to 1 indicate the fault presence. The probability of 0.5 means that the section's two states (faulty and healthy) are equally likely. Finally, measured probabilities of fault presence according to post-processing methods are gathered, under the assumption that the selected post-processing methods are sequentially performed, and are entirely independent.

Table 2.1 – Summary table of some post-processing methods

Method	Advantage(s)	Inconvenience(s)
Joint Time Frequency Domain Reflectrometry (JTFDR)	- Allows more flexibility - Allows detecting soft faults in a coaxial cable	-Complexity -False alarms
Cluster Time Frequency Domain Reflectometry (CTFDR)	Mitigate the cross-term that can mask the real peaks or generate false positives	-Complexity -Calculation time
Self-Adaptive Correlation Method (SACM)	Made an important gain on the amplitude of the soft fault signature	-Gain sensitivity to noise
Signature Magnification by Selective Windowing (SMSW)	-Don't need to increase the power or the bandwidth of the injection signal -Robust to noise	-Performance depends strongly on the window width
Fusion Algorithm	- Eliminates false alarms -Automated (decision making is based on a probability of detection)	-The performance depends on the weight affected to each post-processing algorithm

The detailed mathematical calculations are found in [13]. The result of this method for a $30m$ shielded twisted pair TWINLINK 50 FA with a shield fault of $-8mm$ long and $3mm$ wide present at $10.9m$ from the injection point, shows the efficiency of the developed demonstrator to detect and locate the soft fault with high accuracy (1%).

Table 2.1 summarizes the post-processing methods above, with their advantages and inconveniences. Although several interesting methods have been proposed to enhance soft fault diagnosis, they are prone to test signal attenuation and dispersion phenomena.

2.3.2 . Attenuation and Dispersion

The propagation of the TDR pulse at a high frequency in a cable can undergo two phenomena [1]: dispersion and attenuation [30], as shown in Figure 2.4 where the reflected signal obtained has undergone deformation and attenuation. This section will introduce these two phenomena.

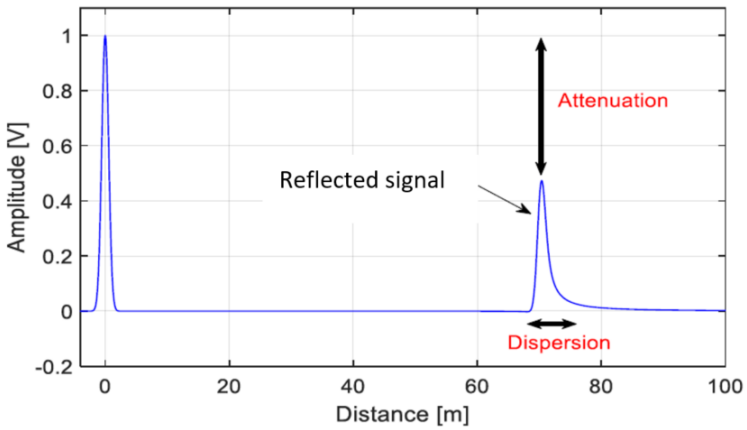


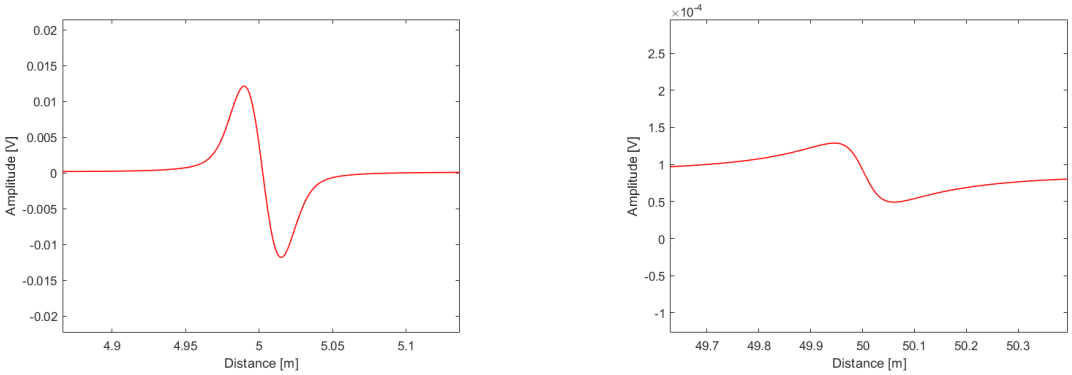
Figure 2.4 – Reflectogram of a $100m$ coaxial cable terminated with an open circuit. A Gaussian pulse is injected into the cable

[1] A third and frequent phenomenon concerns transmission lines arranged in close proximity to each other. This is crosstalk: part of the energy that travels on one line is transmitted to the other one by electromagnetic coupling. Two types of crosstalk could be distinguished [103], the effects of which overlap: capacitive crosstalk and inductive crosstalk.

2.3.2.1 Attenuation

Signal attenuation is mainly due to the cable resistance. It is observed by an exponential decrease in the amplitude of the signal during propagation. Consequently, it limits the diagnosis performance in terms of cable length. Attenuation is a limiting factor for reflectometry methods.

In order to demonstrate the attenuation phenomenon caused by the cable length, a shielding fault defined by the following parameters: length $l_f = 5mm$ and width $\theta_f = 180°$, is placed at the middle of a $10m$ and a $100m$ cable. The simulations consist of injecting a Gaussian pulse U_i, of unit amplitude, in the frequency band $[0;\ 4GHz]$, into one extremity of the wire and then recording the reflected signals at the injection point. Figure 2.5 shows the amplitude variation of this fault for the different cable lengths. It is shown that, as the length of the cable increases, the attenuation of the signal increases and the amplitude of the fault signature decreases as described by (1.16) in chapter 1.



(a) $10m$ cable (b) $100m$ cable

Figure 2.5 – Soft fault signature for different cables' length

The cable resistance increases with the used test signal frequency, so the narrower the pulse (wide-band), the greater the attenuation. For a coaxial cable, the resistance R is given by the following equation:

$$R = \frac{1}{2}\sqrt{\frac{f\mu_0}{\Pi}}\left(\frac{1}{R_{ext}\sqrt{\sigma_{ext}}} + \frac{1}{R_{int}\sqrt{\sigma_{int}}}\right) \tag{2.2}$$

where μ_0 is the magnetic permeability of vacuum, R_{int} is the radius of the inner conductor and R_{ext} is the radius of the outer conductor of the coaxial cable, σ_{int} and σ_{ext} are the inner and outer conductor conductivity's, respectively. Figure 2.6 shows how varying the maximal frequency will cause the resistance variation of the cable and thereby affecting the soft fault signature. Here, the soft fault is the shielding damage introduced above at $70m$ in a $100m$ coaxial cable. It is obvious that the fault signature amplitude is affected by the excitation frequency. As the frequency increases, the attenuation increases.

In multi-branched wiring networks, using a single sensor may no longer be relevant. The signal attenuation may explain this due to the distance and multiple junctions. Although the distance between the injection point and the fault may be determined, identifying the

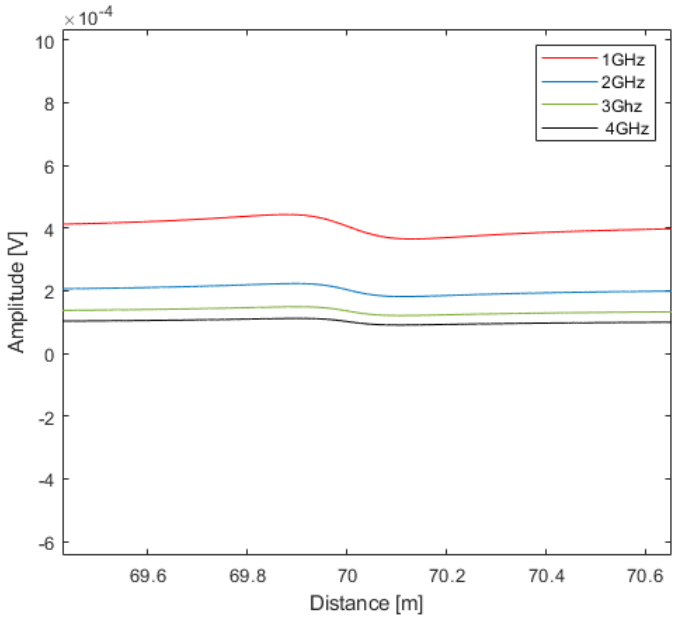


Figure 2.6 – Soft fault signature at four different excitation frequencies in a $100m$ cable

faulty branch remains ambiguous. To illustrate this, we consider the multi-branched network in Figure 2.7 with five branches B_i; $i \in \{1, 2, 3, ..., 5\}$ with the same characteristics. The branches B_2, B_4 and B_5 are terminated by open circuits. An open circuit is observed on branch B_3 at $25m$ from the injection point. Only one reflectometer is placed at the extremity of B_1 to diagnose the whole network. Figure 2.8 shows the reflectogram corresponding to the network presented in Figure 2.7. It is obtained by injecting a Gaussian pulse at the input of the network. The first negative peak corresponds to the first junction at $20m$ from the injection point. Next, a fault is detected at a distance of $25m$ from the reflectometer. From the reflectogram, the detected fault could be located either on branch B_2 or branch B_3. Although the distance between the injection point and the fault may be determined, identifying the faulty branch remains ambiguous.

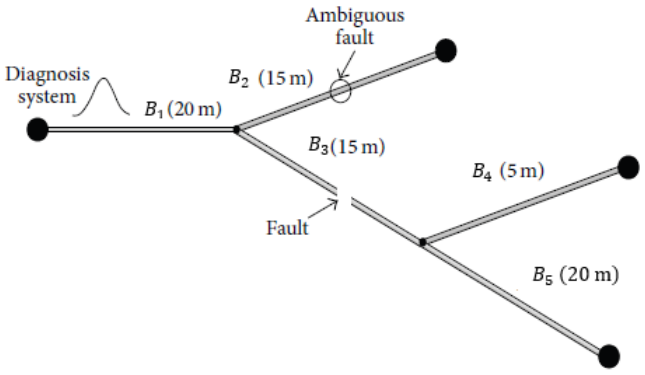


Figure 2.7 – Fault location ambiguity in a branched network [14]

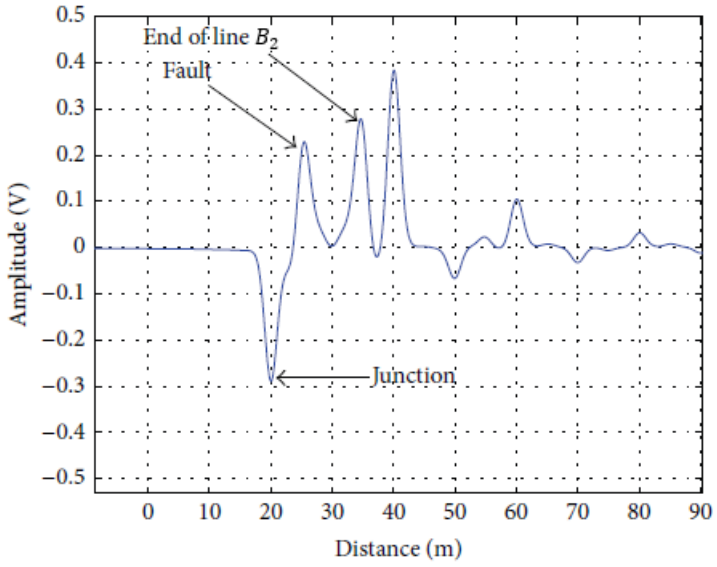


Figure 2.8 – Multi-branched network reflectogram using TDR [14]

This ambiguity becomes worse if there are other faults in the network. In addition, the test signal will be attenuated due to the multiple reflections. This makes the detection of a fault at a significant distance from the injection point difficult if not impossible, especially in the case of a soft fault represented by a peak of very low amplitude. Moreover, the ramifications can obscure the detection of a fault when it is found in a blind zone [17]. If the TDR system is not matched to the intrinsic impedance of the cable, a part of the energy of the injected signal is directly sent back, without even going inside the cable. This creates an additional high amplitude peak (impedance mismatch) at the beginning of the reflectogram which may hide a fault close to the injection point.

2.3.2.2 Dispersion

Dispersion is the consequence of the absorption of the high frequencies of the signal by the cable. The narrower the pulse (wide-band), the greater the dispersion [31, 32]. In the general case, the presence of the factor β is equivalent to introducing a variation in the speed of the signals as a function of their frequency; thus, a non-sinusoidal signal associating a set of harmonics of different frequencies is deformed.

In the case of dispersive lines, the propagation speed (v_p) depends on the frequency as shown in equation (2.3), where $\omega = 2\pi f$(rad/s) is the frequency and β(radian/m), defines the phase constant, causing the high-frequency components of the test signal to propagate at a faster speed, faster than low-frequency components. This further distorts the reflected signal and leads to the location accuracy reduction, especially for long-distance cables [9].

$$v_p(f) = \frac{\omega}{\beta} \tag{2.3}$$

In order to show the effect that the cable length can have on the fault location accuracy,

Figure 2.5 demonstrates this phenomenon for longer cables where the obtained reflected signal has undergone deformation. The whole problem lies in the fact that in a dispersive medium (such as an electric cable), the propagation time of an electromagnetic wave varies with the frequency, which deforms the reflected signal all the more as the path traveled by the wave is large, thus seriously affecting the quality of the measurement. This deformation affects the fault location accuracy. Scattering on a pulse signal, for example, turns that fault signature more flattened. However, when we want to locate a singularity (or even a fault), we must measure the peak abscissa position of the injection point and that of the echo to calculate the difference "$\Delta(t)$" which is used to determine the distance separating the wave generator from the fault. Nevertheless, if the reflected signal is very distorted, choosing the exact peak abscissa position of the echo is tricky. The obtained localization error is not constant but varies according to the distance.

The effect of frequency variation on the fault signature, the dispersion is observed between the signatures of the same fault at different frequencies on the reflectogram in Figure 2.6. Therefore, improvements in measurement and processing are necessary to overcome these limitations and make the use of the reflected signal more efficient.

In addition to the application of particular processing techniques, the choice of the mode and type of injection (nature of the signal, rise time, duration) can also improve the detection performance. Thus [33] has shown that in TDR the detection of a soft fault required the use of a signal presenting a low rise time and therefore wide band. This is however limited by the frequency losses and the dispersion. This is therefore difficult to apply for long cables or presenting significant losses (the reflected signal will be strongly attenuated). In addition, increasing frequency requires more expensive equipment.

Some wave-forms are more suitable and more robust to propagation effects in cables. It is shown that the exponential-looking wave-forms are optimized to introduce the smallest distortion compared to a rectangular pulse. In [33], it is shown that, at the same bandwidth, the distortion rate introduced by a Gaussian pulse is better than that of a rectangular pulse, especially for longer lengths. Therefore, the Gaussian pulse is more robust to the propagation and better suited for an application in time domain reflectometry. Therefore, in our study, we will use Gaussian pulse.

In the literature, we have found a method to reduce signal dispersion as it propagates in a transmission line. This method involves injecting a type of wave called "Speedy Delivery (SD)" [104], shown in Figure 2.9. This signal is minimally distorted and preserves the waveform as it travels through the cable. This exponential wave family improves measurement accuracy and resolution, ensuring the better fault detection and localization.
Another more efficient method to overcome the limitations (attenuation, robustness to noise, etc.) of the "Speedy Delivery" method to better compensate for signal dispersion is known as the adaptive correlation method [97]. This method performs the correlation operation more cleverly, so-called "dynamic or adaptive" correlation, where the signal that we use as a reference is deformed equivalently iteratively as a function of the correlation step. Indeed, if we know the laws of the signal dispersion in the cable (propagation equations), we know that the more we go towards the cable end, the more the signal dispersion will be important. So, if we consistently modify the curve that we use as a reference to consider the dispersion for each point

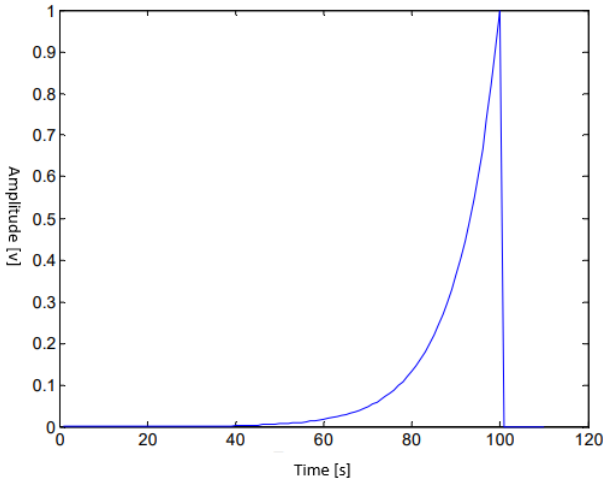


Figure 2.9 – "Speedy Delivery" waveform

(or group of points), we will obtain better results (this process increases the injected/reflected signal ratio). The advantages provided by this method over existing methods are its ability to locate the fault with better precision, and robustness to noise. Also it does not require to know the length of the tested cable. However, the major disadvantage is that it requires large memory resources to store the propagated signals for a given cable model. For this reason, it is not considered in this work.
Recently, authors in [105] have proposed a new post-processing approach that aims to compensate the dispersion effect of the wave during its propagation in the cable. It includes three steps:

1. Estimate the cable propagation velocity v_p.
2. Make a first injection (without dispersion compensation) of the test signal into the cable: fault position estimation.
3. Make a second injection of the test signal by compensating the dispersion undergone by the signal at the fault position pre-estimated in step (2): fault position correction.

The test results have shown the advantages of this new approach in fault localization with better precision, especially for long-distance cables. This suggested method can be applied to any test signal injected in the wire under test. However, this approach was only validated in the point-to-point cable. It should be evaluated in the case of multi-branched networks. For this reason, it is also not considered in this work.

Table 2.2 summarizes the methods used to reduce signal dispersion, their advantages, and inconveniences.

The problems associated with diagnosis, described previously, are aggravated in the case of a multi-branched network and a simple measurement at the single injection point is not sufficient. Indeed, this result alone can lead to ambiguities about the exact location of a fault. Distributed reflectometry seems to be a good candidate to address this problem. This is the subject of the next section.

Table 2.2 – Summary table of the methods used to reduce signal dispersion

Method	Advantage(s)	Inconvenience(s)
Speedy Delivery (SD)	- Preserves the waveform as it travels through the cable -Improves measurement accuracy and resolution	-Attenuation -Robustness to noise
Adaptive Correlation	-Locate the fault with better precision -Robustness to noise	Requires large memory resources to store the propagated signals for a given cable model
Method to Improve Fault Location Accuracy Against Cables Dispersion Effect	Locate the fault with better precision, especially for long-distance cables	Only validated in the point-to-point cable case

2.4 . Distributed Reflectometry for Soft Fault Detection and Location

From the previous section, we have shown that there may be fault location ambiguity that worsens if there are other faults in the network. In the case of a branched wire network affected by multiple faults, only the first fault closest to the reflectometer can be detected and located. The others can be interpreted as secondary echoes. In addition, the test signal is affected by attenuation and dispersion phenomena. These phenomena make it challenging to detect a fault located at a significant distance from the injection point, especially in the case of a soft fault characterized by a peak of very low amplitude.

Distributed reflectometry [9,33,92] addresses this problem very well. It consists of measuring several points of the network. Each reflectometer provides a different view of the network, thus making it possible to cover the blind area [21] and remove ambiguities as to the location of a fault.

In the case of Figure 2.7 in section 2.3.2.1, it is possible to add another reflectometer (sensor) at the end of B_2 using distributed diagnosis. The ambiguity disappears thanks to this new sensor but would recur upon the occurrence of a new fault on B_4. So, another sensor should be added to overcome this ambiguity. Then, distributed reflectometry is a suitable method to overcome ambiguity problems. However, this implementation technique in a multi-branched wired network raises new challenges related to the number of sensors such as computational complexities, sensor fusion and energy consumption.

A minimum number of sensors is required to set up a distributed reflectometry system in a multi-branched network. These sensors must be placed at strategic points in the network to precisely locate a fault. Generally, for any junction jun connecting $N_b(jun)$ branches, it is sufficient to place $(N_b(jun) - 1)$ sensors. In a network with N_{jun} junctions, the minimum number of sensors depends on the number of junctions and the number of branches at each junction. The minimum number of sensors required is given by the following relation [92]:

$$N_s = \sum_{jun=0}^{N_{jun}-1} (N_b(jun) - 1) \tag{2.4}$$

Figure 2.10 shows an example of the placement of sensors in a multi-branched network where $N_s = 6$. Therefore, it is sometimes advisable to reduce the cost of the diagnosis by optimizing the number of sensors without degrading the quality of the diagnosis.

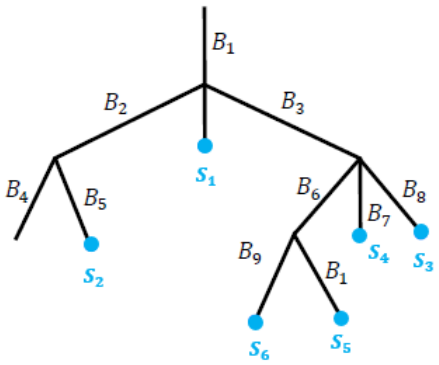


Figure 2.10 – Placement of sensors (in blue) in a multi-branched network

In order to overcome these constraints in section 2.3 and to enhance the detection and location of soft faults in multi-branched topologies, several methods have been used in the literature and are introduced in section 2.5. These tools combined with distributed reflectometry are used for the detection, location, and characterization of several faults in different types and topologies of wired networks.

2.5 . Enhanced Methods for Soft Fault Diagnosis

Reflectometry methods make it possible to test wire networks by injecting an electromagnetic wave. This test is done by examining the observation data either from the measurement or from an analytical or numerical model. Obtaining observation data related to the condition of the cables and the configuration of the considered networks is commonly called the forward problem. However, in most applications, what is of interest to an engineer or scientist is to inspect the condition of the cable. Observation data must allow the physical parameters of these cables to be traced. The means of obtaining electrical parameters of cables or the structure of networks from observation data is called inverse problem. In the case of a hard fault, these parameters are the lengths of the branches and the loads at the end of the lines. In the case of a soft fault, these parameters are the impedance and the position of the fault.

In the 2 thesis studies, [90, 106], two methods are used to solve the inverse problem: iterative or direct inversion. The first type uses Genetic Algorithm (GA). The second method is based on the response from a parametric inverse model. To achieve this inverse model, Artificial Neural Networks (ANN) are used.

Recently in [90], the authors presents a method [107] to solve ambiguities due to multiple paths and to detect and locate accurately electrical faults in a multi-branched network. It is based on the combination of distributed reflectometry and graph theory.

The following subsections describe these mentioned techniques and their pros and cons

given in the literature.

2.5.1 . Genetic Algorithm

Iterative inversion uses genetic algorithm to minimize the error between the response estimated by the direct model (FDTD modelization), and the measurement [15, 77, 108, 109]. Genetic algorithms are global optimization techniques based on the principles of genetics, and natural selection [110]. This algorithm is applied iteratively until a stop condition is satisfied. It is generally linked either to the number of iterations or to the error criterion, which must be less than a predefined threshold. The flow chart of the characterization method used in [15] is presented in Figure 2.11. The detailed process is explained in the following:

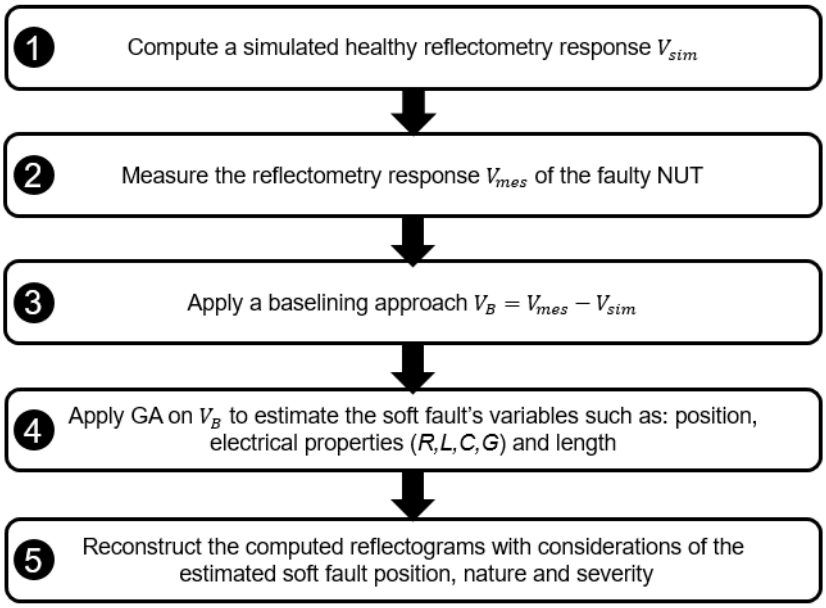


Figure 2.11 – Characterization method [15]

1. First, three-dimensional numerical modeling of coaxial cable is developed and validated by distributed parameters, including the frequency-dependent effects such as the skin effects. The TDR response noted V_{sim} is obtained based on the computed R, L, C, G parameters.

2. Second, the reflectometry response of the faulty NUT is measured. The signature for the faulty network is denoted as V_{mes}.

3. Third, the baselining approach will be applied as it enhances the weak signature of the soft faults. A NUT may mask a soft fault by other impedance discontinuities such as splices, loads, etc. By performing the difference between the faulty system and the reference one, baselining will enhance the visibility of the existing soft faults. The baselined reflectometry response will be computed as $V_B = V_{mes} - V_{sim}$, where V_{mes} is the reflectometry response of the faulty NUT containing the signature of the soft fault (stage 2) and V_{sim} is the reflectometry response containing the contribution of the healthy NUT (i.e. splices, loads, etc.) (stage 1).

4. At the end of step 3, soft faults are detected and located. This step proposes to estimate its characteristics: determine its per-unit length parameters such as the resistance, the capacitance, the inductance, and conductance. For this, a GA is applied as it proves its efficiency in calculation time and simplicity. It has been applied with success to the blind reconstruction of black-boxed unknown complex wire networks [111].

   The proposed algorithm considers three set of variables to run:

   (a) $Nature = \{R, L, C, G\}$ corresponds to the nature of the fault where a vector includes the starting values (in the case of a healthy cable): the resistance R, the inductance L, the capacitance C and the conductance G. These variables are computed in step 1 of the considered methodology.

   (b) $O = \{L_f, x_f\}$ is the fault parameters, where L_f is the fault length and x_f is the fault position.

   (c) Δ_v is the variation with respect to the $RLCG$ parameters of the cable under test. Here, Δ_v is a vector of 4 elements including the variation of R, the variation of L, the variation of C and the variation of G. i.e. $\mathbf{\Delta}_v = [\Delta R\ \Delta L\ \Delta C\ \Delta G]$. At the output, one can define, for example, the resistance of the fault as $R_d = R(1 + \Delta_v(1)/100)$.

   The algorithm iteratively improves the simulated baselined reflectometry response V^B_{sim} by changing the set of variables $\{Nature\}$, $\{O\}$ and $\{\Delta v\}$ in order to reduce the differences between the hypothesized faulty network model and the baselined measured one demonstrated by V_B in step 3. An objective function $fval$ is a set that contains the difference between V^B_{sim} and V_B. This optimization algorithm aims to minimize $fval$ by varying the set of variables of the soft fault. Once the algorithm converges to a minimum, that is to say, $fval \approx 0$, the algorithm is stopped, and the best candidates of the soft fault properties are reached. Based on the electrical faults parameters obtained from the GA, i.e. the resistance R_d, the inductance L_d and the capacitance C_d, the impedance of the fault Z_d is calculated.

5. The final step of the proposed methodology is to reconstruct a NUT based on the data gathered from the previous procedures. The described methodology can blindly characterize the soft fault in the network and reconstruct the NUT embedding it, with the lengths of the branches composing it and the load impedance connected to the NUT.

The major drawback of this type of inversion is its computation time which strongly depends on the complexity of the network topology. Therefore, it is not suitable for real-time diagnosis.

2.5.2 . Artificial Neural Networks

The ANN training (adjustment of their internal parameters) requires the use of datasets. In cable diagnosis, the datasets are constituted of examples linking the reflectometry response of the transmission line to the position of the fault. In [112], ANN is used to localize and characterize the fault (impedance) of hard faults in complex networks. However, they cannot solve the problem of fault location ambiguity using a single measuring point [106], especially in the case of a soft fault characterized by little variation. To overcome this limitation, authors in

[113] propose to distribute the measurements to several points of the network and, consequently, to merge the data between the distributed reflectometers using and ANN. Figure 2.12 illustrates the block diagram of this method.

For a complex wired network monitored by several reflectometers S_i, $i \in \{1, ..., N\}$, N is

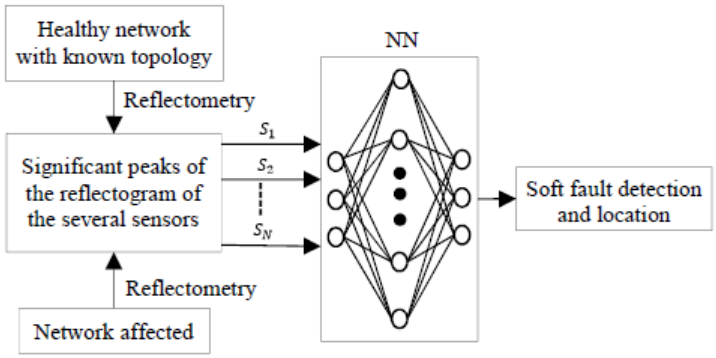


Figure 2.12 – Block diagram of the fault location procedure using ANN

the number of the distributed reflectometers. This is done by using Multi-Layer Perceptron Neural Networks (MLP-NN). At first, to facilitate the reflectogram analysis, each sensor must calculate the difference between the response of a healthy and faulty network. This processing avoids the different peaks caused by the nodes in the network. Then, only the main peaks are detected, and for each peak, their magnitude and location (distance from the origin) is stored. These peaks of each sensor are then injected into the ANN. Its output gives information about the detection and location of the soft faults.

In order to select the main peaks in the differential reflectogram, detection thresholds is used to detect positive and negative peaks. These thresholds are represented by an exponential decay according to the cable attenuation coefficient. Their equation are as follows:

$$positive\,threshold = \frac{A_{min} e^{e^{-2\alpha L}}}{(1 - \Gamma_0^2)} \tag{2.5}$$

$$negative\,threshold = -\frac{A_{min} e^{e^{-2\alpha L}}}{(1 - \Gamma_0^2)} \tag{2.6}$$

Where $A_{min} = 3.10^{-3}$v is the sensor minimum detection threshold. $(1 - \Gamma_0^2)$ represents the coupling compensation term between the sensor and the network. L is the length of the network, and α is the linear attenuation.

A neural network is formed for each fault parameter to be estimated. The number of output neurons corresponds to the number of estimated quantities. For the ANN design for an example network of five branches, a fully connected three-layer (input, hidden, and output) feed-forward neural network has been used with hyperbolic tangent activation functions in the hidden layer, and an output layer constituted of five neurons having a linear activation function.

The output layer of the network is composed of five neurons corresponding to the number of branches in the wiring network. It provides information about the faulty branch and accurately locates the soft faults in the network. If a soft fault is detected, the corresponding NN output gives the location of the soft fault on the faulty branch; otherwise, it is "zero".

The architecture type, number of neurons, size of the hidden layer, and training dataset size play a key role in determining the best NN architecture [114]. The design of the NN has to be achieved carefully with the appropriate number of hidden neurons. An NN with insufficient neurons will not learn the training database correctly. On the other hand, using too many neurons can lead to an over-fitting phenomenon.

First, the training data set examples are presented to the ANN. After a first training iteration, the output of the ANN is compared to the one contained in the data set. To reduce the error obtained at the output, the variables of the ANN are adjusted according to Levenberg–Marquard algorithm. Finally, the generalization capability of the ANN is assessed by calculating the Mean Square Error (MSE) obtained on the test set, which evaluates the quality estimation of the ANN.

The major disadvantage of the ANN method used in [112] is the significant time required for database creation, network settings adjustment, and training (approximately 2 hours using a PC equipped with Pentium 4 Core Due Processor and 4 GB of RAM for a Y-network increase to 4 hours for a network with five branches). However, the inversion step is quick (about $1second$). Even though in [113], by working with most significant peaks of differential reflectogram rather than the whole reflectometry response, they were able to reduce the size of the required data and consequently the number of neurons in hidden layer and the time of processing. However, they require the knowledge of the topology of the network and its components (socket, connector, junctions, cable, etc.). Moreover, it requires increasing the number of examples in the database and optimizing the number of neurons in the hidden layer. Generally, these responses have a high dimension, leading to complex ANN including a lot of internal parameters to be adjusted. As a result, a large number of examples is needed in the database. This corresponds to the "cruse of dimensionality". Moreover, a high sensitivity to noise can be expected.

2.5.3 . Graph Theory

The objective of the developed method in [90] is to combine the MCTDR method with the graph theory in distributed diagnosis architecture. The main idea is to represent the complex wired network by a graph with nodes and edges. Several graph theory algorithms are used to facilitate the faulty branch's decision-making. Then, a connection matrix is used to ensure the connection between different nodes.

Graph theory is a mathematical representation of a network that describes the relationship between lines and points [115]. Graph theory is considered a useful tool that allows graphs to represent many practical problems. It deals with problems that have a graph or network structure.

A graph $Gr = (No \quad E)$ can represent a complex wired network. The loads, junctions, and reflectometers can be identified as the nodes 'No', and the connecting transmission lines can be modeled as the edges 'E'. The wired network topology can be modeled by the use of $n_{No} \times n_{No}$ connection matrix $cm = a_{ij}$, also known as adjacency matrix, where n_{No} is the number of nodes in the network. The component a_{ij} indicates a connection from node j and to node i. A length B_{ij} of a line at a position $(i; j)$ indicates that the node i is adjacent to

node j. Therefore we define cm by:

$$a_{ij} = \begin{cases} B_{ij} & if\ there\ is\ an\ edge\ from\ node\ j\ to\ node\ i, \\ 0 & else. \end{cases} \tag{2.7}$$

The different reflectometers diagnose the complex network. Then, the result is sent to a central processing unit that processes the reflectograms to decide the fault location. To facilitate the reflectogram analysis, each reflectometer must calculate the difference between healthy and faulty network responses. This processing avoids the different peaks caused by the ramifications in the network. To identify faults peaks, post-processing is necessary. Two reflectometers at a minimum must confirm the detection of faults peaks. Then, the fault distance from the different injection points is calculated to locate a fault, and then the faulty branch is identified. This can be done in the following steps:

(a) Analyze the reflectograms obtained to detect different faults peaks.

(b) Model the wired network as defined above and represent the distance between different nodes in a connection matrix cm.

(c) Use the reflectogram of the first reflectometer and the graph theory's different algorithms to find the first node connected to the faulty branch, noted j.

(d) Simplify the network with only the nodes connected to the node j (explained below). The simplified network can be stored in a vector b_j, containing the jth row values of cm.

(e) From the reflectogram of the second reflectometer and the different algorithms of the graph theory, the second node connected to the faulty branch is founded, noted i.

(f) Simplifies the network with only the nodes connected to the node i (explained below). The simplified network can be stored in a vector b_i, containing the ith column values of cm.

(g) The intersection between two vectors obtained in steps (d) and (f) gives the faulty branch connected to nodes j and i.

The information obtained by the different distributed reflectometers is merged. This fusion is performed using different graph theory algorithms such as: Breadth-First Search (BFS), which is an algorithm that explores all accessible nodes in a graph from the source node [116]. Dijkstra's algorithm, is a graph search algorithm that solves the shortest path problem from a single source to many destinations [117], and Nearest Neighbor Algorithm (NAA), which pre-processes a given dataset da in such a way that for an arbitrary forthcoming query vector q we can quickly find its nearest neighbors in da, where nodes correspond to the elements of da, and each node is connected to its nearest neighbors by directed edges. Then, for a given q, one first takes an element in da (either random or fixed predefined) and makes greedy steps towards q on the graph: at each step, all neighbors of a current node are evaluated, and the one closest to q is chosen [118].

The disadvantages of the graph theory application to distributed diagnosis is that it requires knowledge of the topology of the network and its components (socket, connector, junctions, cable, etc.).

These tools combined with distributed reflectometry allow detection, location, and characterization of several faults in different types and topologies of wired networks. The major disadvantage of these method is the significant time required for database creation, network settings adjustment, and training. In addition, they lack autonomy in the analysis and interpretation of measurements to extract the state of the wired network. Moreover, some of them requires knowledge of the topology of the network and its components (socket, connector, junctions, cable, etc.). For this, in this work, we need a simple technique which allows to save time, reduce the computing complexities and doesn't require the knowledge of the topology of the network.
Data-driven models are implicit empirical models derived from analysis of available data. Their derivation requires a minimal a priori knowledge about process physics, but a significant amount of historical process data which may contain faults and their symptoms. It is mainly based on computational intelligence and machine learning methods, which are referred to as data-driven methods. These include multivariate statistical projection methods such as Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN) and Support Vector Machines (SVM). PCA is a simple multivariate data-driven statistical modeling method, which has gained a remarkable acceptance in industry for statistical monitoring and control of multivariate processes [119].
Although the combination of the reflectometry and PCA in the literature was proposed for an efficient reduction of the reflectometry data space [37], the application of PCA on reflectometry-based data has not been used, yet, for fault detection and diagnosis in wired networks.
The following section will present the PCA method for wire Fault Detection and Diagnosis (FDD).

2.6 . Principal Component Analysis for Wire Fault Detection and Diagnosis

2.6.1 . PCA Principles

PCA is a multivariate data-driven statistical modeling technique [36]. It uses information redundancy in a high-dimensional correlated input space to project the original data set into a lower-dimensional subspace defined by the principal components (PCs). One of the main research objectives of PCA is dimensionality reduction. It considers, in several cases, a substantial variability percentage in the data that can be interpreted using a limited number of components.

The PCA analysis begins with the data matrix $X^* \in \mathbb{R}^{n \times m} = [\mathbf{x_1^*}, \ldots, \mathbf{x_m^*}]$ which consists of m variables of the system, with n observations ($n > m$). In our work, each variable corresponds to the reflectometry response of the studied network at a specific excitation frequency.

Data pre-processing is important; First centering: mean center the X^* variables (subtract the mean value of each variable from its true values). Using uncentered data will have unfavorable effects on the statistical tests' confidence limits. Second, scaling: it is used to affect the magnitude of the data by applying a scaling factor to the centered data of each variable. The purpose is to make the numerical values of each variable equally significant (on average).

Since PCA (and other related techniques) is based on the concept of maximizing the variance of the principal components, which are linear combinations of the original data, it is natural to scale the $\mathbf{x}^*$-variables to make their variances equal. This is obtained by dividing the data of each variable by the standard deviation of the variable data.

After normalizing X^* to a matrix X_0 with zero mean and unit variance, PCA determines the optimal linear transformation as follows:

$$CM = \frac{1}{n-1} {X_0}^T X_0 \tag{2.8}$$

where CM is the covariance matrix of X_0, and by means of Singular Value Decomposition (SVD), one can write CM as:

$$CM = P \Lambda P^T, where \quad P^T P = I_m \tag{2.9}$$

where $P = [\mathbf{p_1}, \mathbf{p_2}, ..., \mathbf{p_m}] \in \mathbb{R}^{m \times m}$ is the PCA loading matrix such that its columns $\mathbf{p_j}$ are the orthonormal eigenvectors associated to the eigenvalues λ_j of the covariance matrix CM for $1 \leqslant j \leqslant m$, and $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_m)$ is the diagonal eigenvalues matrix with decreasing order. Each eigenvalue represents the amount of variance that has been captured by one component.

According to [120–122], PCA decomposes the data matrix into two parts; the first explains the system main variation while the second encapsulates the residual information (noise):

$$X_0 = TP^T = T_k P_k^T + \tilde{T} \tilde{P}^T \tag{2.10}$$

where $T = [\mathbf{t_1}, \mathbf{t_2}, ..., \mathbf{t_m}] \in \mathbb{R}^{n \times m}$ is the score matrix, highlighting the relationship between the samples in X_0. The superscript ($\tilde{}$) is the residual matrix operator and k ($k \leqslant \max(n, m)$), is the number of Principal Components (PC).

The data matrix X_0 is decomposed as the principal space and the residual one. It is of great importance to decide how many components to retain in each space. k is determined using one of the proposed approaches in the literature: the scree plot, [123, 124], Cross-validation method [125], Kaiser- Guttman method [126], Parallel analysis method [127, 128], and Cumulative Percent Variance (CPV) method [129]. CPV is commonly used in fault detection, retaining k components having a cumulated variance greater than a prescribed threshold (as an example 90%) of the total variance. The variance of the PC scores, preferably given in percent of the total variance of the original variables, is an important indicator of how many PCs to include.

$$CPV(k) = 100 \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \tag{2.11}$$

In this study, we preferably use this latter criterion to define the size of the principal subspace. As proposed in the literature [36], we consider the Principal subspace defined with CPV higher or equal than 90%.

2.6.2 . PCA-based Fault Detection and Diagnosis

For fault detection, the PCA model of the system is developed based on healthy operating system data and then used to check new measurement data X. After scaling and centering the

new data, the residuals between the new measurement data and their projections in ((2.12) and (2.13)) are then analyzed through statistical test. Usually the Q statistic ((2.14)), also known as Squared Prediction Error (SPE), and the Hotelling's T^2 statistic ((2.15)) are used to represent the variability in the residual subspace and principal component subspace, respectively [130].

$$T_{new_k} = XP_k \tag{2.12}$$

$$\tilde{T_{new_{(m-k)}}} = X(I - P_k P_k^T) \tag{2.13}$$

$$Q = \tilde{T_{new_{(m-k)}}} \tilde{T_{new_{(m-k)}}}^T \tag{2.14}$$

$$T^2 = \sum_{1}^{k} \frac{\mathbf{t_j}^2}{\lambda_j} \tag{2.15}$$

The Q statistic method shows how well a new sample fits into the PCA model. It measures the difference (residual) between the sample and its projection onto the $m-k$ residual components retained in the model. The T^2 statistic method detects variation within the PC subspace information. T^2 represents the squared length of the projection of the current sample into the space spanned by the PCA model. It indicates how the PCA estimate of the sample is far from the multivariate mean of the data, i.e., the intersection of the principal components. Therefore, if a sample has an abnormal value of T^2 but a Q value below the limit, it is not necessarily a fault – it can also be a change of the operating region [121].

Hence, Q and T^2 statistical values are used for evaluating the fault presence. In the case of an abnormal event, the Q and T^2 statistic values will be greater than the confidence limits Q_α and T^2_α, respectively. Those limits ((2.16) and (2.19)) are calculated using the healthy original data X^* that is used to build the PCA model.

$$Q_\alpha = z_1 [\frac{c_\alpha \sqrt{2 z_2 {h_0}^2}}{z_1} + 1 + \frac{z_2 h_0 (h_0 - 1)}{{z_1}^2}]^{\frac{1}{h_0}} \tag{2.16}$$

$$z_i = \sum_{k+1}^{m} \lambda_j^i, \ i = 1, 2, 3 \tag{2.17}$$

$$h_0 = 1 - \frac{2 z_1 z_3}{3 z_2^2} \tag{2.18}$$

$$T^2_\alpha = \frac{k(n-k)}{n-k} F_{k,n-k,\alpha} \tag{2.19}$$

where c_α is the critical value of the normal distribution at α significance level and $F_{k,n-k,\alpha}$ is the Fisher–Snedecor distribution critical value [121].

The PCA model can be used with new system data to detect changes in the system. After detecting a probable outlier due to extreme T^2 or Q values, we can investigate the inputs (responsible variables) that highly influence their residual. Contribution plots are used for this purpose.

PCA is used in the literature for process fault detection and diagnosis, as in [36, 131, 132]. In the literature, the combination of the reflectometry and PCA was proposed in [37] for an efficient reduction of the reflectometry data space where the significant parameters of the reflectometry response data are extracted using PCA. However, so far, PCA has not been used

for fault detection and diagnosis in wired networks.
PCA meets our needs in terms of simplicity, which allows us to save time, reduce the computing complexities, and don't require the knowledge of the network's topology. Furthermore, it enables automatic computerized decision-making for the considered networks in the presence of faults.

2.7 . Problematic

This work intends to improve the soft fault diagnosis in multi-branched networks using this PCA-based reflectometry method. Two main research questions are tackled. The first is related to the best frequency selection (chapter 3) to be used in the TDR construction, and the second one is related to the sensor selection (chapter 4) in multi-branched wiring networks using the distributed reflectometry.

2.7.1 . Best Frequency Selection for Reflectometry-based Soft Fault Diagnosis using PCA

The frequency of the reflectometry signal is a critical parameter in detecting and locating a fault in a cable. Indeed, the higher the frequency of the reflectometry signal, the better is the resolution of the reflectogram and the localization accuracy of small faults. However, in the case of fault detection in long cables, increasing the signal frequency is not recommended as it introduces dispersion and increases signal attenuation, as discussed in section 2.3.2.

Although interesting methods have been proposed to enhance soft fault diagnosis in section 2.3, most of them are prone to test signal attenuation and dispersion phenomena. Moreover, the signature of the soft fault may be invisible on the corresponding reflectogram. Hence, the choice of the test signal bandwidth is critical and affects the diagnosis performance. Although bandwidth increases the resolution for detecting soft faults, it provides more attenuation. Thereby, a compromise between those two quantities should be done.

In practice, the expert configures and calibrates the Vector Network Analyzer (VNA) at a given frequency and records the healthy cable measurement. Measurements at the same frequency are then done on a faulty cable. Analysis of the measurements is established at this frequency on the computer. If the fault is not detected, this operation must be repeated. Therefore, there is a loss of information and time in addition to the subjectivity of the decision-making.

In this context, chapter 3 introduces a new approach for frequency selection in the case of soft fault diagnosis. The proposed method permits to configure and calibrate the VNA at different frequencies. It performs measurements at different frequencies for the healthy case. After which, the PCA model is established. It performs the new measurements at different frequencies. If a difference is detected between the model and the new data, the contribution of each variable (i.e., frequencies) to this difference is calculated. The algorithm then chooses the most relevant frequency to monitor the soft fault. The advantages are thus time saving and objectivity of the decision-making.

2.7.2 . Sensors Selection for Distributed Reflectometry-based Soft Fault Diagnosis using PCA

As it have been explained in section 2.4, the distributive measurements is the best solution to remove ambiguities in fault location. However, the injection of multiple signals into the NUT leads to computational complexities and sensor fusion problems. Moreover, energy consumption is a significant drawback of this method with respect to environmental constraints. The study on the reduction of the sensors' number in complex networks and its impact on the diagnosis quality is provided in [34]. However, it shows further challenges related to bandwidth allocation, communication protocol, and noise interference mitigation.

Thus, in [35], the cable life profile is included, allowing to reduce the diagnosis cost by avoiding the use of too many sensors in the network. With this solution, the reliability of the sensors in emission and reception is considered in the obtained statistics. This reliability differs from a sensor to another, and impacts the fault location.

In this context, chapter 4 introduces a new approach for selecting the relevant sensors to monitor and diagnose soft faults in multi-branched wired networks. It combines TDR distributed reflectometry with PCA. Indeed, for a given NUT, a distributed reflectometry approach is considered where the sensors perform their reflectometry measurements. These collected data are used to establish a PCA model coupled with statistical analysis tests to evaluate new measurements status. Whenever a fault is detected, the relevant sensors for monitoring and diagnosis are identified with high accuracy. Based on these results, the sensor's number could be reduced, and the non-selected ones could be inactivated temporarily, reducing energy consumption, computing complexities, and sensor fusion problems.

For this study, we assume that a reference network exists, for which the PCA model is established based on its TDR response and will be used to examine new measured data.

2.8 . Conclusion

Reflectometry methods are well suited to diagnosing wired networks. However, there are several limitations inherent to the method or physics of the problem, such as the measurement noise, in-homogeneity of propagation, blind zone, ambiguity of localization, attenuation and dispersion phenomena. Soft faults are challenging to detect because a very weak signature characterizes them. They are often the premises of a more severe fault, which can appear in a short period of time, and will be a potential source of incidents or accidents. The diagnosis of soft faults is, therefore, an important objective of research.

This chapter has discussed the different problems of soft fault detection and localization in multi-branched networks. Several post-processing methods have been revised. However, they remain prone to the attenuation and dispersion phenomenon. The limitations concerning the location accuracy can sometimes be very troublesome, especially in the case of very long cables. Indeed, the phenomena of attenuation and dispersion significantly reduce the location accuracy when the propagation distance is important. Hence, the choice of the test signal bandwidth is critical and affects the diagnosis performance. Although a wide bandwidth increases the reso-

lution for detecting soft faults, it provides more attenuation. Thereby, a compromise between those two quantities have to be done.

In the case of a multi-branched network, distributed reflectometry is used to solve the location ambiguities. However, several challenges related to bandwidth allocation, communication protocol, and noise interference mitigation are presented. Thus, reducing the number of sensors in the network is relevant to reduce energy consumption, computing complexities, and sensor fusion problems.

Finally, a throughout study about the existing enhanced methods in the literature for soft fault detection, localization, and characterization as well as their limitations is presented. This state of the art has made it possible to identify the main weaknesses of these methods, linked in particular to complex topologies and the detection of soft faults. Thus new tools used for fault detection and diagnosis are needed; for that, Principal Component Analysis (PCA) principle is presented. PCA meets our needs in terms of the simplicity which allows to save time, reduce the computing complexities and it doesn't require the knowledge of the topology of the network.

In this context, this work intends to improve the soft fault diagnosis in multi-branched networks using this PCA-based reflectometry method. Two main research questions are tackled. The first is related to the best frequency selection (chapter 3) to be used in the TDR construction, and the second one is related to the sensor selection (chapter 4) in multi-branched wiring networks using the distributed reflectometry.

3 - Best Frequency Selection for Reflectometry-based Soft Fault Diagnosis using PCA

Contents

3.1 . Introduction

This chapter introduces an approach to select the best test signal frequency for soft fault detection in wired networks. In the literature, the reflectometry method has been well investigated to deal with the problem of soft fault diagnosis (i.e., chafing, bending radius, pinching, etc.). A small impedance variation characterizes soft faults resulting in a low amplitude signature on the corresponding reflectograms. Accordingly, the detection of those faults depends strongly on the test signal frequency. The increase of test signal frequency enhances the soft fault "spatial" resolution, to the detriment of signal attenuation and dispersion. A combination of reflectometry-based data and Principal Component Analysis (PCA) algorithm is proposed to address this problem. The data are obtained from TDR responses of 3D based-models of faulty coaxial cable RG316 and shielding damages have been simulated at different frequencies. A PCA model is developed and used to detect the existing soft faults based on the obtained reflectograms. The method aims to determine the best frequency of the test signal to have the best soft fault diagnosis performance.

With this approach, we show how to automatically proceed with the frequency selection to detect changing impedance and shielding damage due to common chafing causes in coaxial wire geometry before any possible shorting or open condition in the conductors.

In this chapter, we develop an analytical model of the fault detection performances (Probability of Detection (P_d) and Probability of False Alarm (P_{FA})), where the simulation results will highlight the impact of different network elements on the method's performance. An experimental validation will show the method's advantages regarding frequency selection for soft fault diagnosis.

In this context, the modeling of the cable is presented: 3D EM with Computer Simulations Technology (CST) for short cables and an $RLCG$ model for long cables. Experimental results are provided to validate the models for healthy and faulty cables. Next, the proposed methodology for best test signal frequency selection is illustrated. After which, simulation results in the case of short and long cables are presented. Experimental validation is performed next to show the method's advantages regarding frequency selection for soft fault diagnosis. Finally, performance analysis is inducted to study the efficiency of the proposed method in the presence of noise.

3.2 . Faulty Coaxial Cable Modeling and Electromagnetic Simulations

Although there is a great variety of cable types and faults of potential interest, in this work we focus on coaxial cable and "chafe" faults. Chafe first ablates outer insulation, then shield, leaving inner conductors intact that typically results from vibration or forced translation of the cable against an external component or another cable. As a result of the chafe, the signal return path of a signal on the cable can be disturbed. Fortunately, this disturbance is detectable thanks to time-domain reflectometry.

Underlining the importance of the numerical modeling of soft faults, we will first consider

the precise electromagnetic characterization of a cable zone located near a soft fault. It will be then a question of modeling the phenomena and its propagation along the cable.

The study of numerical methods in electromagnetism must be based on the fundamental basis of electromagnetism represented by Maxwell's equations. Numerical methods are used to solve them in an approximate way. These equations are used in particular way to model the problems of propagation of electromagnetic waves. The results of these methods will lead to a precise expression of the disturbances generated by the fault on the propagation of the signals along the line.

3.2.1 . Coaxial Cable Modeling and Simulation

To validate and generate noise-free data, we have employed a commercial microwave simulator: CST (Computer Simulations Technology) Microwave Studio. Coaxial cables are considered in this work, where they are used for the numerical model to validate the proposed method and in the conducted experimental measurements. These cables are widely used, so the comparison can be made between the simulated and the measured data.

A 3D EM model of a coaxial cable is developed using CST as shown in Figure 3.1. This is a healthy RG316 coaxial cable model. The cross-section shows four concentric cylinders defining the geometry: outer plastic jacket, shield, inner dielectric insulator made from PTFE, and the core. The coaxial cable construction specifications are listed in Table 3.1.

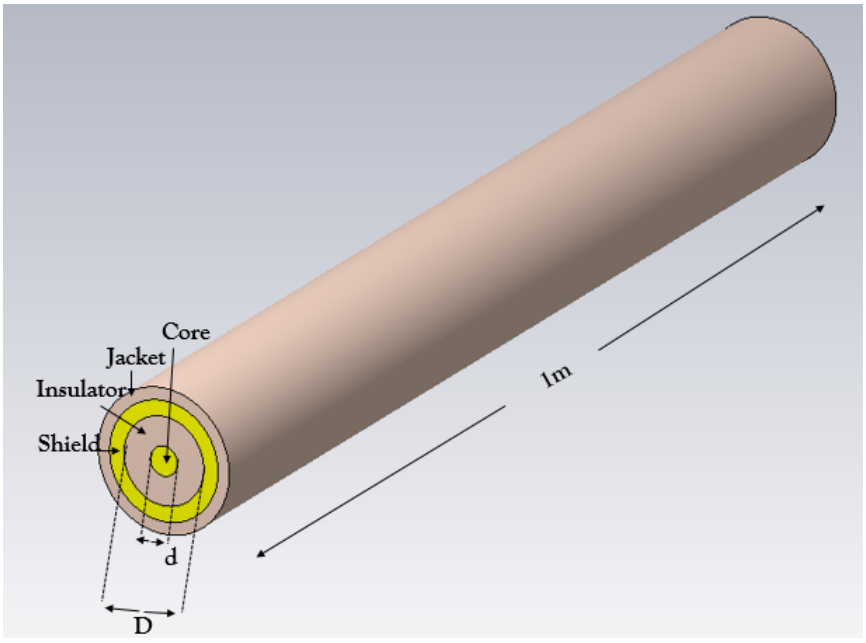


Figure 3.1 – The developed 3D EM model of a shielded coaxial

The external diameter of the inner conductor is $d = 0.51mm$ and the internal diameter of the outer conductor is $D = 1.52mm$. The conductors are made of copper with a conductivity $\sigma = 5.81 \times 10^{-7}(S/m)$. The insulation is made with Teflon (PTFE) with a relative permittivity $\epsilon_r = 2.1$, a relative permeability $\mu_r = 1$. The thickness of the shielding is $0.54mm$. Finally, there is an outer Teflon sheath of thickness $0.43mm$ which envelops all the components. The

Table 3.1 – RG-316 Coaxial cable construction specifications

RG-316 coaxial cable	Material	Diameter [mm]
Inner conductor (Core)	Copper	0.51
Dielectric	Polytetrafluoroethylene (PTFE)	1.52
Outer conductor (Shield)	Copper	2.06
Jacket	Fluorinated Ethylene Propylene (FEP)	2.49

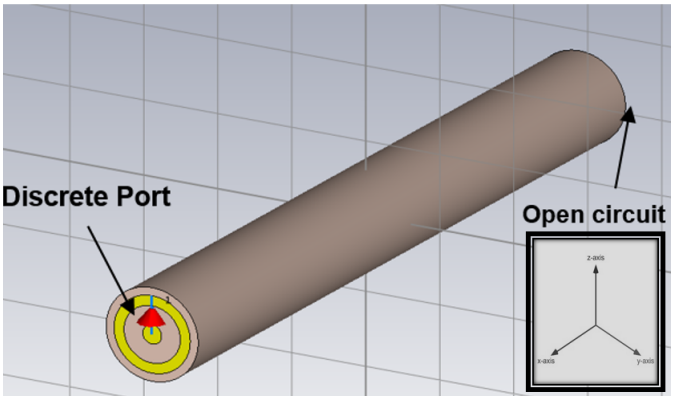


Figure 3.2 – The 3D model port

length of the model is $l = 1m$. The cable characteristic impedance is $Z_c \approx 45\Omega$ using the following equation:

$$Z_c = \frac{138 \times \log_{10}\left(\frac{D}{d}\right)}{\sqrt{\epsilon_r}} \tag{3.1}$$

A discrete port with a $Z_0 = 50\Omega$ impedance connects the two conductors through the center point as shown in Figure 3.2 to inject a test signal and receive the reflected one. The transmission line is supposed to be left open-circuited at its end, and the power supply is a Gaussian excitation source located at x= 0 with a maximum frequency of $4GHz$, lower than the cable cut-off frequency ($f_c \approx 15GHz$) to avoid the propagation of the higher modes [133]. The simulations consisted of injecting a Gaussian pulse into the wire and recording the reflected signals.

Figure 3.3 represents the TDR response for the fault-free cable that is recorded at the injection point. The signal propagates from the source point to the end. As there is no discontinuity along the transmission line, the perfect transmission of the electric field signal through a faultless medium is observed. The peak at $0m$ is due to the reflection from the impedance mismatch between the cable and the test circuitry (the port), such that the reflection coefficient at the input of the cable ($\Gamma_E = \frac{Z_c - Z_0}{Z_c + Z_0} = -0.0538$). The peak at $1m$ corresponds to the open circuit at the end of the cable. As mentioned in section 1.4.1.1, in case of an open circuit, $Z_l = \infty$ thus $\Gamma_l = 1$, here in practice $\Gamma_l < 1$ due to attenuation through the transmission line. The amplitude of the peak at the end of the cable is equal to $0.83V$, compared to the amplitude of the Gaussian pulse $1V$.

There are several sources of perturbations in reflectometry measurements, such as the perturbation caused by the connectors, as seen above. Since the reflectometer must be connected to a wide variety of cables or anchors, it is not generally feasible to match the impedance of

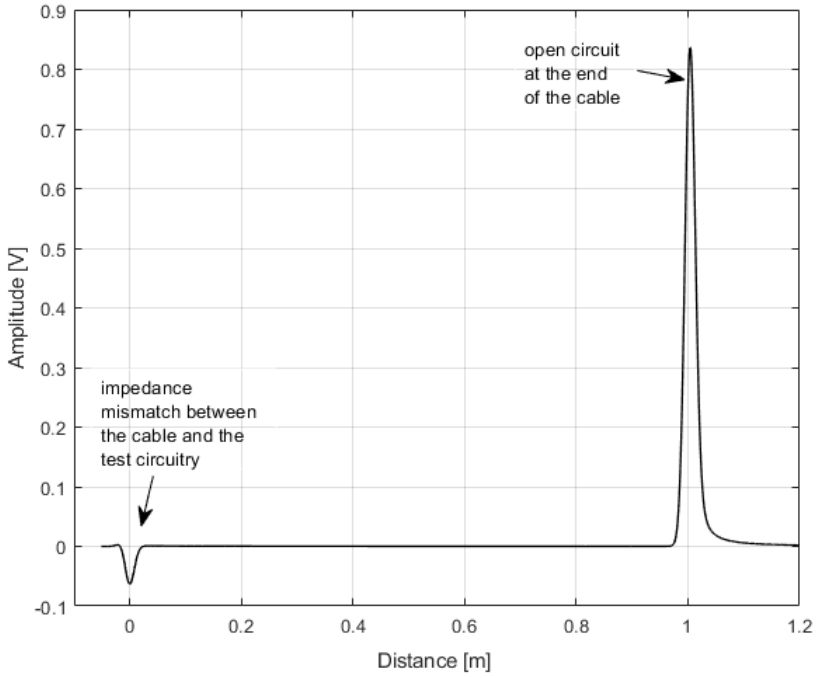


Figure 3.3 – TDR simulation result for the $1m$ modeled coaxial cable with an open circuit at the transmission line end

the reflectometer to the wire. This means there will always be a reflection between the testing system and the wire under test. The test-lead, connectors, adapters, etc., all contribute to this reflection in different ways. However, these perturbations can be reduced by using baselines, where comparing the response of the faulty network with either the pre-measured or simulated response of its (known) healthy condition is done [59].

3.2.2 . RLCG based Model

For the cable modeled in the previous section, the R2RLCG method explained in section 1.4.1.1 is applied to extract its $RLCG$ parameters. Figure 3.4 presents the extracted $RLCG$ parameters. The results are consistent with the previous assumption as they show the dependence on the square root of the frequency of the resistance R, and a linear dependence of the conductance G with respect to the frequency. The capacitance $C = 1.08{\times}10^{-10}(Farad/m)$ and inductance $L = 2.18{\times}10^{-7}(Henry/m)$ are constant over the reflectometry working frequency band. Indeed, this measurement had to be made with the VNA which stopped at $3.6GHz$. For the constant values (L, C and Z_c) we will have the same values at $4GHz$ and for R and G, it would be necessary to look at the leading coefficients of the trends (line for G and square root for R) to know the values. Figure 3.5 shows the real and the imaginary parts of the characteristic impedance as a function of the frequency. It is shown that Z_c tends to be constant at high frequency and the effects of R and G in Z_c tends to disappear.

Figure 3.6 shows the TDR reflectograms of the RLCG-based model and the 3D EM models. For these simulations, the used bandwidth is ranged from DC to $4GHz$, and the width of

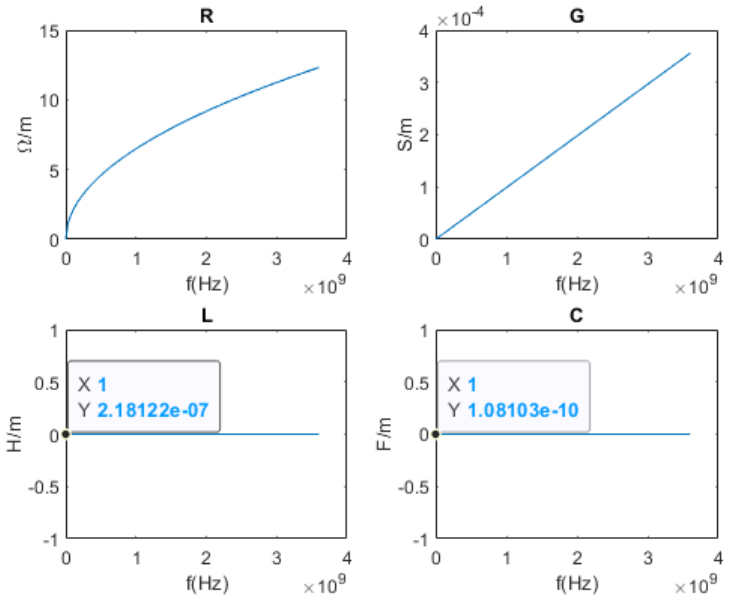


Figure 3.4 – Variation of estimated cable $RLCG$ parameters as a function of the frequency

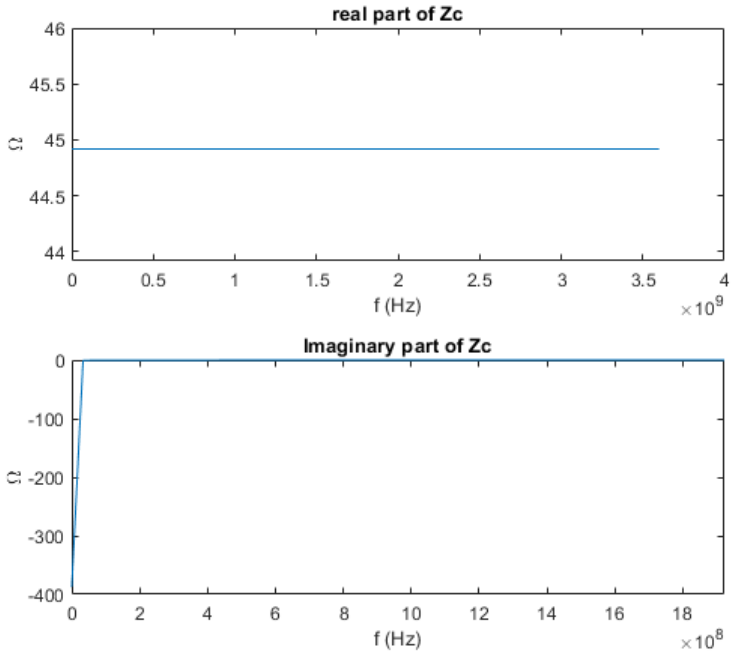


Figure 3.5 – Variation of the characteristic impedance Z_c as a function of the frequency

the Gaussian pulse, used as a test signal, is $250ps$. The inset in this figure represents the difference between the two reflectograms resulting from the calculated $RLCG$ model and the 3D model. The correlation coefficient is used to evaluate the similarity between the two models. The correlation coefficient between the two reflectograms resulting from the calculated $RLCG$

model and the 3D model equals 0.9584. Here, it is considered that the data is strongly correlated if the correlation coefficient is between 0.9 and 1 [15]. One can thus consider that the correlation coefficient is strong. It validates the model with distributed constants. Therefore, the computed R, L, C, and G parameters of the healthy cable will be used for soft fault characterization in the next section.

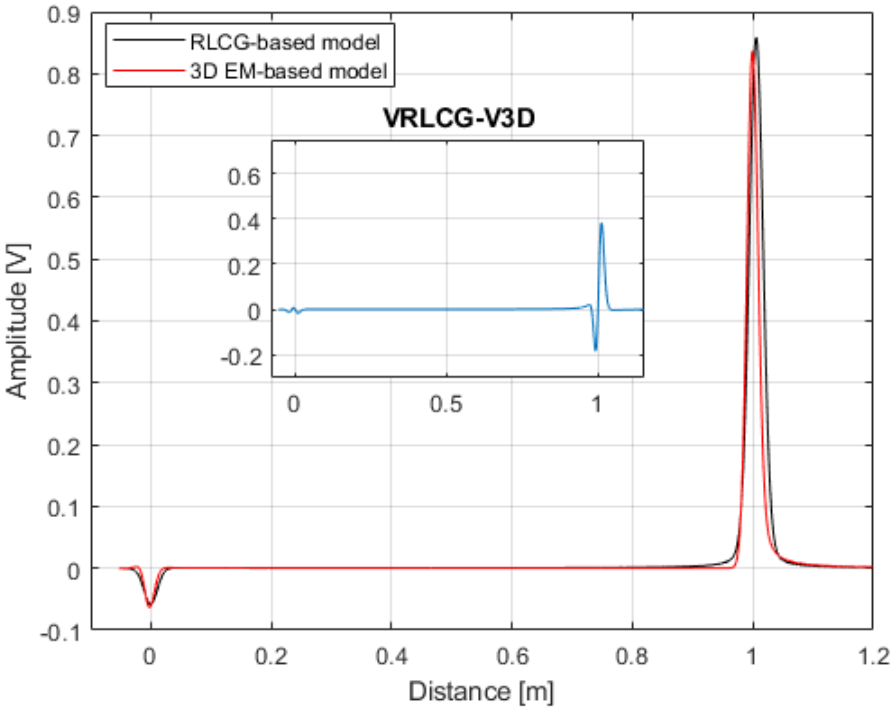


Figure 3.6 – TDR reflectograms of the $RLCG$-based model and 3D EM model

3.2.3 . Soft Fault Modeling and Simulation

In Section 1.3, an assessment of the most prominent faults in electrical wiring has been made. The Federal Aviation Agency (FAA), Naval Systems Air Command (NAVAIR), and National Aeronautics and Space Administration (NASA) have all identified wire chafing as the most significant factor contributing to electrical wiring and interconnect system failures [7]. This fault type is a precursor to more significant problems such as open and short circuits, and arcing, which cause smoking, fires, etc. In this context, our case study will use this fault type.

TDR is an industry-standard method used for detecting chafing faults. Given the input signal, the cable dimensions, and materials, the model estimates the signal reflected by the wire. This estimation is derived from Maxwell's equations using numerical methods and minimal assumptions.

This part aims to study this type of degradation, which is among the most representative of the faults encountered. We will focus on the impact of the degradation of the cable shielding, taking into account the three-dimensional characteristics of the faults. Figure 3.7 shows a

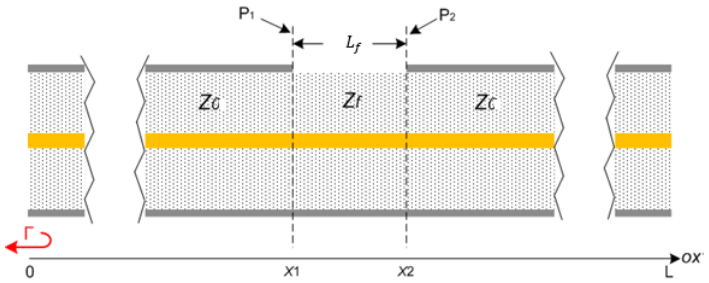


Figure 3.7 – Coaxial cable model with a shielding fault between x1 and x2

coaxial transmission line with a discontinuity on the outer conductor of length L_f. The model presents the damaged area located between x1 and x2.

As the size of the fault and the degree of opening on the shield increase, the radiation losses due to the propagation of waves through the opening of the shield may also increase. The more severely the cable shielding breaks, the closer the fault will approach an open circuit. The faulty area will then have an impedance Z_f different from the cable's characteristic impedance Z_c. Therefore, when the incident wave enters the faulty zone, there will be reflections (at the beginning and end of the fault) with positive and then negative signs on the P_1 and P_2 interfaces (Figure 3.7), respectively.

The coaxial cable modeled in section 3.2.1 is degraded (a portion of the shield is removed). A representation of the faulty coaxial cable is depicted in Figure 3.8.

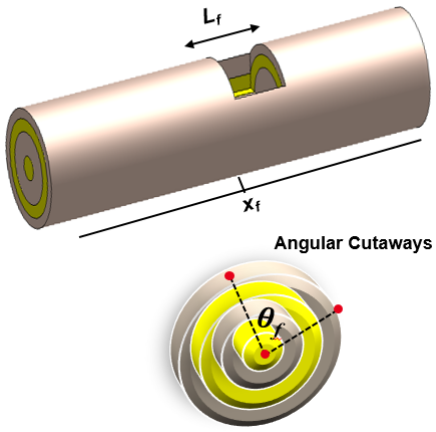


Figure 3.8 – Coaxial cable with a shielding damage model

The fault is characterized by three different parameters: the position x_f =x1, the length L_f and the angular cutaways θ_f. We will seek to understand the influence for each parameter on the shape of the reflectogram. We consider here the case where only the cable's shield is degraded with different widths (θ_f) and constant length (L_f). Holes ranging from $\theta_f = 45°$ to

$\theta_f = 180°$ in width are created in the shielding for individual simulations to emulate the growth of experimentally created shield flaws.

The considered cable length is $l = 1m$ and a soft fault occurs at $x_f = 0.5m$ from the beginning of the transmission line, with length $L_f = 5mm$, and various widths (θ_f) of $45°$, $90°$, and $180°$, as shown in Figure 3.8. The simulations consist of injecting a Gaussian pulse U_i, of unit amplitude, in the frequency band $[0; 4GHz]$, into one extremity of the wire and then recording the reflected signals at the injection point. The reflectometry response will reveal the disturbance generated by the fault on the signal propagation along the line. Figure 3.9 presents the TDR response for the faulty case of $\theta_f = 180°$. Due to the fault presence, when the incident signal arrives at the discontinuity of length L_f between the points x1$= 0.5m$ and x2$=0.505$, multiple reflections between the two interfaces and the defective zone are produced. The signal amplitude at the end of the cable illustrates that the incident signal still propagates along the line but with a lower amplitude due to the reflections in the faulty zone.

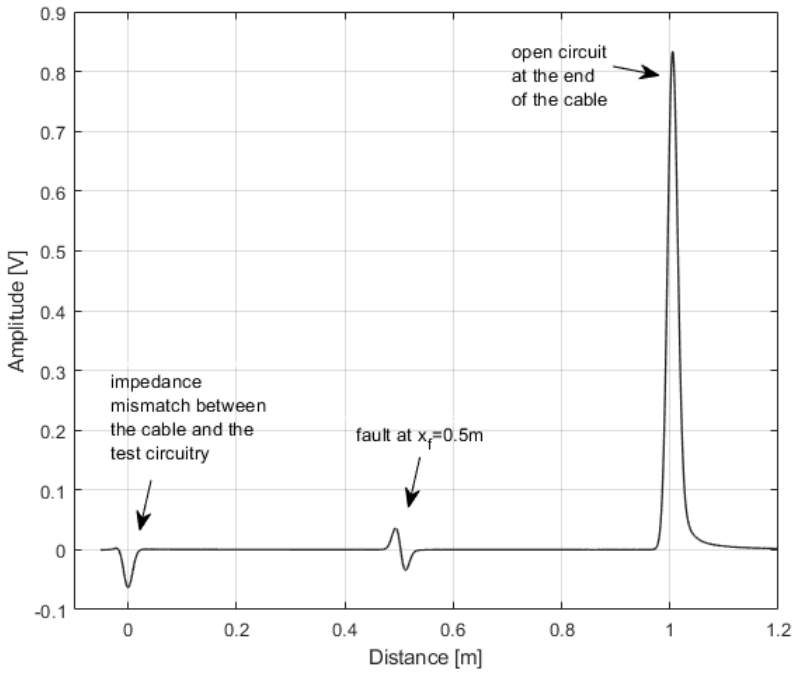


Figure 3.9 – TDR response of a shielding fault with $x_f = 0.5m$, $L_f = 5mm$ and $\theta_f = 180°$

Figure 3.10 illustrates the TDR fault signature as a function of the angular cutaways of the damaged shield. The shape of the reflected signal exhibits positive and negative reflections due to the sign of the reflection coefficient at the discontinuity interfaces at P1 and P2. The shape and amplitude of the time domain response due to the fault depend on the nature and severity of the fault. Moreover, as shown in this figure, the amplitude of the reflected waves is generally very small (here about 5% of U_i), and therefore, the reflected pulse cannot be detected easily by means of the measurement devices. The reflected signal can change significantly due to the radiation leakage that increases with the width of the shielding fault θ_f.

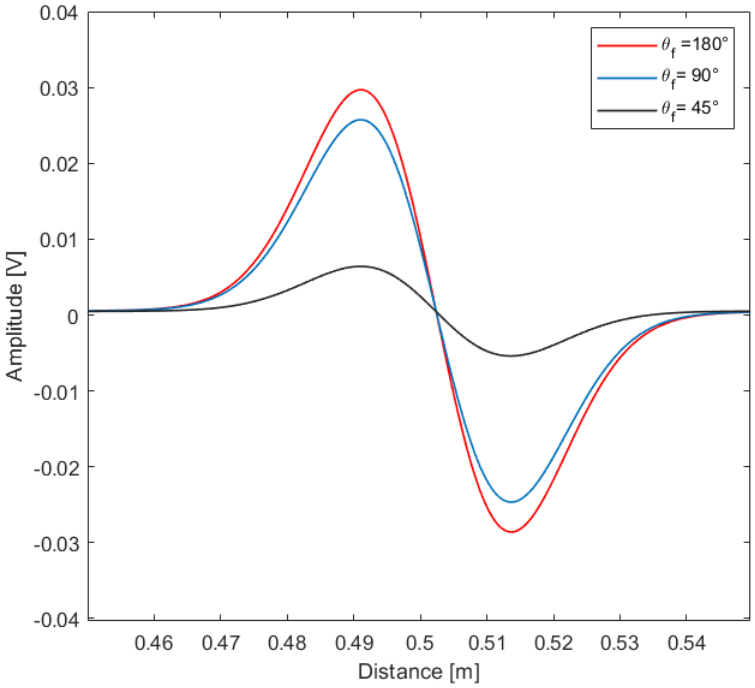


Figure 3.10 – Shielding fault signatures for different fault width's at $x_f = 0.5m$ and $L_f = 5mm$

In CST Studio Suite, the accuracy of the solution is strongly related to the number of meshes. As a consequence, for long cables ($\approx 100m$ cables), the number of meshes is very high, which increases the computational burden and the simulation time. Therefore, at first, CST will be used for the $1m$ faulty coaxial cable modeled in this section. Then the fault parameters are extracted using a characterization method [15], and are implemented in a Mathwork's Matlab©language code to simulate different scenarios where the following parameters are changed:

- The cable length l
- The fault position x_f

The simulation results of this method are compared to the output of the CST simulator with a high degree of success, and the Matlab code was tested for its ability to retrieve (estimate) the TDR response. The detailed process of the characterization method has been presented in chapter 2, section 2.5.1. The first step has been done in section 3.2.2 where the $RLCG$-based model is developed and validated. The $RLCG$-based model TDR response in the flowchart Figure 3.6 is V_{sim} defined in the Figure 2.11 in chapter 2: the reflectometry response containing the contribution of the healthy NUT. Second, we propose to study the effect of increasing the width θ_f and the length L_f of the shielding fault on the $RLCG$ parameters. The three faults previously defined in this section are used, with a constant length of the fault ($L_f = 5mm$). Figure 3.10 is used as the V_{mes} in the flowchart: the reflectometry response of the faulty NUT containing the signature of the soft faults. For these simulations, the maximum frequency

of the signal is $4GHz$. Third, the baselined signature of the three fault cases is obtained. Fourth, based on the electrical faults parameters obtained from the Genetic Algorithm (GA), the impedance Z_d of the fault is calculated. From the results in Table 3.2, we can observe an increase of the variation of the characteristic impedance and the reflection coefficient when the fault width increases.

Table 3.2 – The estimated parameters of the shielding damage for different widths

Width of the fault	x_f	L_f	ΔZ_c	Γ
$\theta_f = 45°$	$0.5m$	$5mm$	5Ω	0.0526
$\theta_f = 90°$	$0.5m$	$5mm$	23Ω	0.2987
$\theta_f = 180°$	$0.5m$	$5mm$	27Ω	0.3698

The results in Table 3.3 show the absence of the impact of the fault length on the impedance variation when the fault width is set constant at $45°$. Nevertheless, the increase of the fault length L_f increases the amplitude of the fault signature based on equation (1.35).

Table 3.3 – The estimated parameters of the shielding damage for different lengths

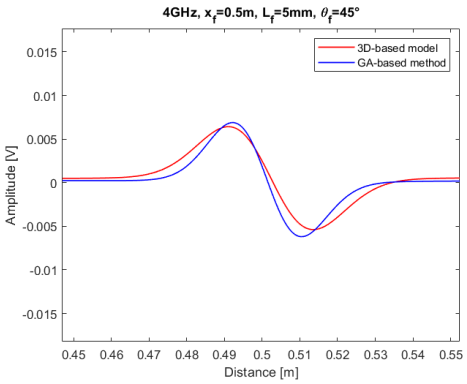
Length of the fault	x_f	θ_f	ΔZ_c
$L_f = 5mm$	$0.5m$	$45°$	5Ω
$L_f = 10mm$	$0.5m$	$45°$	5Ω
$L_f = 20mm$	$0.5m$	$45°$	5Ω

To sum up, the characterization method within the Matlab code was tested against the commercial 3D simulator (CST) by simulating a cable which had a shielding fault. The TDR results from CST as well as the Matlab code for three different widths are depicted in Figure 3.11. The correlation coefficients for the different cases are $= 0.9841$, 0.9839 and 0.9797 respectively. The outputs of both the simulator and characterization method code are nearly identical (from the correlation coefficient calculations). These encouraging results prove that we can retrieve the TDR response for longer cables from the Matlab code, which is far less time consuming than using CST.

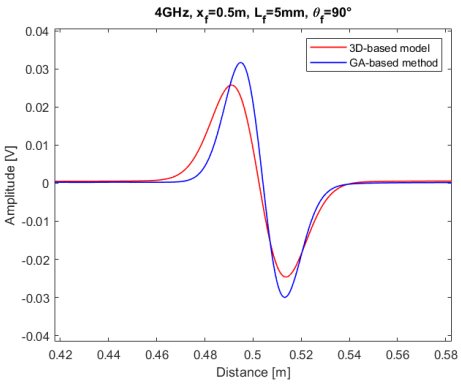
Figure 3.12 shows the TDR fault signatures at $4GHz$ for the following fault ($\theta_f = 180°$ and $L_f = 5mm$) located at $x_f = 0.5m$, $5m$ and $50m$, for $1m$, $10m$ and $100m$ cables, respectively. Table 3.4 shows the amplitude variation of the soft fault signature with the change in the fault position along the cable. It illustrates the attenuation phenomenon explained in chapter 2 and hence, the selection of frequency issue addressed in the following sections.

3.2.4 . Model Validation with Experimental Results

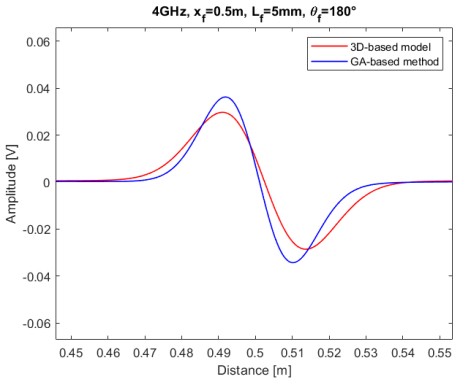
This section aims at validating the developed models with experimental results. Experimental results are encouraging for finding shielding flaws before any disruptive damage of the inner conductors. The experimental data collection starts with focusing on the coaxial cable



(a)

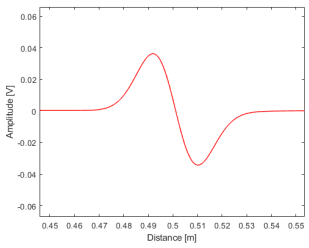


(b)

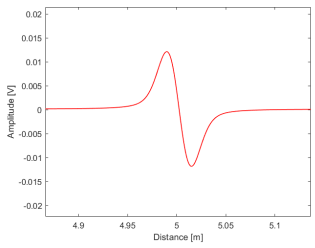


(c)

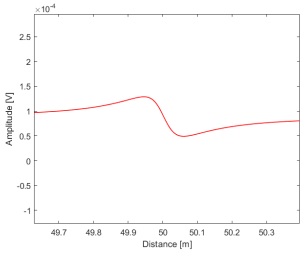
Figure 3.11 – Signatures of the reconstructed and simulated fault with width (a) $\theta_f = 45°$, (b) $\theta_f = 90°$ and (c) $\theta_f = 180°$



(a) $1m$ cable



(b) $10m$ cable



(c) $100m$ cable

Figure 3.12 – Fault signature for different cable lengths

Table 3.4 – Variation of the Fault ($\theta_f = 180°$ and $L_f = 5mm$) amplitude with its position along the cable

Cable length l_f [m]	1	10	100
Fault position x_f [m]	0.5	5	50
Amplitude [V]	0.036	0.012	0.00012

type, measuring the TDR response of the healthy cable, then the cable is damaged using an abrasive apparatus to chafe a small section of the wire and finally measure the TDR response of the faulty cable. In the study described here, we are interested in how the TDR signal changes as a function of the width of chafing into the shielding. At $\theta_f = 180°$, the shielding has worn away sufficiently to expose the inner core conductor insulation.

The measurement was carried out using an RG-316 coaxial cable ($Z_c = 50 \mp 2\Omega$ at $4GHz$). The exterior insulation on the cable has a relative permittivity of around 2.1. In this study, we are not interested in measuring scrapes of the exterior insulation, but instead, we are interested when chafing starts to create holes in the shielding. In this part, the results of the simulation with CST are compared to the experimental ones.

A $5mm$ long, $180°$ wide shielding damage is created at $0.5m$ from the injection point as shown in Figure 3.13. The length of the Cable Under Test (CUT) is $1m$. In Figure 3.14, the

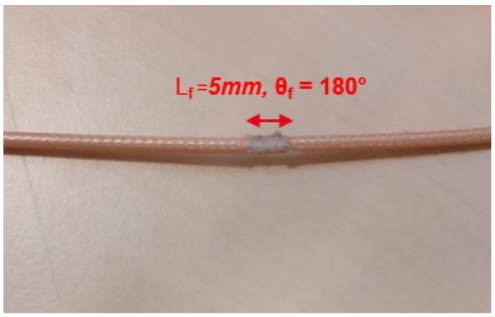


Figure 3.13 – Experimental chafed cable

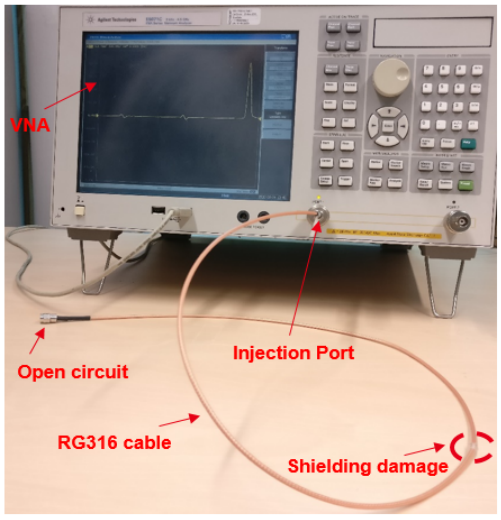


Figure 3.14 – Experimental Setup

cable input terminal was attached with a 50Ω N series RF coaxial cable connector. The other end is left open-ended. The reflected signals and the corresponding reflectograms are obtained using a VNA Agilent E5071C, with calibration to acquire highly accurate measurements. A

TDR signal is considered on a total bandwidth defined from DC to a maximal frequency f_{max} ($1GHz$, $2GHz$, $3GHz$ and $4GHz$).

Figure 3.15 represents reflections from TDR Gaussian input by the digital simulation in CST and the experiment in a no-fault case. Figure 3.16 represents reflections from TDR Gaussian

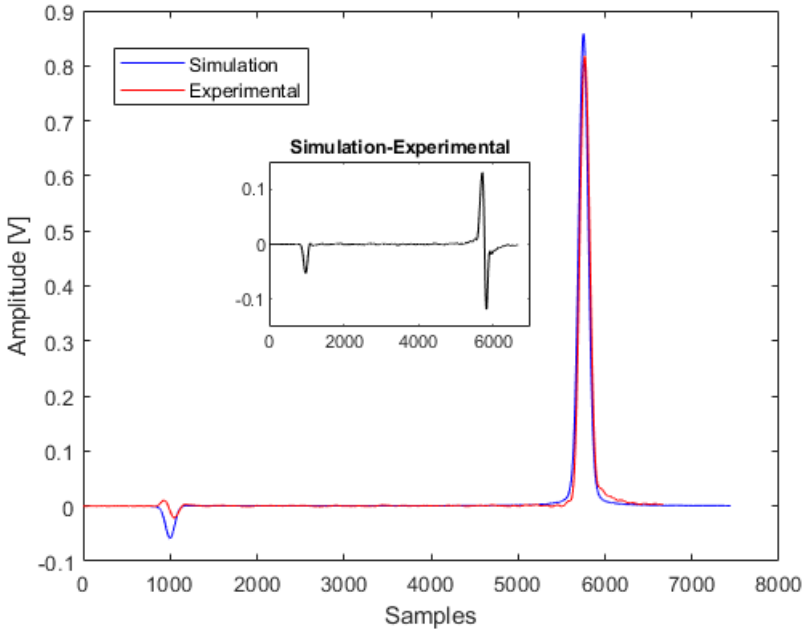


Figure 3.15 – Healthy TDR responses

input by the digital simulation in CST and the experiment in the case of the shielding fault defined in Figure 3.13. The inset in this figure represents the difference between the simulation and the experimental TDR responses. The difference at the input is due to the presence of a connector and a generator in the experiments connected to the cables' input, whereas in simulations, only a generator is used without a connector. Three sections are considered in the experimental and simulation TDR responses: input, faulty part, and the end (here, open circuit). Thus, $s(t)$, defined in chapter 1 (1.2) as the reflected signal, is divided into three sections. The spectral energy density of each section is given by E_s and it is calculated using equation (3.2) and is presented in Table 3.5. It shows that the energy of the input section is identical for both the experimental and the simulation responses, regardless of the difference in the signature shape and its amplitude. The difference in the amplitude at the cable end is due to the open circuit impedance in the experiment, which is different from the infinite value used for the simulations.

$$E_s = |X(f)|^2 \tag{3.2}$$

where $X(f)$ is the Fourier transform of $s(t)$. Since the signal $s(t)$ is in $volts$, then the dimensions of X(f) would become $[volts * seconds]$ and E_s would be in $[v^2 * s^2]$, so this gets us to $[power * s/Hz]$, that is $[J/Hz]$.

The Root Mean Square Error (RMSE) between the two reflectograms in Figure 3.16 is equal to 0.024. This value is relatively low since the observations mean is 0.168. The RMSE based only on the fault signature section is 0.001. Since it constitutes only 4% of the total

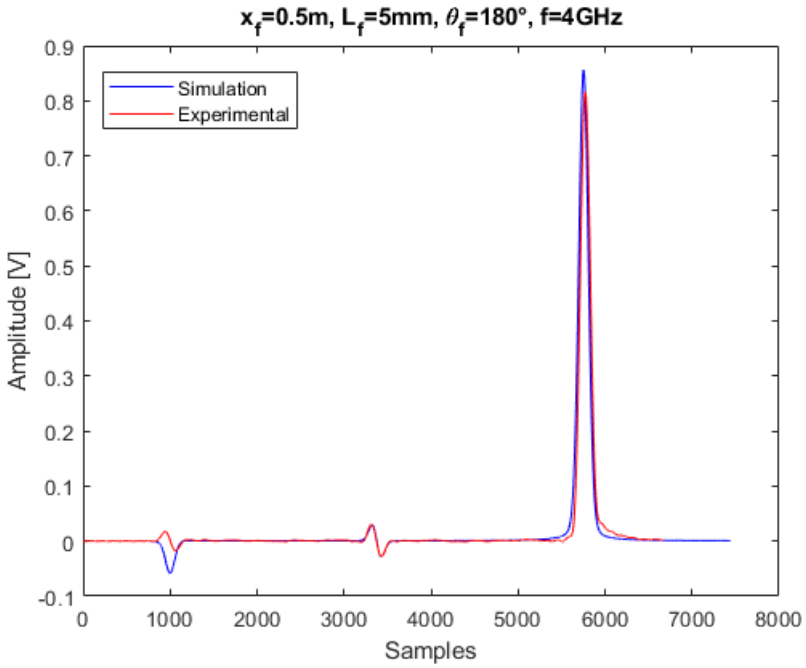


Figure 3.16 – Faulty TDR responses

RMSE (0.024), the majority of the error is due to the output section. This plot shows that we are within the same range of magnitude between the simulation and the experiments.

Table 3.5 – Energy calculations

Energy $[J/Hz]$	**Input**	**Fault (middle)**	**Cable end**
Simulation	0.31	0.11	71.65
Experimental	0.31	0.11	67.06

Now that the model is validated, it can be used to generate the data used to evaluate the frequency selection approach, presented in the following sections.

3.3 . Frequency Selection Algorithm for Soft Fault Diagnosis

In practice, the expert configures, calibrates a VNA, selects the excitation frequency and records healthy cable measurement. Measurements at the same frequency are done on a faulty cable. These four steps are repeated as long as the fault is not detected. Therefore, there is a loss of information and time in addition to the subjectivity of the decision-making.

Our objective is to propose a method for selecting the best test signal frequency for soft fault detection. It performs measurements at different frequencies for the healthy case. After which, the PCA model is established. Then, new measurements are performed at different frequencies. If a difference is detected between the model and the new test data, the contribution of each variable (i.e., frequencies) to this difference is calculated. The algorithm then chooses the

most relevant frequency to monitor the soft fault. The advantages are thus time saving and objectivity of the decision-making. Furthermore, this methodology could be able to monitor the evolution of detected faults in the prognosis perspective.

The first step here is to have the healthy operating data upon which the PCA model will be established. This model is then used to evaluate new measurement data. Two statistical methods are used here; the Q (or SPE) and the Hotelling T^2 statistics [121]. If a fault is detected, the best frequency corresponding to the ones with the highest contribution in the detection is then selected. Figure 3.17 shows the flowchart of the proposed method, which

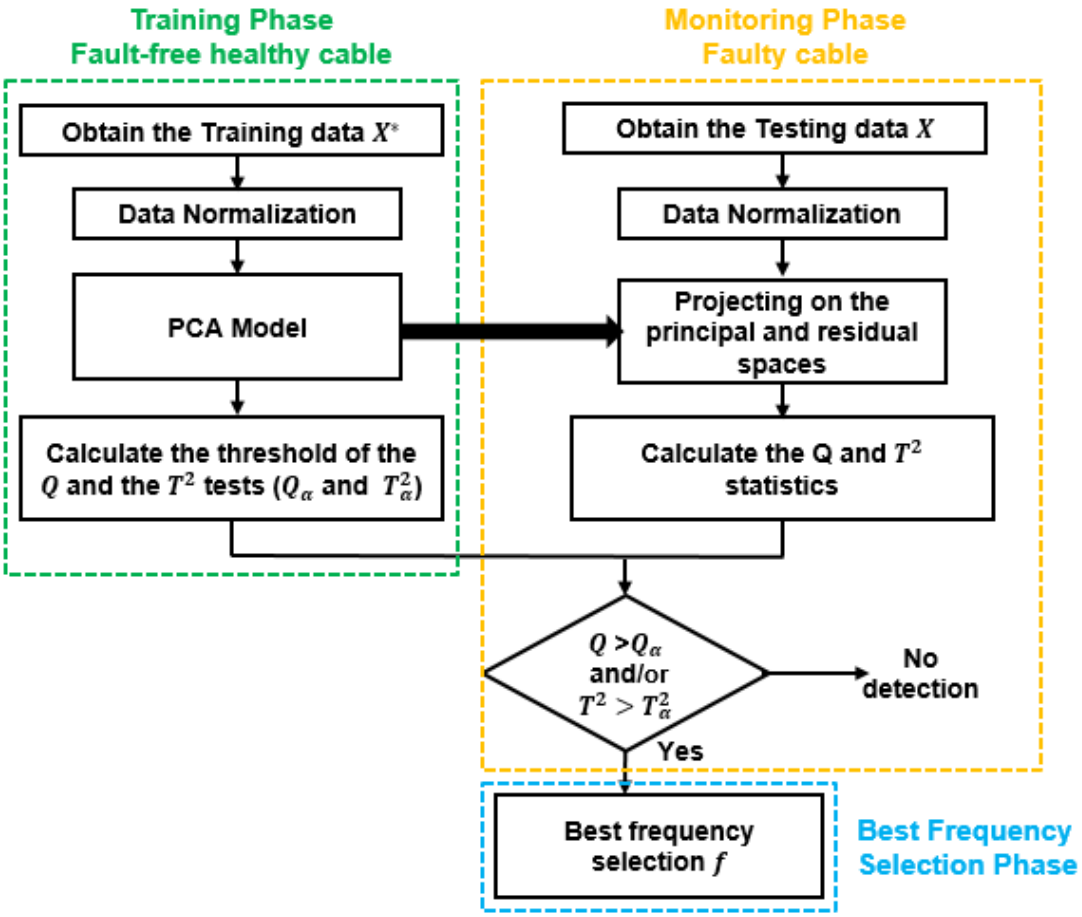


Figure 3.17 – Flowchart of the proposed approach

consists of the following steps:

Step 1: Collect training data representing healthy process operations and scale(normalize) them using their mean and standard deviation (Training phase).

Step 2: Develop a statistical PCA model for the system using the singular value decomposition (Training phase).

Step 3: Define a confidence level α and calculate the upper control limits for Q and T^2 statistics (Training phase).

Step 4: Acquire new sample measurement and scale it using the same factors (mean and standard deviation) from step one (Monitoring phase).

Step 5: Generate Q and T^2 statistics based on the obtained PCA model (Monitoring phase).

Step 6: Decide if the new sample is considered faulty when one or more residuals exceeds the threshold (Monitoring phase).

Step 7: Inspect the inputs (original variables) that highly influence the residual. Contribution plots are used for this purpose, and the best frequency is then selected (Best frequency selection phase).

This method consists of three phases:

1. Training phase, where data X^* are collected during fault-free operation, and the PCA model is developed as in section 2.6.1. After developing a model using healthy (training) data, the reduced dimension model can be used to detect and diagnose abnormalities.
2. Monitoring phase, i.e., fault detection, is handled using the monitoring statistics T^2 and Q test. For fault detection, the developed PCA model based on healthy operating system data is used to check new measurement data X as explained in section 2.6.2. If the Q or T^2 value falls outside the confidence limit at a specific sample, then there is an abnormality.
3. Best frequency selection phase, where fault diagnosis will be managed through contribution plots. Since the reference matrix, X^* is built so that each variable corresponds to a specific frequency; then one can inspect the inputs (frequencies in this case) that highly influence the abnormal sample Q or T^2 value. After which, the most relevant frequency to monitor the detected soft fault is selected. For this purpose, contribution plots are used. It is necessary to find the variable (frequency) that contributes the most to the Q or T^2 value. We can look at the contribution of each input to this large statistical value. This analysis can determine which variable is responsible for the unusual behavior of Q or T^2 and, hence, the most relevant frequency is retained.

3.4 . Simulation Results

3.4.1 . Short Cable Analysis

3.4.1.1 Training phase

The $1m$ 3D coaxial cable model developed using CST in section 3.2.1 is used. The healthy cable reflectograms are simulated for four different bandwidths. Each bandwidth ranges from DC to a maximal frequency f_{max}. The cable is excited using a Gaussian pulse with different maximal frequencies ($1GHz$, $2GHz$, $3GHz$ and $4GHz$).
For our proposed methodology, we first consider the reference, healthy performance, representation of the data given by (3.3). X^* is formed up of four variables (step 1). Each variable $\mathbf{R_f}$

is a column vector of the matrix and corresponds to the cable healthy TDR response at the frequency $f = 1GHz$, $2GHz$, $3GHz$ and $4GHz$ respectively.

$$X^* = [\mathbf{R_{1GHz}} \quad \mathbf{R_{2GHz}} \quad \mathbf{R_{3GHz}} \quad \mathbf{R_{4GHz}}] \tag{3.3}$$

The obtained data matrix X^* is then used for the construction of the PCA model according to (2.8), (2.9) and (2.10) (step 2).

Table 3.6 indicates that the cumulative variance of the first two scores is 98.64% that is greater than the lower limit (i.e. 90%). This implies that the observed variables are highly correlated. Using (2.11), the data are well described with the first two principal components. Thus, the number of the retained components, k, is equal to two. For step 3, the 95%

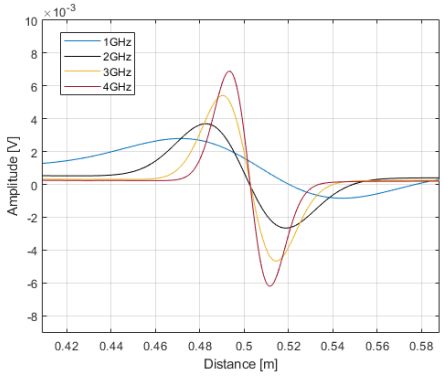
Table 3.6 – PCA model variances

PC number	Variance percentage (%)	Cumulated percentage (%)
1	95.37	95.37
2	3.27	98.64
3	1.03	99.68
4	0.31	100.00

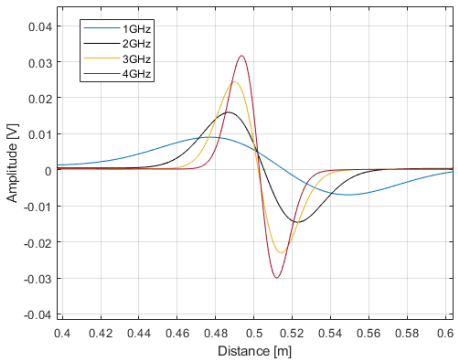
confidence limits of those tests are calculated according to ((2.16)-(2.19)). Thus, $Q_\alpha = 22.25$ and $T^2_\alpha = 26.92$.

3.4.1.2 Monitoring phase

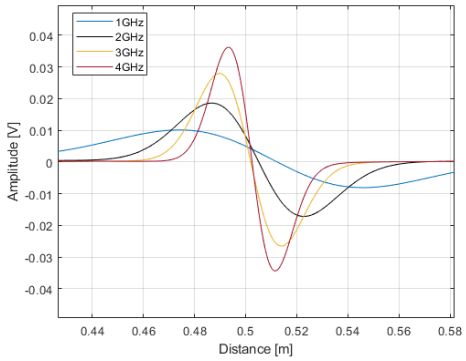
Second, the soft fault model in section 3.2.1 is used, where the angular cutaways (fault width) parameter θ_f is set to one of the three values: $45°$, $90°$ or $180°$. We consider the case where only the sheath of the cable is degraded with different widths, and the length of the fault remains constant ($5mm$). Figure 3.18 (a), (b) and (c) shows the signature of the shielding damage with angular cutaways $\theta_f = 45°$, $\theta_f = 90°$ and $\theta_f = 180°$ respectively, for the different maximal frequencies.



(a) $\theta_f = 45°$



(b) $\theta_f = 90°$



(c) $\theta_f = 180°$

Figure 3.18 – Soft fault signatures at four different excitation frequencies

We observe the impact of varying the width of the fault on the signal reflection. The figure illustrates that as the fault's width increases, the incident wave will be more affected by the fault. Since velocity changes with frequency as described in (2.3), a shift between the signatures of the same fault at different frequencies is observed on the reflectograms. As shown in Figure 3.19, both the fault severity and the excitation frequency lead to higher peak values.

Step 4 corresponds to the acquisition of a new measurement data set X in faulty case. It has four vectors such that $X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3} \quad \mathbf{x_4}]$. Each of the four variables is a concatenated vector of the fault signature data in the three considered cases ($\theta_f = 45°, \theta_f = 90°$ and $\theta_f = 180°$), and for the same operating frequency. The new data matrix is defined as:

$$X = \begin{bmatrix} \mathbf{R_{45°_{1GHz}}} & \mathbf{R_{45°_{2GHz}}} & \mathbf{R_{45°_{3GHz}}} & \mathbf{R_{45°_{4GHz}}} \\ \mathbf{R_{90°_{1GHz}}} & \mathbf{R_{90°_{2GHz}}} & \mathbf{R_{90°_{3GHz}}} & \mathbf{R_{90°_{4GHz}}} \\ \mathbf{R_{180°_{1GHz}}} & \mathbf{R_{180°_{2GHz}}} & \mathbf{R_{180°_{3GHz}}} & \mathbf{R_{180°_{4GHz}}} \end{bmatrix} \tag{3.4}$$

where $\mathbf{R}_{\theta°_f}$ describes the fault signature data vector with angular cutaway θ_f and frequency f.

The PCA reference model obtained in section 3.4.1.1 is used to project the new measurement data. The differences between the projection $\hat{X}$ and X are evaluated through the Q and Hoteling's T^2 statistical tests. New measurements X will be projected into the framework

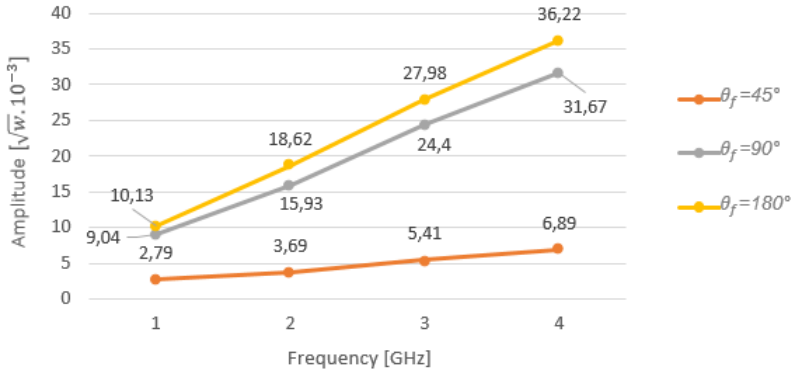


Figure 3.19 – The different fault signature positive peaks (maximum value)

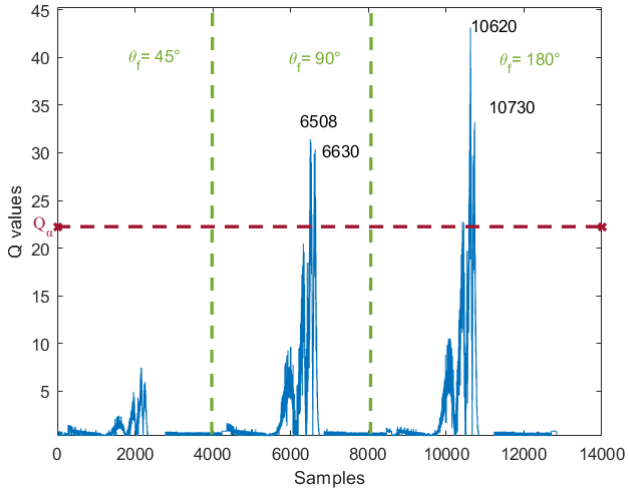


Figure 3.20 – Scores plot: Q values of new measurement samples

spanned with the loading matrix P_k. The new scores T_{new} and the residual $\hat{T}_{new}$ are calculated according to (2.12) and (2.13). According to ((2.12)-(2.15)), the Q and the T^2 values for each new measurement sample are calculated (step 5). Figure 3.20 and Figure 3.21 present the Q and the T^2 control charts respectively, with the dashed red line representing the 95% confidence limit. It is observed that Q_α is crossed by several samples (6508, 6630, 10620 and 10730) and T^2_α by the samples (10620 and 10730). This indicates that faults have occurred (step 6).

3.4.1.3 Best Frequency Selection phase

The contributions to the Q statistics are plotted in Figure 3.22. It can observed that the variable $\mathbf{x_4}$ has the highest contribution to the unusually large Q statistic. Thereby, the selection of the relevant frequency is performed by Q and Hotelling T^2 control tests (step 7).

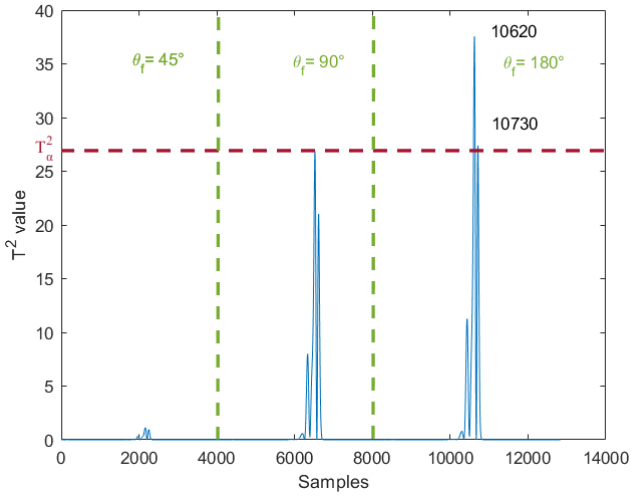


Figure 3.21 – Scores plot: T^2 values of new measurement samples

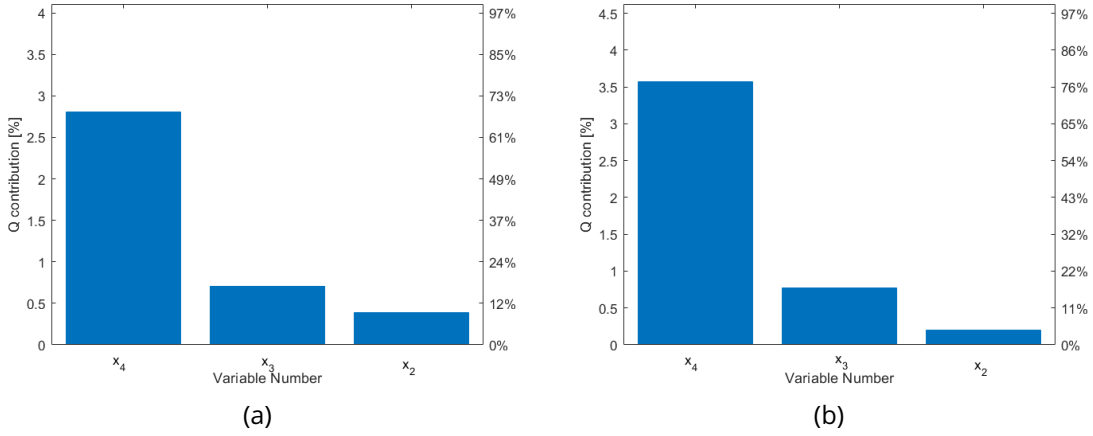


Figure 3.22 – Contribution plot of the samples: (a) 6508 and (b) 10620

The first occurrence of the fault has the highest impedance variation in the reflectogram, and for each round-trip of this fault, its peak amplitude on the reflectogram decreases. Due to the presence of several round-trip peaks for the same fault, several abnormal samples will be produced in the Q and T^2 chart. Therefore, we consider the samples with the highest Q and T^2 values as the abnormal ones corresponding to the fault. Hence, according to Figure 3.20 and Figure 3.21, samples 6508 and 10620 are investigated. The contribution of the first variable $\mathbf{x_1}$ is almost neglected with respect to the other variables; hence, it does not appear in the plots. Sample 6508 in Figure 3.22 (a) is detected only by the Q test whereas sample 10620 in Figure 3.22 (b) is detected by the two tests.

If we look at Figure 3.20 (Q chart) and Figure 3.22, we can draw the following concluding remarks:

- for $\theta_f = 45°$, the fault cannot be detected whatever the frequency.

- for $\theta_f = 90°$ and $\theta_f = 180°$ the faults are detectable and $f = 4GHz$ is the best selected frequency to monitor this fault.

However, Figure 3.21 (T^2 chart) and Figure 3.22 show that the fault is only detectable for $\theta_f = 180°$ and the best frequency is $4GHz$. Thanks to the simulation results of a $1m$ length RG316 coaxial cable with a shielding damage with 3 severity levels ($45°, 90°$ and $180°$), the combination of reflectometry and PCA coupled to Q and T^2 statistics shows that:

- the highest frequency ($4GHz$) leads to the best fault detection capability for the two largest severities if Q test is used and the largest severity detection if T^2 test is used.
- for the lowest fault severity, the Q and T^2 in the PCA framework fail to detect the fault.
- Q test is more suitable than T^2 to be used in our study since it can detect more faults. Thus, in the following sections, only Q test will be used.

The simulation results are in coherence with the well known rule that for short cables, the higher the frequency is, the better it is for the fault detection. This approach will be applied to another set of cables with different operating conditions and fault types in the following sections.

3.4.2 . Long Cable Analysis

3.4.2.1 Data-base Building

The process of data collection simulation is divided into two parts where CST model is used for short cable and $RLCG$ model will be used for longer cables:

(1) CST simulations: for the cable defined in 3.2.1 of length $l = 1m$ (the same cable used in section 3.4.1.1, we simulate fault cases characterized by three parameters at four different frequencies $f = [1GHz, 2GHz, 3GHz,$ and $4GHz]$. We have in total nine total cases for each frequency:

   (a) Position $x_f = 0.5m$

   (b) Length $L_f = [5mm, 10mm, 20mm]$

   (c) Width $\theta_f = [45°, 90°, 180°]$

(2) Usually, the location of the fault is unknown; the fault length is also unknown. It is necessary to analyze the sensitivity of the proposed method for different values of fault length to evaluate its robustness. For that, the fault parameters are extracted using the characterization method described in section 3.2.3 and are implemented to simulate different fault cases for which the cable length l and the fault position x_f are changed:

   - Extract the R, L, C, G parameters from the healthy cable data
   - Derive ΔZ_d for the nine fault cases defined in (1)

- Insert the signatures in the simulation Matlab code
- Choose a cable of length $l = 10m$
- Simulate fault cases characterized by three parameters (36 total cases):
  (a) Position $x_f = [20\%, 50\%, 70\%, 90\%]$ (vary the fault position along the cable length in $\%$)
  (b) Length $L_f = [5mm, 10mm, 20mm]$
  (c) Width $\theta_f =$ [45°, 90°, 180°]

A significant source of problems in reflectometry analysis is the so-called "blind zone." This is particularly problematic for wires or cables that are very short or when the fault is near the front of the cable. This is caused by the reflected signal overlapping the incident signal because the time delay is small. This makes it challenging to identify the reflected signal. Several methods can be used to mitigate this effect. For example, using a longer test lead to connect the reflectometer to the wire under test. This would effectively delay enough the reflected signal reducing or eliminating the overlap. This may be practical for handheld applications, but not for industrial applications, where the reflectometer is actually embedded in the system. Another method is to use baselining to identify the overlapping signals and extract the reflected response [134]. The blind zone at the start and the end of the cable in our study is 5%. The baselining approach is used to deal with this problem.

Now, the reference TDR response for the $10m$ cable at the different frequencies ($1GHz$, $2GHz$, $3GHz$,$4GHz$) is simulated. In addition, for the 36 faulty cases, the corresponding TDR response at the same frequencies is also obtained.

3.4.2.2 Training phase

For our proposed methodology, we first consider the reference, healthy data given by (3.5). X^* is formed up of four variables. Each variable $\mathbf{x}_\mathbf{j}^*$ corresponds to the $10m$ cable healthy TDR response at the frequency $f = 1GHz$, $2GHz$, $3GHz$ and $4GHz$ respectively.

$$X^* = [\mathbf{x_1^*} \quad \mathbf{x_2^*} \quad \mathbf{x_3^*} \quad \mathbf{x_4^*}] \tag{3.5}$$

The obtained data matrix X^* is then used for the construction of the PCA model according to (2.8), (2.9) and (2.10).

Table 3.7 indicates that the cumulative variance of the first score is 99.9412% that is greater than the lower limit. This implies that the observed variables are highly correlated. Using (2.11), the data is well described by one principal component. Thus, k is equal to one.

3.4.2.3 Monitoring phase

Our methodology's second step leads to considering a new measurement data set X corresponding to the faulty cases. It has four vectors such that $X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3} \quad \mathbf{x_4}]$. Each

Table 3.7 – PCA model variances

PC number	Variance percentage (%)	Cumulated percentage (%)
1	99.941	99.941
2	0.0575	99.9985
3	0.0014	99.999
4	3.5377e-05	100.000

variable is a concatenated vector of the fault signature data for the 36 use cases and the same operating frequency. The new data matrix is defined as:

$$X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3} \quad \mathbf{x_4}] \tag{3.6}$$

where the variable $\mathbf{x_j} = [F1 \quad F2 \quad ... \quad F36]$ describes the fault signature data vector at frequency j such that $F1$ corresponds to the fault case 1.

The constructed PCA reference model in section 3.4.2.2 is then used to check the new measurement data. To do so, the differences between the new measurement data and their projections into the constructed model are then subjected to the Q test. The 95% confidence limit of this test is calculated according to (2.16). Thus, $Q_\alpha = 0.0118$.

According to ((2.12)-(2.14)), the Q value for each new measurement sample is calculated. Figure 3.23 presents the Q control chart, with the dashed red line representing the 95% confidence limit. The Q value for each detected fault case is given in the Appendix A. It is observed that several samples have crossed Q_α. This indicates that faults have occurred. 28 fault cases out of the 36 are detected. Table 3.8 shows the fault cases with the varying position along the cable, their severities, and if they are detected or not.

3.4.2.4 Best Frequency Selection phase

According to Figure 3.23, 28 samples are detected. We use the contribution plot of the abnormal sample related to each case of the detected cases to select the best frequency. For example, for the abnormal sample 34280 corresponding to the fault case $F9$ ($L_f = 20mm$,$\theta_f = 180°$), the contributions to the Q statistics is plotted in Figure 3.24: $\mathbf{x_4}$ had the highest contribution, so the highest frequency ($4GHz$) leads to the best fault detection capability in this case. Table 3.8 shows the selected frequency for the detected fault cases.

For the fault cases at $x_f = 90\%$, the selected frequency is $3GHz$ for the two severities $F30$ and $F33$ and $2GHz$ for $F35$ and $F36$. Figure 3.25 shows the contribution plots of the abnormal sample corresponding to each one of these four fault cases. In order to explain the obtained results, the energy of the faulty part at each frequency for each of the indicated fault case is presented in Table 3.9. E_f corresponds to the fault energy at f. It is noticed that the method chooses the frequency with the highest energy.

The combination of reflectometry and PCA coupled to Q statistic shows that:

- For $x_f = 20\%$, all the faults are detected whatever their severity level, and the selected

Table 3.8 – Fault detection by PCA for different fault scenarios for a $10m$ cable

Fault cases	Parameters(x_f,L_f,θ_f)		Detected/Not detected	If detected, selected f
$F1$	$x_f = 20\%$	$L_f = 5mm$ $\theta_f = 45°$	Yes	4GHz
$F2$		$L_f = 5mm$ $\theta_f = 90°$		
$F3$		$L_f = 5mm$ $\theta_f = 180°$		
$F4$		$L_f = 10mm$ $\theta_f = 45°$		
$F5$		$L_f = 10mm$ $\theta_f = 90°$		
$F6$		$L_f = 10mm$ $\theta_f = 180°$		
$F7$		$L_f = 20mm$ $\theta_f = 45°$		
$F8$		$L_f = 20mm$ $\theta_f = 90°$		
$F9$		$L_f = 20mm$ $\theta_f = 180°$		
$F10$	$x_f = 50\%$	$L_f = 5mm$ $\theta_f = 45°$	No	
$F11$		$L_f = 5mm$ $\theta_f = 90°$	Yes	4GHz
$F12$		$L_f = 5mm$ $\theta_f = 180°$		
$F13$		$L_f = 10mm$ $\theta_f = 45°$		
$F14$		$L_f = 10mm$ $\theta_f = 90°$		
$F15$		$L_f = 10mm$ $\theta_f = 180°$		
$F16$		$L_f = 20mm$ $\theta_f = 45°$		
$F17$		$L_f = 20mm$ $\theta_f = 90°$		
$F18$		$L_f = 20mm$ $\theta_f = 180°$		
$F19$	$x_f = 70\%$	$L_f = 5mm$ $\theta_f = 45°$	No	
$F20$		$L_f = 5mm$ $\theta_f = 90°$	Yes	4GHz
$F21$		$L_f = 5mm$ $\theta_f = 180°$		
$F22$		$L_f = 10mm$ $\theta_f = 45°$	No	
$F23$		$L_f = 10mm$ $\theta_f = 90°$	Yes	4GHz
$F24$		$L_f = 10mm$ $\theta_f = 180°$		
$F25$		$L_f = 20mm$ $\theta_f = 45°$		
$F26$		$L_f = 20mm$ $\theta_f = 90°$		
$F27$		$L_f = 20mm$ $\theta_f = 180°$		
$F28$	$x_f = 90\%$	$L_f = 5mm$ $\theta_f = 45°$	No	
$F29$		$L_f = 5mm$ $\theta_f = 90°$		
$F30$		$L_f = 5mm$ $\theta_f = 180°$	Yes	3GHz
$F31$		$L_f = 10mm$ $\theta_f = 45°$	No	
$F32$		$L_f = 10mm$ $\theta_f = 90°$		
$F33$		$L_f = 10mm$ $\theta_f = 180°$	Yes	3GHz
$F34$		$L_f = 20mm$ $\theta_f = 45°$	No	
$F35$		$L_f = 20mm$ $\theta_f = 90°$	Yes	2GHz
$F36$		$L_f = 20mm$ $\theta_f = 180°$		

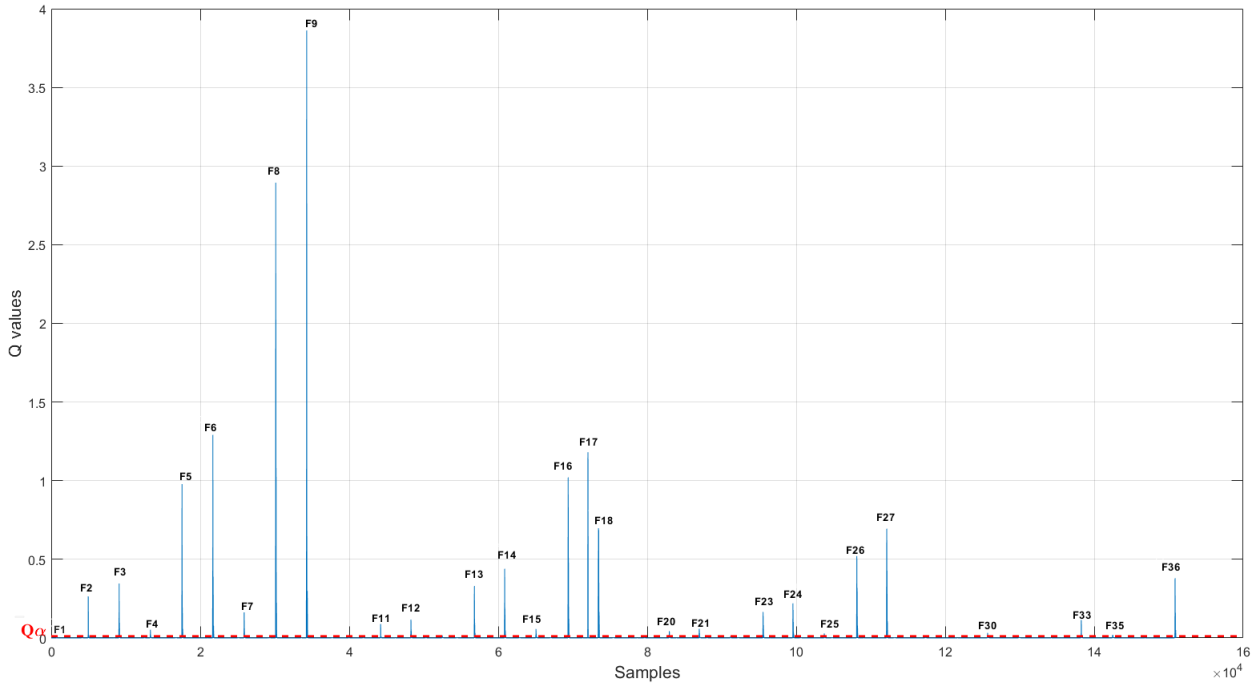


Figure 3.23 – Scores plot: Q values of new measurement samples

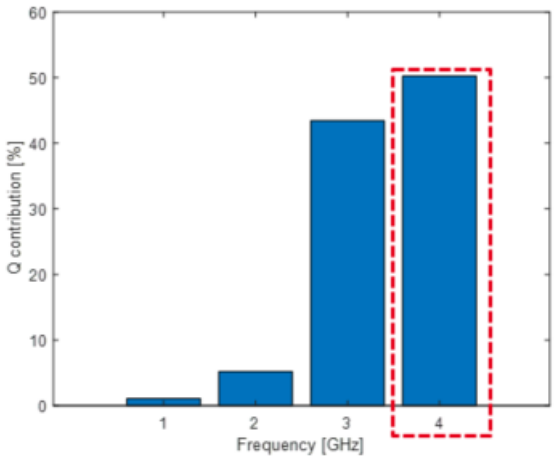
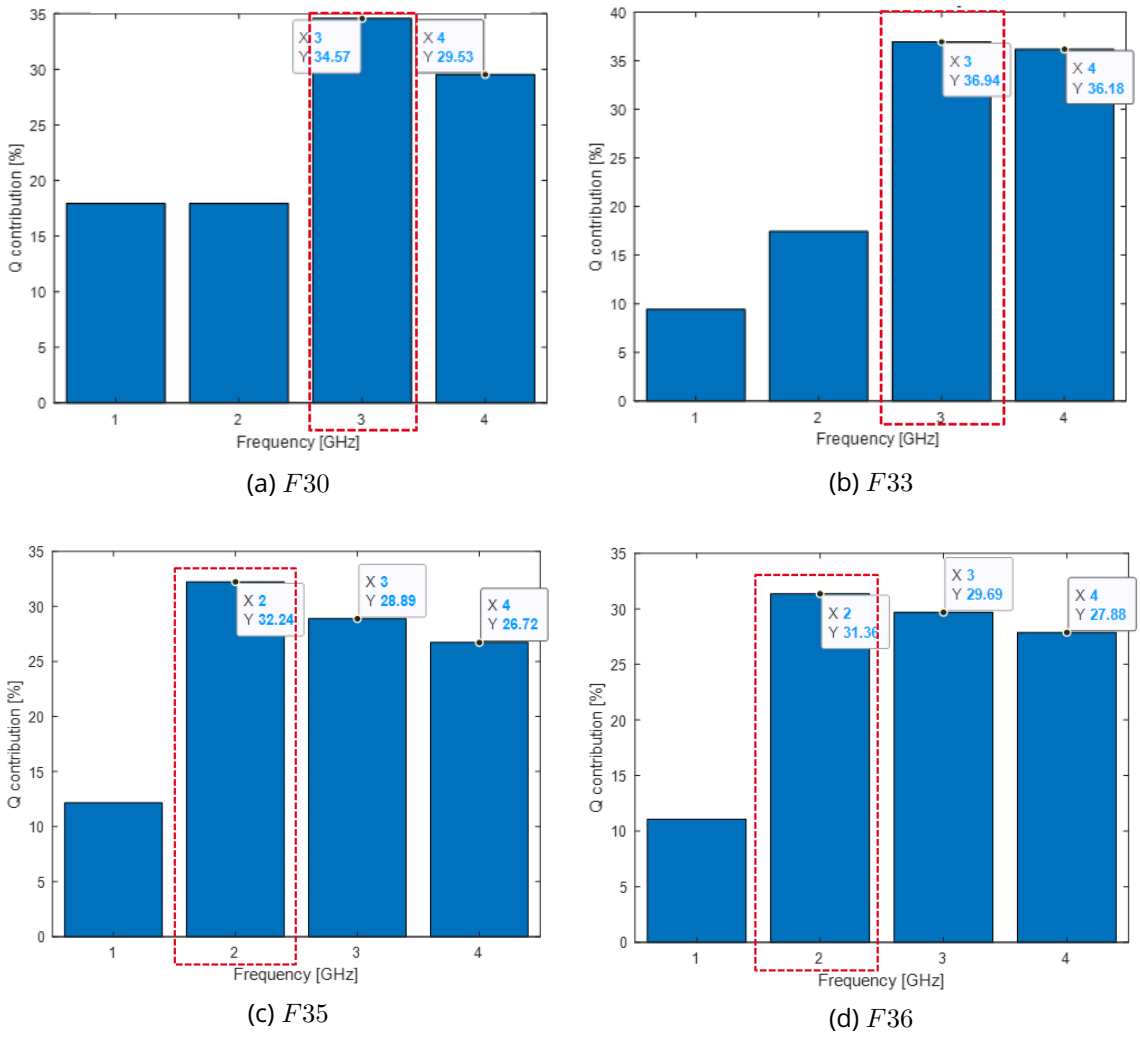


Figure 3.24 – Contribution plot of the fault $F9$

frequency is $4GHz$. These simulation results are in coherence with the rule that as the fault is close to the beginning of the cable, the higher the excitation frequency is, the better is the fault detection.

- For $x_f = 50\%$, the fault with the smallest severity ($\theta_f = 45°$, $L_f = 5mm$) cannot be detected whatever the excitation frequency. All other severities are detected with $4GHz$ as the best frequency.
- For $x_f = 70\%$, the faults of width $\theta_f = 45°$ and with 5 and $10mm$ long cannot



(a) $F30$ (b) $F33$

(c) $F35$ (d) $F36$

Figure 3.25 – Contribution plot of the fault cases at $x_f = 90\%$

be detected whatever the frequency. The selected frequency for the detected cases is $4GHz$.

- For $x_f = 90\%$, the faults with $\theta_f = 180°$ are all detected. For the cases where the lengths are $5mm$ and $10mm$, the selected frequency equals $3GHz$, and for the $20mm$ case, it is $2GHz$. Only one case with $\theta_f = 90°$ and $20mm$ long is detected and the best frequency is $2GHz$.

To show the effect of the fault position on the retained frequency for longer cables, the simulations for a $100m$ cable of the fault case $F3$ ($L_f = 5mm, \theta_f = 180°$) are done. The position of the fault varies with a step of $2m$. In Figure 3.26 the selected frequency decreases as the fault position approaches the end of the cable.

3.5 . Experimental Validation

Table 3.9 – Energy calculation for the fault cases at $x_f = 90\%$

Fault cases	$E_1[J/Hz]$	$E_2[J/Hz]$	$E_3[J/Hz]$	$E_4[J/Hz]$
F30	0.011	0.013	0.073	0.063
F33	0.028	0.035	0.081	0.076
F35	0.003	0.024	0.0107	0.0097
F36	0.0423	0.1406	0.1248	0.113

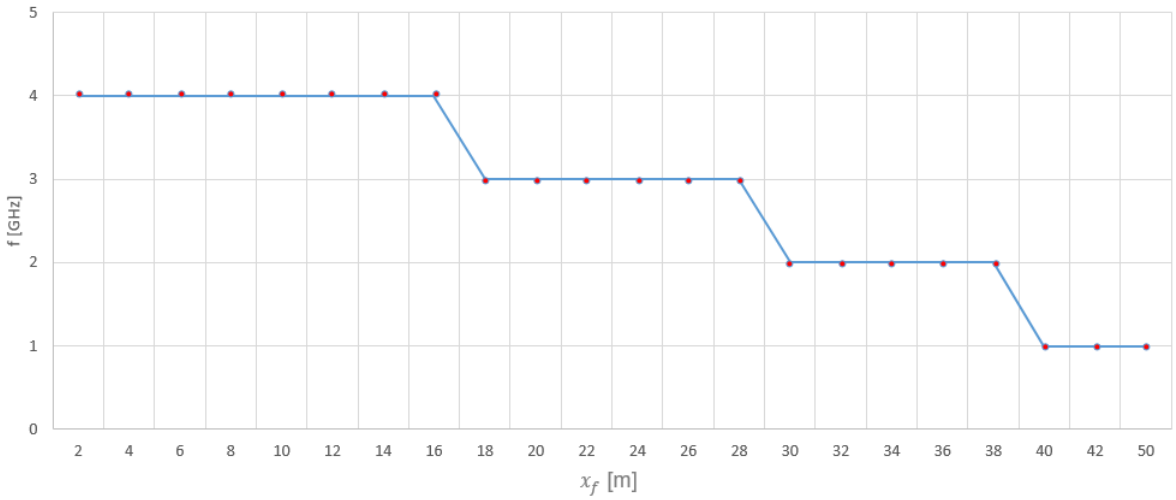


Figure 3.26 – Variation of the selected frequency with the fault position for a $100m$ cable for the fault case ($L_f = 5mm, \theta_f = 180°$)

The methodology for fault detection using PCA, and frequency selection is now evaluated on real cables using the setup in Figure 3.14. The reflectometry responses of the healthy cable and the shield damage located at $0.5m$ on a $1m$ cable are measured, and are the ones displayed in Figure 3.15 and Figure 3.16.

We first consider the reference representation of the data given by X^* as in section 3.4.1.1. Table 3.10 indicates that the cumulative variance of the first two scores is 99.99% that is greater than the lower limit. Using (2.11), k is equal to two. The value of the limit is $Q_\alpha = 0.628$. The new measurement data set is defined as X in section 3.4.1.2. Each vector $\mathbf{x_j}$ contains the

Table 3.10 – PCA model variances

PC number	Variance percentage (%)	Cumulated percentage (%)
1	79.09	79.0959
2	20.904	99.994
3	3.0823e-31	99.994
4	7.1047e-63	100.00

faulty signature of the shielding fault defined above at frequency j.

Figure 3.27 shows the Q test results of this experiment where the fault has been detected. The contribution plot of the abnormal sample is represented in Figure 3.28. The variable

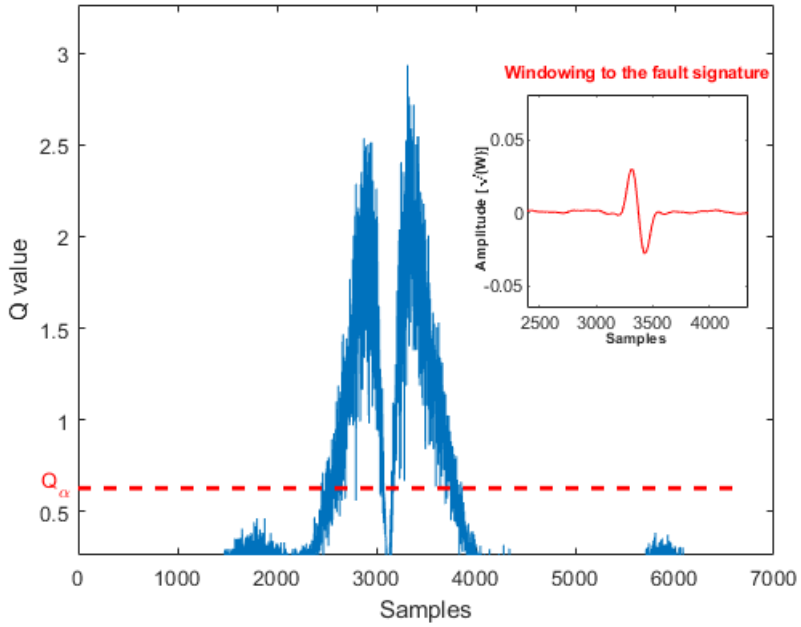


Figure 3.27 – Q values of experimental measurement samples

corresponding to frequency $4GHz$ contributes the most to this sample. Hence, $f = 4GHz$ is the selected maximal frequency. The experimental and the simulation results are coherent in terms of fault detection and frequency choice.

In the following, the performance of this approach will be evaluated for a set of cables with different operating conditions. Two questions will be tackled:

(1) What is the performance of the Q test for different SNRs?

(2) What is the performance of the frequency selection for different SNRs in the fault case detection?

3.6 . Performance Analysis in the Presence of Noise

The performance of the Q test in the presence of noise is studied in this section. Figure 3.29 reveals the flowchart of the used methodology. This study is divided into three parts; First, the soft fault detection, where fault detection using PCA is applied and P_d is calculated in the presence of noise. Second, the false alarm evaluation, where the P_{FA} for each SNR is calculated. Finally, the best frequency selection where the frequency selection in the noisy environment is performed.

3.6.1 . Soft Fault Detection

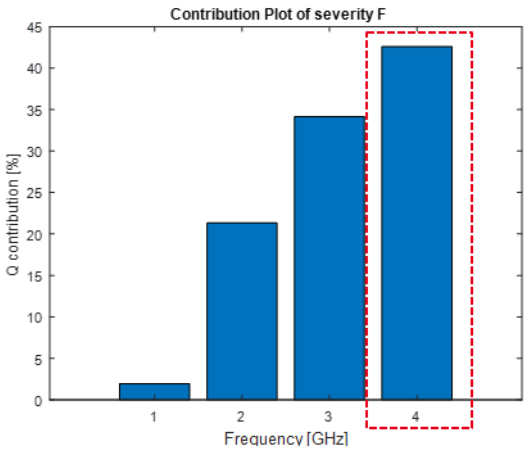


Figure 3.28 – Contribution plot of the abnormal sample in the experimental data corresponding to a $5mm$ long, $180°$ wide shielding fault

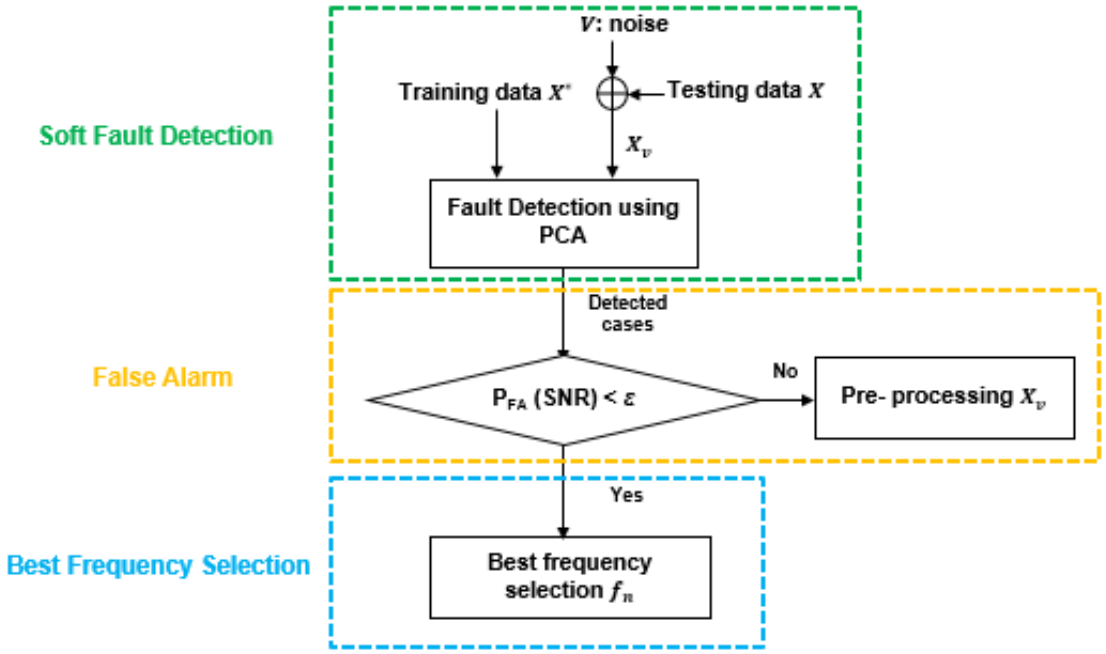


Figure 3.29 – Flowchart of the Q test performance analysis

3.6.1.1 Training Data Generation

For the training data X^*, the star mark (*) refers to the fault-free and noise-free environment. This data matrix is the one used in section 3.4.2.2 where a $10m$ cable is used and simulated at four different excitation frequencies.

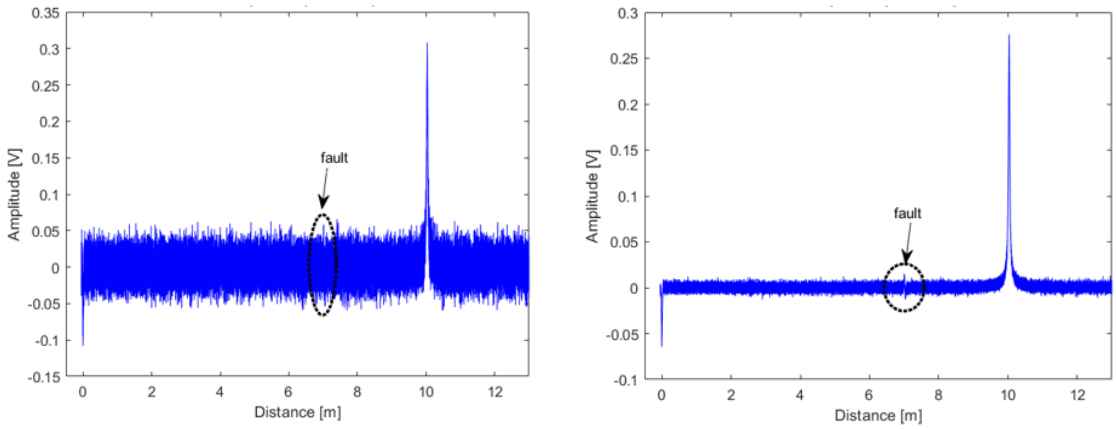
3.6.1.2 Testing Data Generation: Noise introduction

The testing data used here is X in 3.4.2.3 where the different fault scenarios (36 cases by varying the fault position along a $10m$ cable while changing its severity (L_f and θ_f)) at different frequencies are used for the performance analysis of the proposed method.

Now, noise (V) is added to the testing data X, resulting in X_ν where the used Signal to Noise Ratio (SNR) levels are $-5dB$, $0dB$, $5dB$, $10dB$ and $15dB$. The variable $\mathbf{x_j}$ is a column vector of X taken for the jth variable. The noise is assumed to be Additive White Gaussian Noise (AWGN). The noise vector added to the variable $\mathbf{x_j}$ is $\nu_j \sim N(0, p_\nu)$ ($\mathbf{x_{\nu j}} = \mathbf{x_j} + \nu_\mathbf{j}$). Its power p_ν is related to the signal's power p_s by 3.7. Then, 500 realizations are performed at each SNR level.

$$SNR = 10\log\left(\frac{p_s}{p_\nu}\right) \tag{3.7}$$

Figure 3.30 shows the TDR responses when the noise in the presence of the fault ($L_f = 5mm$, $\theta_f = 180°$ at $x_f = 70\%$) at SNR= $-5dB$ and SNR= $10dB$ is added, respectively. It is noticed that the noise can mask the soft fault at SNR=$-5dB$, thus it is impossible to find the fault peak just by looking at the TDR response.



(a) SNR= $-5dB$ (b) SNR= $10dB$

Figure 3.30 – TDR response in the presence of F at $x_f = 70\%$ with the effect of the added noise

3.6.1.3 Fault Detection

For each realization, we apply the Q test to find if the fault case is detected or not. Then, the fault detection probability, P_d, is calculated for each fault case at each SNR, and if $P_d \geqslant \epsilon_d$, the fault is considered detected at this SNR. ϵ_d is a domain-specific threshold. Its value is determined by the application domain in which the faults need to be detected. For example in the military, medical and health domains, anomaly detection is a very critical problem and requires a high degree of accuracy [135]. The same 8 non-detected fault cases without noise

(in section 3.4.2.3 where 36 fault cases are studied in noise-free environment, 8 cases are not detected) are not detected here. The 28 detected faults without noise, are also detected in the presence of noise.

For case study, let fault F corresponds to the fault case with the following parameters $L_f = 5mm$ and $\theta_f = 180°$. Consider a $10m$ cable, and vary the fault position, from 20%, 50%, 70% to 90% with respect to the cable length.

P_d at a specific SNR is calculated as follows: the Q test is applied for each realization. If an anomaly is detected, P_d of this realization equals one else zero. For the total number of realizations, P_d is the summation of the individual P_d divided by 500:

$$P_d = \frac{1}{500} \sum_{i=1}^{500} P_{di} \tag{3.8}$$

Table 3.11 gives the P_d of the fault case F for the different SNRs. It is shown that the fault F, in the presence of noise, is detected with a $P_d \geqslant 0.85$. Here, ϵ_d is not considered as it depends on the application domain. As an hypothesis, we assume that the obtained detection probabilities are sufficient for the next steps to analyze the performance.

Table 3.11 – P_d for different noise levels

SNR [dB]	$x_f = 20\%$	$x_f = 50\%$	$x_f = 70\%$	$x_f = 90\%$
-5	1	1	0.853	0.859
0	1	1	1	0.862
5	1	1	1	0.896
10	1	1	1	0.983
15	1	1	1	1

3.6.2 . False Alarm Analysis

The false alarm probability, P_{FA}, at a specific SNR is calculated as follows: first, the TDR signal is divided into healthy and faulty intervals. Then, for each realization, the Q test is applied. If an anomaly is detected in the healthy interval, then P_{FA} of this realization will be equal to one else zero. By averaging the 500 realizations, P_{FA} is calculated. Now, P_{FA} value for the different SNR is shown in Table 3.12. It is 0.034 for an SNR of $-5dB$ and 0.018 at

Table 3.12 – P_{FA} for different noise levels

SNR [dB]	P_{FA}
-5	0.034
0	0.018
5	0
10	0
15	0

$0dB$ and 0 for the other SNR values. The value of the false alarm threshold ϵ depends on the

application domain. Hence, if $P_{FA} > \epsilon$, the data X_ν needs to undergo a pre-processing step; else, we proceed to the best frequency selection part. Here, ϵ is not considered as it depends on the application domain. As a hypothesis, we assume that the obtained false alarm probabilities are sufficient for the next steps to analyze the performance.

3.6.3 . Best Frequency Selection

3.6.3.1 Robustness to Noise Evaluation

For each detected fault case, at a given SNR and for a specific realization, the best frequency selection is attained using the contribution plot of the detected abnormal sample related to the fault case. Taking into account the 500 realizations, the selected frequency f_n is the frequency with the highest occurrence rate among the four used frequencies.

For the case study above, Figure 3.31 represents the variation of the selected frequency with the SNR for different fault positions. It is noted that the selected frequency to monitor F depends on its position and the current SNR. The two fault positions, $x_f = 20\%$ and $x_f = 50\%$ have exactly the same variation. As the fault reaches the cable end, the selected frequency decreases. This change is predicted due to the attenuation throughout the cable and agrees with the rule for long cables: the lower the frequency is, better is the fault detection.

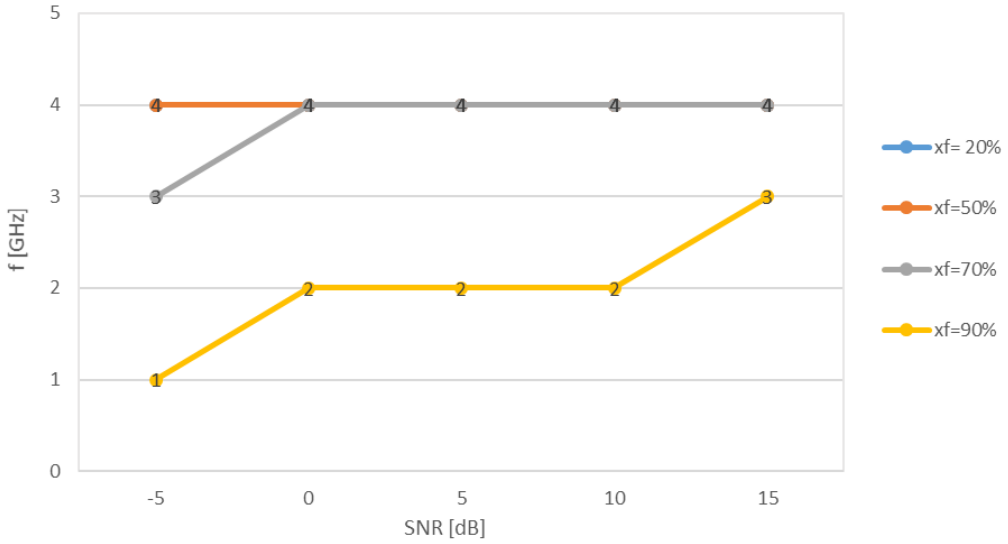


Figure 3.31 – Frequency variation with the SNR for a $10m$ cable of a $5mm$ and $180°$ fault

Now, for another case study, by considering a $100m$ cable, fault F position varies along the cable length such that, $x_f = 10\%$, 20%, 30% and 40%. We take only half of the cable length ($50m$); for the other half, the measurements could be done from the other extremity of the cable, making the proposed method more efficient. Figure 3.32 monitors the variation of the selected frequency with the SNR for different fault positions.

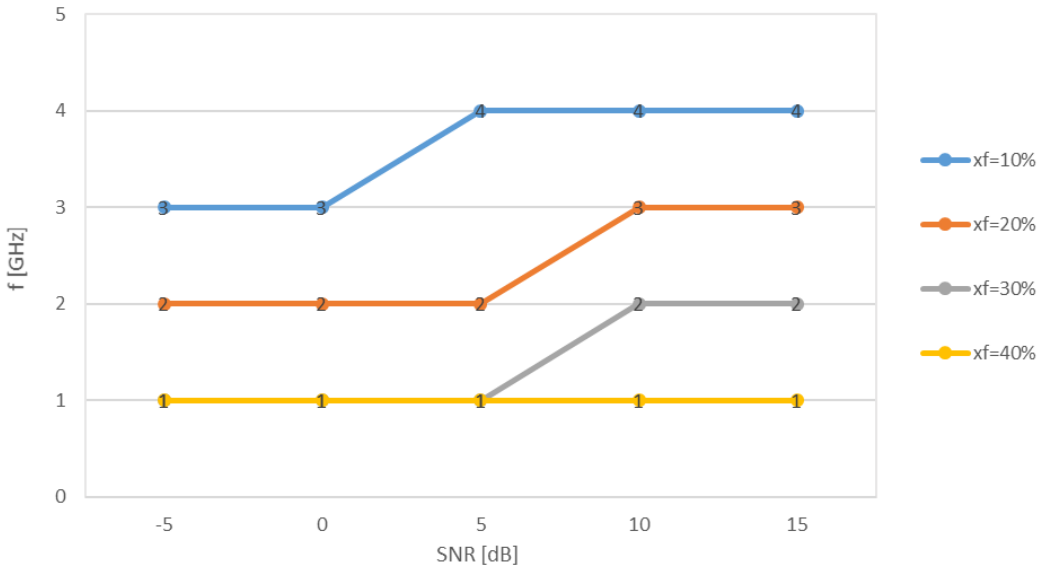


Figure 3.32 – Frequency variation with the SNR for a $100m$ cable of a $5mm$ and $180°$ fault

3.6.3.2 Frequency Occurrence

For all the detected 28 fault cases in section 3.6.1.3, whenever the selected frequency in the noisy environment is different from that in the noise-free environment (i.e. $f_n \neq f$), the occurrence percentage for each frequency will be studied. These fault cases are:

- $F25$ at SNR$= 0dB$
- $F30$ at SNR$= 0dB$, $5dB$ and $10dB$
- $F35$ at SNR$= 0dB$

Those are the critical cases at the positions 70% and 90%:

- $F25$, it is the only detected case at 70% where $\theta_f = 45°$
- $F30$, it is the only detected case at 90% where $L_f = 5mm$
- $F35$, it is the only detected case at 90% where $\theta_f = 90°$

3.6.3.2.1 Fault case F25 ($x_f = 70\%, l_f = 20mm, \theta_f = 45°$ at SNR$= 0dB$), The selected frequency is $f_n = 3GHz$, where as the selected frequency in the noise-free case is $f = 4GHz$. So, in the presence of noise, what is the occurrence percentage for each frequency? Do $3GHz$ and $4GHz$ have close percentages? Table 3.13 presents the different frequency occurrence percentages for 500 realizations. To help understand the overall behavior,

2000 realizations are made. For 500 realizations, it is shown that the $3GHz$ percentage (the highest - the selected frequency f_n) is the closest to the $4GHz$ percentage, which is originally selected in the noise-free condition. For 2000 realizations, no major changes in the results are noticed.

Table 3.13 – Frequency occurrence percentages for the fault case $F25$

SNR [dB]	*f* [GHz]	Occurrence percentage
0	1	3.3%
	2	21.5%
	3	**39.2%**
	4	**36%**

3.6.3.2.2 Fault case F30 ($x_f = 90\%, l_f = 5mm, \theta_f = 180°$ at SNR$= 0dB, 5dB$ and $10dB$), the selected frequency is $f_n = 2GHz$, where the selected frequency in the noise-free case is $f = 3GHz$. Table 3.14 represents the different frequency occurrence percentages for 500 realizations. It is noted that the difference between the $2GHz$ and $3GHz$ percentages decreases as SNR increases. For 2000 realizations, no major changes in the results are noticed.

Table 3.14 – Frequency occurrence percentages for the fault case $F30$

SNR [dB]	*f* [GHz]	Occurrence percentage
0	1	17.8%
	2	**35.2%**
	3	**31.4%**
	4	15.6%
5	1	2.7%
	2	**43.5%**
	3	**41.3%**
	4	12.5%
10	1	0.6%
	2	**45.3%**
	3	**43.4%**
	4	10.7%

3.6.3.2.3 Fault case F35 ($x_f = 90\%, l_f = 20mm, \theta_f = 90°$ at SNR$= 0dB$), the selected frequency is $f_n = 1GHz$, where as the selected frequency in the noise-free case is $f = 2GHz$. Table 3.15 represents the different frequency occurrence percentages for 500 realizations. For 500 realizations, it is shown that $1GHz$ percentage (the highest - the selected frequency f_n) is the closest to the $2GHz$ percentage, which is originally selected in the noise-free condition. For 2000 realizations, no major changes in the results are noticed.

Table 3.15 – Frequency occurrence percentages for the fault case $F35$

SNR [dB]	*f [GHz]*	Occurrence percentage
0	1	**45.6%**
	2	**33.2%**
	3	12.8%
	4	8.4%

3.6.3.3 ROC Curves Investigation

At each SNR level, for each realization, the Q value data is divided into two classes: Healthy and Faulty. Then for the 500 realizations, the data is concatenated in the Healthy and the faulty classes and the ROC curves are calculated.

The comparative study is first done by setting the SNR to $15dB$ and varying the position of the fault F ($5mm$,180°) in a $10m$ cable case. The detection performance results using the Q test for different SNR levels are displayed in Figure 3.33, for $x_f = 20\%$, $x_f = 70\%$ and $x_f = 90\%$, respectively.

Thanks to the simulation results of a $10m$ length RG316 coaxial cable with a shielding damage, studied at four different positions x_f =[20%, 50%, 70%, 90%]. It is noted that:

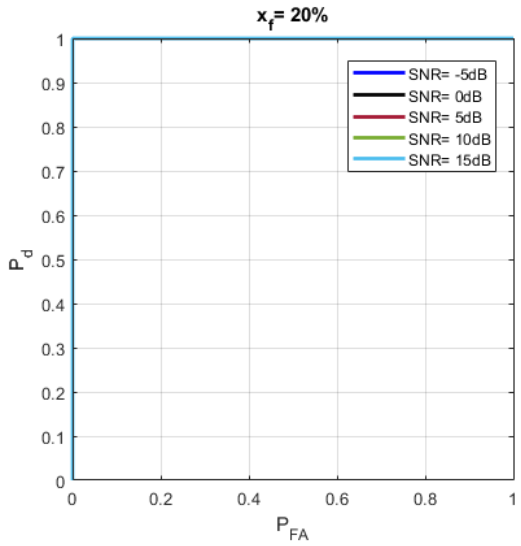
- The Q test has excellent efficiency with 100% detection capability for the noise levels (SNR $\geq 0dB$) with a low false alarm probability for $x_f = 20\%$ and $x_f = 70\%$.
- For the lower noise levels (SNR $< 0dB$), the detection performance of Q test are affected by the noise for $x_f = 70\%$ and it remains 100% for $x_f = 20\%$.
- The Q test has high efficiency with 85% detection capability with a low false alarm probability for $x_f = 90\%$. The fault detection performance decreases along with the increasing noise level from 100% at $15dB$ to 85.9% at $-5dB$.

Q test is then efficient for detecting faults, but its performance is affected by the SNR when the fault position is close to the cable end.

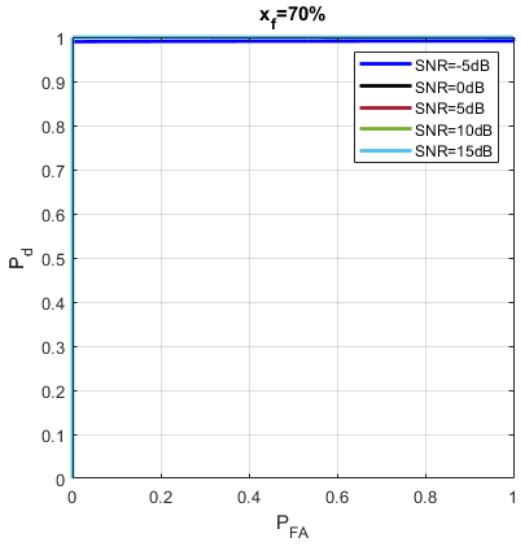
As the performance analysis of this method is investigated, it reveals excellent performance for soft faults detection. The detection capability is equal to 89.6% even when SNR$= 5dB$. As the fault position approaches the end of the cable, the performance is still good and is equal to 85.9% at SNR$= -5dB$.

3.7 . Performance of the Selected Frequency f in the Presence of Noise

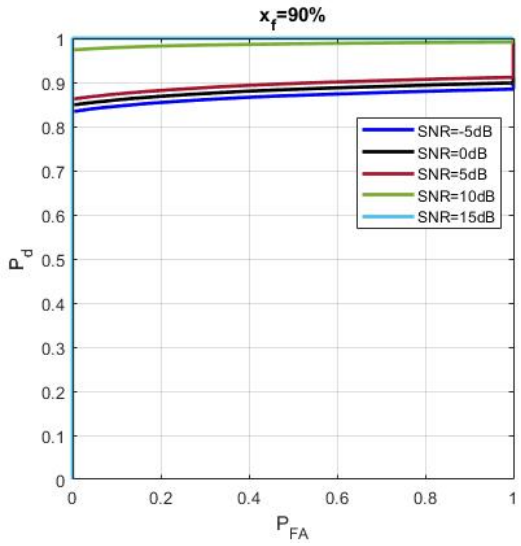
The performance of the selected frequency f in the presence of noise in the fault case detection is now studied. Figure 3.34 represents the flowchart of the proposed methodology. It consists of the following steps:



(a)



(b)



(c)

Figure 3.33 – ROC curves for a $10m$ cable of a $5mm$ and $180°$ fault for (a) $x_f = 20\%$, (b) $x_f = 70\%$ and (c) $x_f = 90\%$

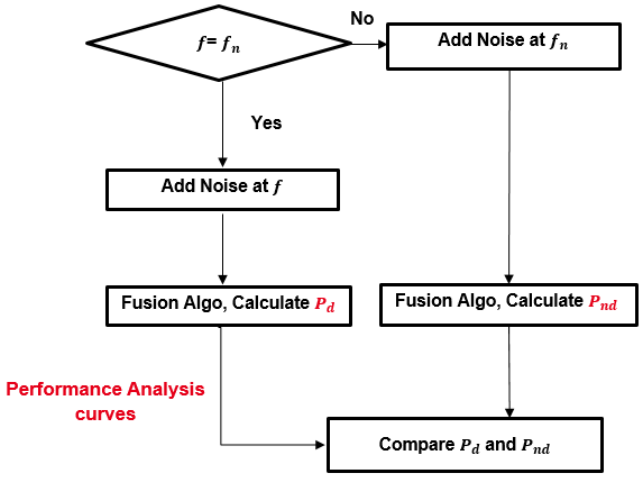


Figure 3.34 – Flowchart of the performance of the selected frequency f in the presence of noise

- From the PCA results is section 3.4.2, 28 fault cases are detected and the selected frequency f for each fault case is indicated in Table 3.8. Noise is added to the signal at frequency f and 500 realizations are performed.
- A fusion algorithm [13] introduced in chapter 2 section 2.3.1 of several post processing methods (Signature Magnification by Selective Windowing (SMSW) [29], subtractive correlation method [97], and the method based on the integral of a reflectogram [136]) is applied to each realization. P_d is calculated for each fault case, at a specific SNR. The results of this method are represented by the performance analysis curves.
- For the cases in section 3.6.3.2, where $f_n \neq f$, noise is added to the signal at frequency f_n. Then the fusion algorithm is used to obtain P_{nd}, the probability of detection calculated at the frequency f_n. Here, the goal is to compare P_d and P_{nd}, and to evaluate the performance of the Q test in section 3.6 using the Fusion algorithm. The results of this study are presented in the Appendix A.

Now, to obtain the performance analysis curves, each detected fault case in Table 3.8, at frequency f, for each SNR, P_d is calculated using (3.9) by averaging the probability of detection of all the realizations $r = \{1, 2, \ldots, 500\}$:

$$P_d = \frac{\sum_{r=1}^{500} p_r}{500} \tag{3.9}$$

where p_r is the probability of detection for each realization, at a specific SNR, calculated using the fusion algorithm.

For the case study, the fault case where $L_f = 20mm$ with three different severities ($\theta_f = 45°$, $\theta_f = 90°$ and $\theta_f = 180°$) is studied. For each severity, the fault position varies along the cable:

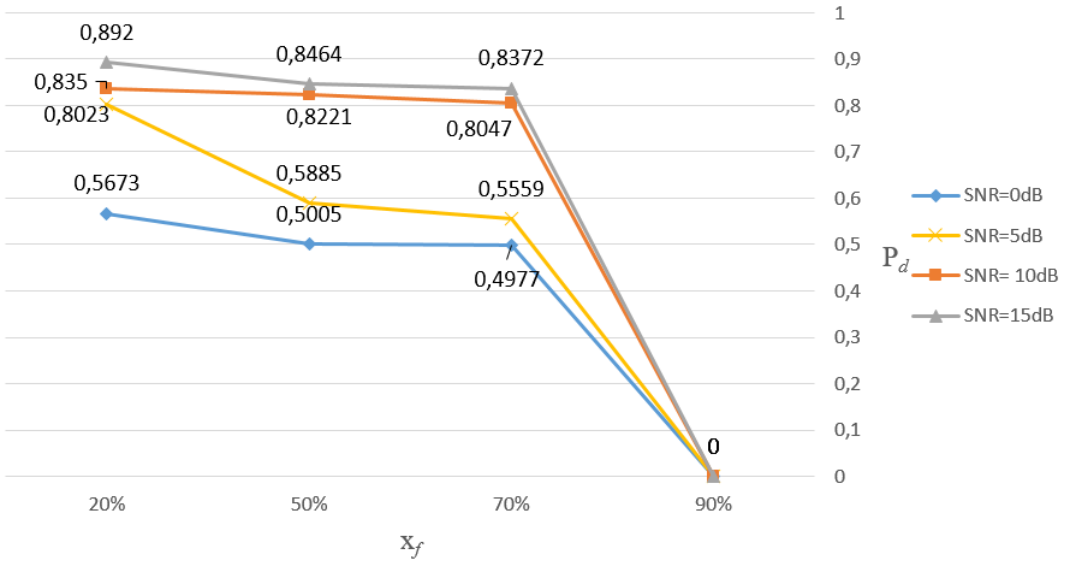


Figure 3.35 – Performance analysis curve in the case of a $20mm$ and $45°$ fault

(1) The performance analysis curve for the first severity ($L_f = 20mm$, $\theta_f = 45°$) is represented in Figure 3.35. P_d is calculated for different fault positions and for each SNR. It is shown that, for each SNR, the detection capability decreases from $P_d = 0.8$ to 0, as this low severity fault reaches the cable end ($x_f = 90\%$).

(2) Figure 3.36 represents the performance analysis curve for the second severity ($L_f = 20mm$,$\theta_f = 90°$). P_d is calculated for different fault positions and for each SNR. It is shown that the detection capability for this severity is better than that for the first severity. It decreases to $P_d = 0.0306$ at $-5dB$ for 90% position. However, it is greater than 0.8005 when fault position is less than 90%, even at a very high noise level.

(3) Figure 3.37 represents the performance analysis curve for the third severity ($L_f = 20mm$,$\theta_f = 180°$). P_d is calculated for different fault positions and for each SNR. It is shown that the detection capability for this severity is better than that for the second severity. It decreases to $P_d = 0.0438$ at $-5dB$ for 90% position. However, it is greater than 0.916 when the fault position is less than 90%, even at a very high noise level.

For this analysis, it is shown that the probability of detection increases as the fault position approaches the beginning of the cable and as the SNR increases for the same fault severity. Thus, the selected frequency f shows improved detection capability as the fault approaches the cables' beginning with a smaller noise level. Moreover, for lower fault severities, the probability of detection P_d increases.

3.8 . Conclusion

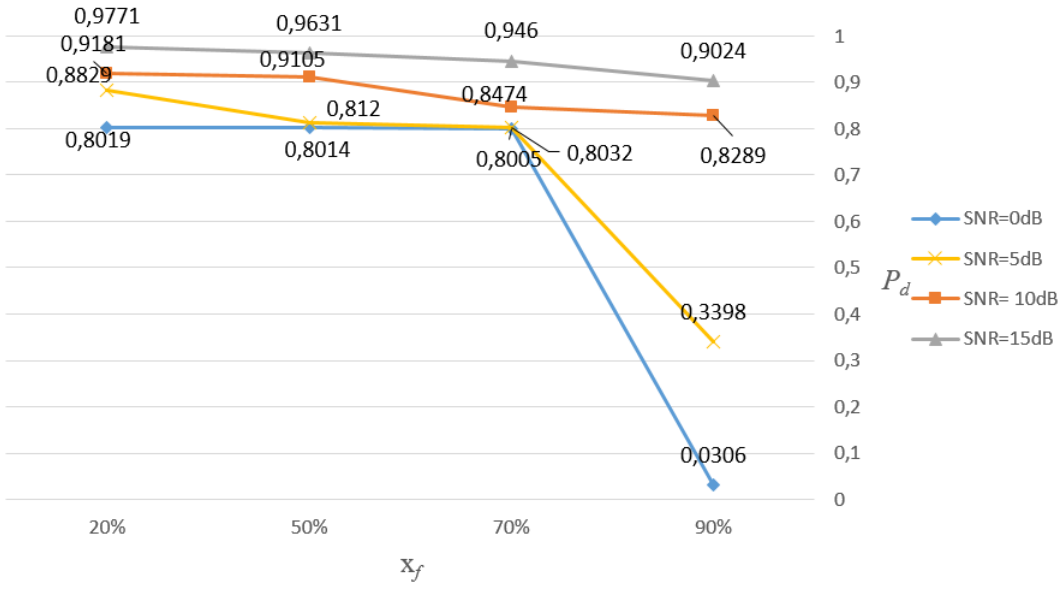


Figure 3.36 – Performance analysis curve in the case of a $20mm$ and $90°$ fault

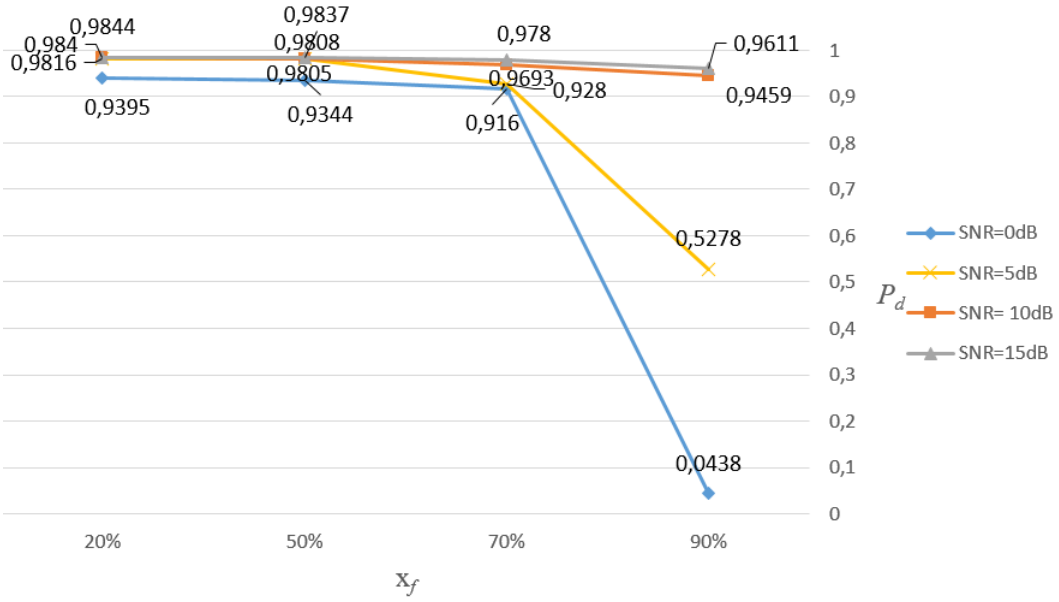


Figure 3.37 – Performance analysis curve in the case of a $20mm$ and $180°$ fault

This chapter has presented an efficient approach to select the best frequency bandwidth for soft fault detection in wired networks based on a judicious combination of reflectometry and Principal Component Analysis. In practice, the expert configures and calibrates the VNA at a given frequency and records the healthy cable measurement. Measurements at the same frequency are then done on a faulty cable. Analysis of the measurements is established at this frequency on the computer. If the fault is not detected, this operation must be repeated. Therefore, there is a loss of information and time in addition to the subjectivity of the decision-making.

The proposed method allows to configure the VNA at different frequencies. It performs measurements at different frequencies for the healthy (reference) case. After which, the PCA model is established. It performs the new measurements at different frequencies. If a difference is detected between the reference and the measurement data, the contribution for each variable (i.e. frequencies) to this difference is calculated. The algorithm then chooses the most relevant frequency to monitor the soft fault. The advantages are thus time saving and objectivity of the decision-making.

First, the three-dimensional modeling of a soft fault and its influence on the propagation in a cable has been studied with CST. The simulation results made it possible to determine the disturbance generated by the soft fault with different levels of degradations. These disturbances have been represented in terms of the reflection coefficient to determine the fault characteristics. To overcome the increase in the computational burden and the simulation time caused by the CST in the case of longer cables, the fault parameters are extracted using an $RLCG$ characterization method, and are implemented in a Mathwork's Matlab©language code to simulate different scenarios where the cable length l and the fault position x_f are changed. The simulation results of this method are compared to the output of the CST simulator with a high degree of success, and the Matlab code was tested for its ability to retrieve (estimate) the TDR response. Experimental measurements had been carried out to validate the numerical model in the case of shielding damage.

This method is then investigated for a set of cable lengths with different operating conditions and different fault types. The simulation results are in coherence with the known rule that for short cables, the higher the excitation frequency is, the better it is for fault detection. In addition, as the fault reaches the cable end for longer cables, the selected frequency decreases. This change is predicted due to the attenuation throughout the cable and agrees with the common rule for long cables.

Experimental validation was carried out. The experimental and the simulation results are coherent in terms of fault detection and frequency choice.

The performance analysis of this method has also been investigated. It has revealed good performance for soft faults detection. The probability of detection equals to one even when SNR$= 0dB$. As the fault position approaches the end of the cable, the performance is still good, but for lower fault severities, the detection is more tedious. In this study, it is noted that the selected frequency to monitor a fault case depends on several parameters: the cable type and characteristics, the fault severity and position, and the current noise level.

Regarding the statistical tests, the Q criteria is more relevant for evaluating the faults. Thus it will be considered in the following chapter for sensor selection.

4 - Sensors Selection for Distributed Reflectometry-based Soft Fault Diagnosis using PCA

Contents

4.1 . Introduction

Although reflectometry offers promising results in point-to-point topology networks, it introduces ambiguity related to fault location in more complex wired networks due to the presence of multiple branches and junctions. Moreover, the complexity of the wired networks comes with the increase of the signal attenuation. As a solution, distributed reflectometry method is used [33]. It consists of taking measurements at several points of the Network Under Test (NUT) by implementing several sensors at different extremities to maximize the diagnosis coverage. However, the injection of multiple signals down to the NUT leads to computational complexities and sensor fusion problems. Additionally, energy consumption is a significant drawback of this method concerning environmental constraints.

This chapter introduces an approach for the most relevant sensors selection to monitor and diagnose soft faults in multi-branched wired networks.

In this context, the proposed method combines TDR with PCA. The simulation results are provided for a Y-shaped network and a CAN bus connected in a network structure in which sensors perform reflectometry measurements consecutively. The TDR responses are then arranged into a database. With this latter, a PCA model is developed and used to detect the soft faults. Coupled with statistical analysis based on the Q test (also known as SPE), the most relevant sensors for monitoring and diagnosing soft faults in the network are identified with high accuracy. Indeed, the proposed method enables monitoring without interfering with other existing signals (of the monitored network) and without interference between different reflectometers (sensors) signals. Based on these results, the sensor's number could be reduced, and the non-selected ones could be inactivated, reducing energy consumption, computing complexities, and sensor fusion problems.

Experimental validation is performed next to show the method's advantages. Finally, performance analysis is developed to study the efficiency of the proposed method in the presence of noise.

4.2 . Sensor Selection Algorithm for Soft Fault Diagnosis

In this section, we propose a methodology to automate the detection of a fault in a multi-branched wired network and select the most relevant sensors to monitor the detected soft fault and inactivate temporarily the non-selected sensors.

Figure 4.1 describes the principle of the new approach combining TDR distributed reflectometry measurements with PCA. It is composed of three main phases: training, monitoring and sensor selection.

4.2.1 . Training Phase

First, data are collected when the network is considered fault-free (step 1). In that case, each sensor in the distributed network collects its TDR measurements which are used to create

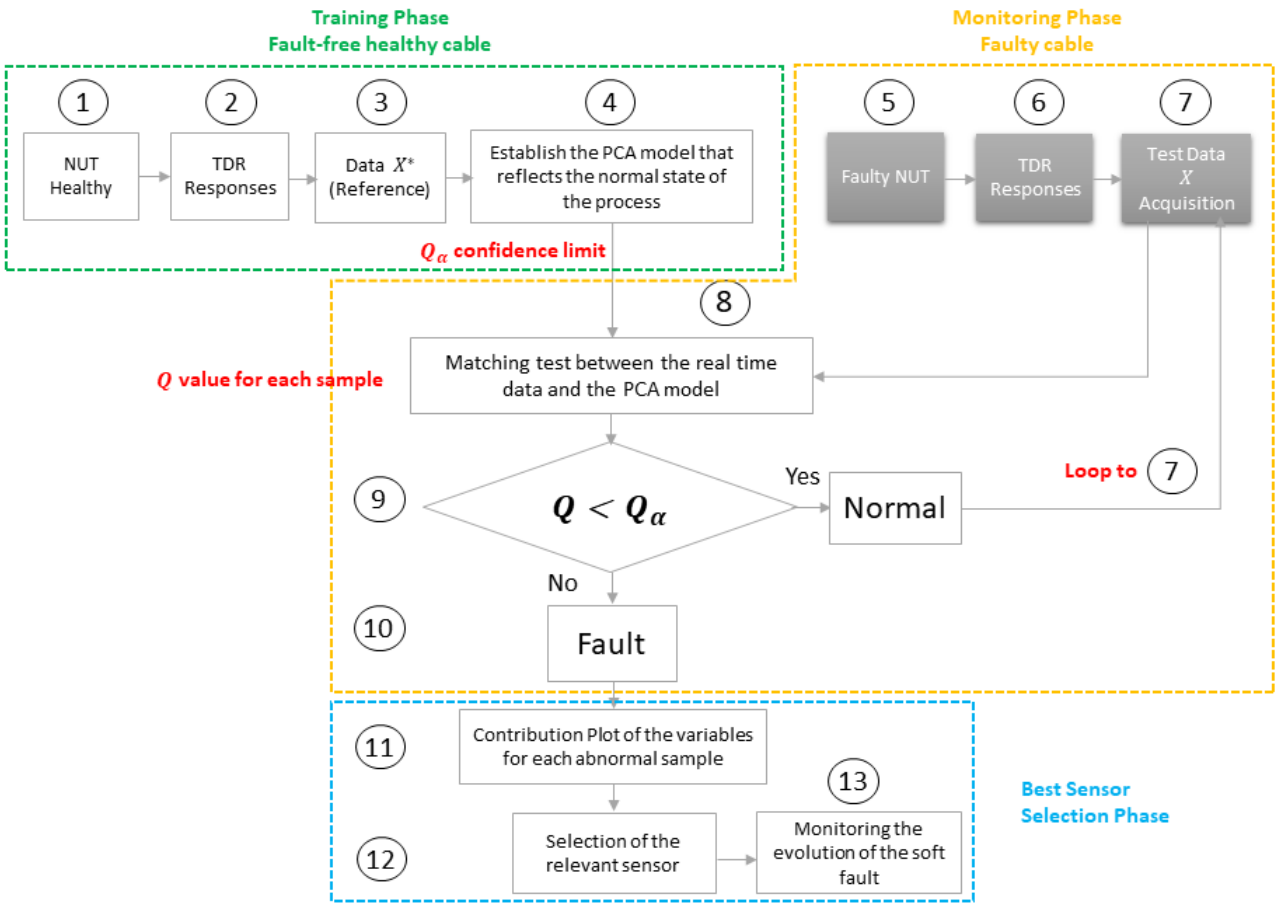


Figure 4.1 – Methodology of the TDR distributed reflectometry PCA-based approach

the database (step 2). Then, a matrix X^* is constructed such that each of its column variables corresponds to a sensor TDR response (step 3). A PCA model is developed based on this database as described in chapter 2, section 2.6.1 (step 4). This model is used in the second step to evaluate new measured data to detects abnormality [121, 122].

4.2.2 . Monitoring Phase

Second, for the NUT, X, the new measured data matrix is built the same way as the reference data matrix X^* was constructed. It is then projected in the PCA reference frame obtained during the training phase. The new scores T_{new} and the residual $\tilde{T_{new}}$ are then calculated. The detailed analytic calculations are similar to those provided in chapter 2, section 2.6.2 (steps 5, 6 and 7).

Monitoring statistics are used next for fault detection, i.e., to find whether a fault has occurred or not. For this purpose, the Q statistical test is used [121]. This test is used for evaluating the fault presence (step 8). The confidence limit Q_α for this test is calculated using

the reference data X^* which is used for constructing the PCA model in the training step. If the Q value falls outside the confidence limit for a specific sample, then a fault exists (steps 9 and 10).

4.2.3 . Sensor Selection Phase

Finally, whenever a fault is detected, we proceed to the sensor selection analysis step. For each detected faulty sample, the analysis starts by plotting the contribution of the variables constituting the new measured data matrix X (i.e., sensors TDR responses) (step 11). Then, we can inspect the variables that highly influence this sample statistics value. Therefore, we can choose the most relevant sensor to monitor the evolution of this fault and inactivate temporarily all other sensors (steps 12 and 13).

4.3 . Simulation Results

To validate the proposed methodology, we consider soft faults diagnosis in a point-to-point network, Y-shaped network, and a Controller Area Network (CAN) Bus. These configurations present topologies with increasing complexity in terms of branches, junctions, echoes, signal attenuation, etc.

4.3.1 . Obtained Results in Case of Point-to-Point Network

The cable model introduced in chapter 3, section 3.2.1 is used where two sensors **S1** and **S2** were implemented at each extremity of the cable (Figure 4.2) and the TDR measurements were

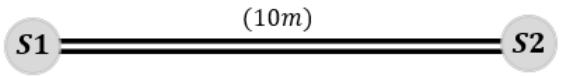


Figure 4.2 – Point-to-Point topology

performed. The length of the cable is $10m$. The reference fault-free data matrix is created:

$$X^* = [\mathbf{x_1^*} \quad \mathbf{x_2^*}] \tag{4.1}$$

X^* is formed up of two variables. Each variable $\mathbf{x_j^*}$ corresponds to the TDR response of the reference healthy cable for the sensor **Sj** ($j \in \{1, 2\}$).

Shielding damage is introduced on the cable under test as in section 3.2.3. In this case, the length of the fault $l_f = 5mm$ and its width is $\theta_f = 180°$. The fault position x_f is set to one of five values: $2m$, $4m$, $5m$, $7m$ and $9m$. The TDR response for each sensor is then obtained for each fault position. Using these faulty model reflectometry data, the new measured data matrix X is defined as:

$$X = [\mathbf{x_1} \quad \mathbf{x_2}] \tag{4.2}$$

where $\mathbf{x_j}$ describes the concatenated fault signature data vector for the sensor **Sj** ($j \in \{1, 2\}$) at each fault position. E.g, for $j = 1$:

$$\mathbf{x_1} = [\mathbf{R_{S1@2m}} \quad \mathbf{R_{S1@4m}} \quad \mathbf{R_{S1@5m}} \quad \mathbf{R_{S1@7m}} \quad \mathbf{R_{S1@9m}}]^T \tag{4.3}$$

where $\mathbf{R_{S1@x_f}}$ describes the fault signature data vector measured by sensor **S1** at fault position x_f.

Now, the Q value of each new measured sample is calculated. Figure 4.3 represents the Q

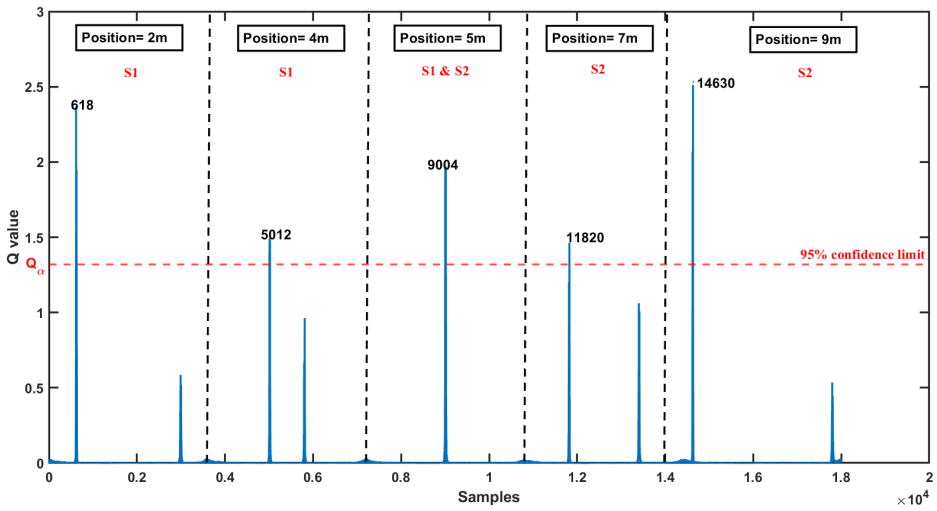


Figure 4.3 – Q chart of the new measured samples of the Point-to-Point topology

control chart, with the dashed red line representing the 95% confidence limit ($Q_\alpha = 1.32$). It is observed that several samples have crossed Q_α. This means that faults have occurred. Due to the presence of five different fault position cases, the method should allow to choose the most relevant sensor(s) to monitor the fault for each case. Plotting the contribution charts of the indicated faulty samples permits us to know the sensor that highly influences its Q value. Hence, the selection of the relevant sensor is performed by the Q control test. Sensor **S1** is selected for the positions $2m$ and $4m$ and sensor **S2** is selected for the positions $7m$ and $9m$. In the case where $x_f = 5m$ (fault at equal distance for the two sensors), the method propose to choose the two sensors.

4.3.2 . Obtained Results in Case of Y-Shaped Network

We consider the Y topology in Figure 4.4 with three branches B_1, B_2 and B_3. Two different cases were studied where we change the branches lengths. Three sensors were placed at the extremities of this topology. $X^* = [\mathbf{x_1^*} \quad \mathbf{x_2^*} \quad \mathbf{x_3^*}]$ is the reference fault-free (healthy) data matrix. It is formed up of three variables. Each variable $\mathbf{x_j^*}$ corresponds to the TDR response of the Y network for the sensor **Sj**.

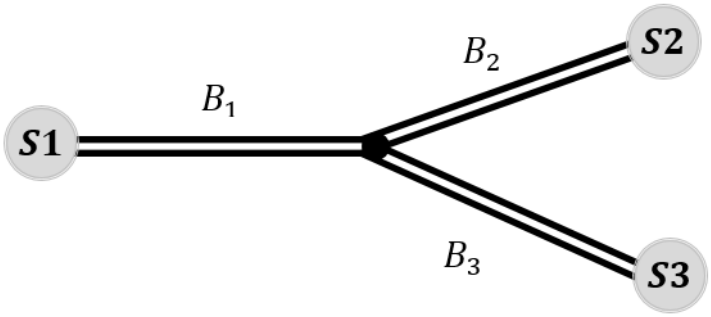


Figure 4.4 – Y network topology

4.3.2.1 First Case: Equal Distance Branches

Figure 4.5 shows the first case topology and the different studied fault positions. In that case, $B_1 = B_2 = B_3 = 5m$. The fault positions are studied one fault at a time. $X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3}]$ is the new measured data matrix.

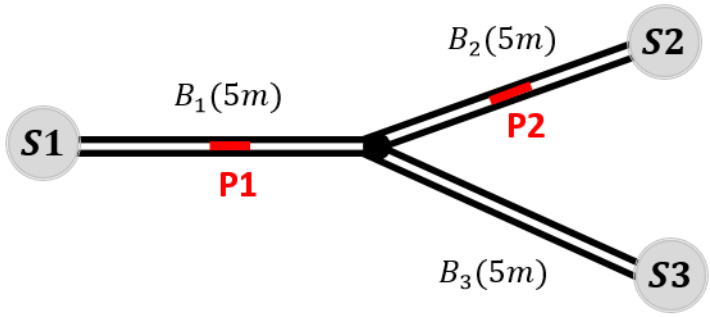


Figure 4.5 – Y network topology: First case

4.3.2.1.1 Fault in Position P1 The first simulated fault is located at $3m$ from sensor **S1** in B_1. The TDR responses of the three sensors are shown in Figure 4.6. After calculating the Q values of these measurements (Figure 4.7), the contribution plot of the highest peak is represented in Figure 4.8.

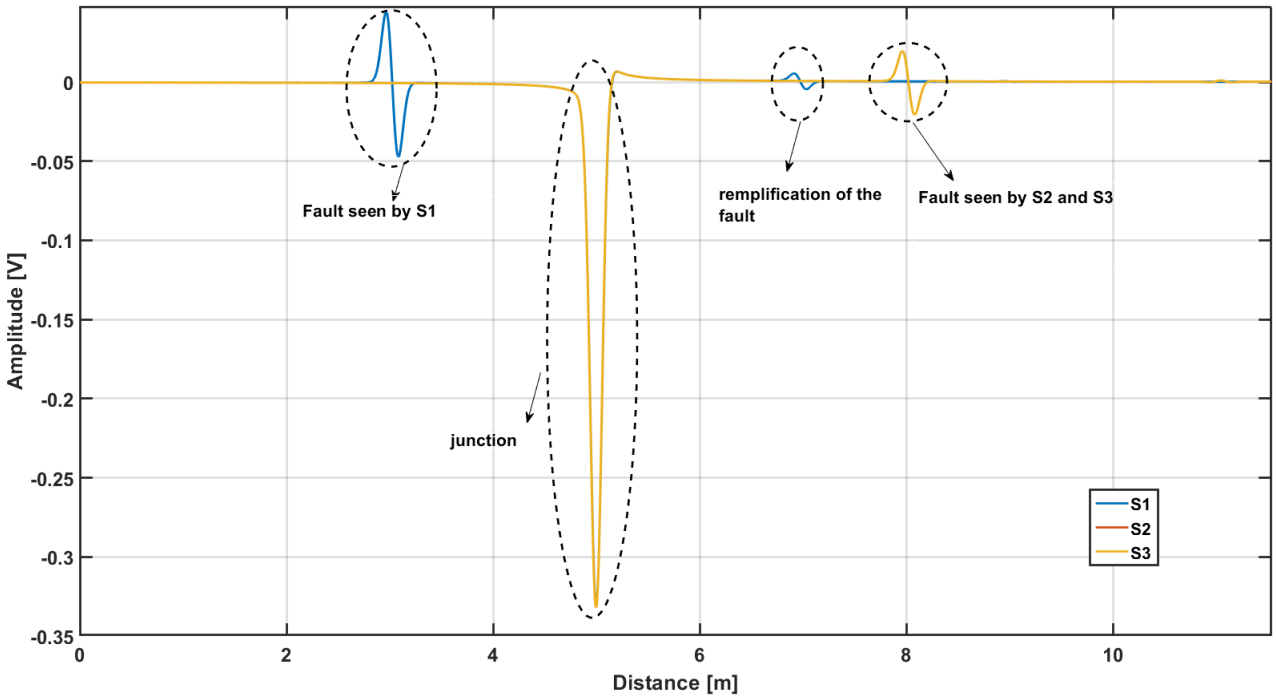


Figure 4.6 – TDR responses of the Y topology, first case, position P1

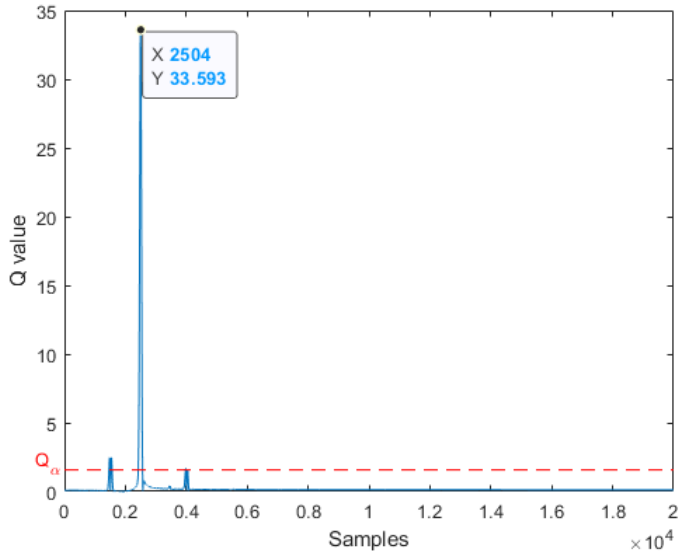


Figure 4.7 – Q chart of the new measured samples of the Y topology, first case, position P1

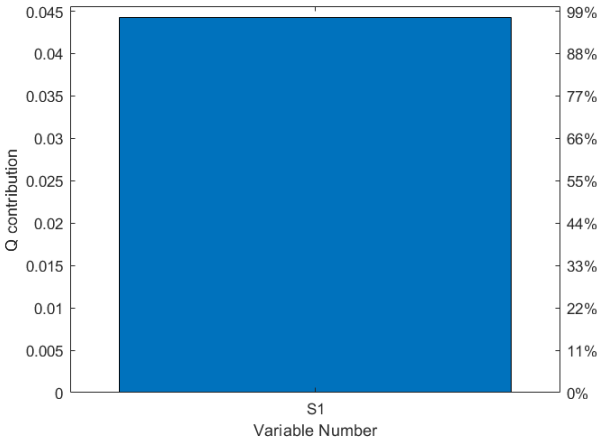


Figure 4.8 – Contribution plot of the sample 2504

Each abnormal sample in the Q plot corresponds either to the fault peak with respect to one sensor or the fault's round trips. Hence, the contribution plot of this abnormal sample corresponds to the sensor **S1**. So, it is incorrect to see the contribution plot for this sample as all the fault peaks are coincidental at the same sample. Therefore, we should retain from the contribution plot just the highest contributions (in this case, **S1**) and not the other contributions since they are unrelatable. For our application example, the contributions of the second and third variable corresponding to **S2** and **S3**, respectively, are almost neglected with respect to the first variable (0.96% for each compared to 98.07%). Hence, it does not appear in the contribution plot.

4.3.2.1.2 Fault in Position P2 The second simulated fault is located at $2.5m$ from sensor **S2** in B_2. The TDR responses of the three sensors are shown in Figure 4.9. After calculating the Q value of these measurements (Figure 4.10), the contribution plot of the highest peak is represented in Figure 4.11. It is shown that sensor **S2** is selected.

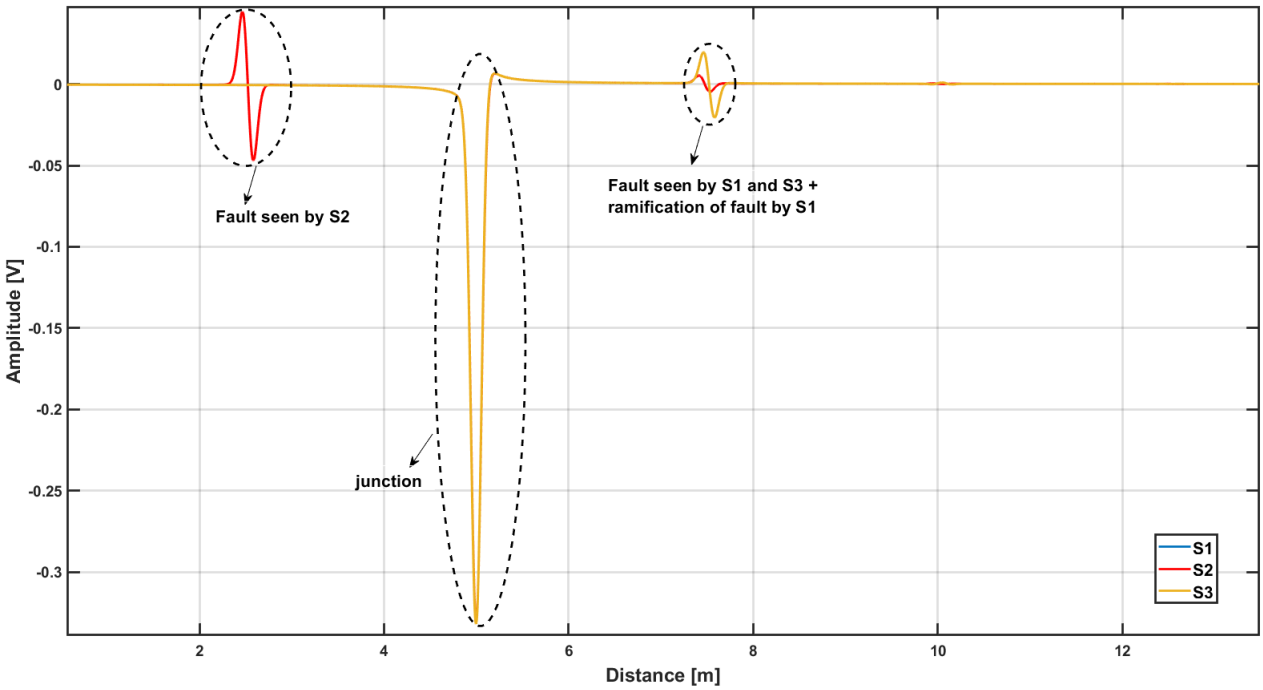


Figure 4.9 – TDR responses of the Y topology, first case, position P2

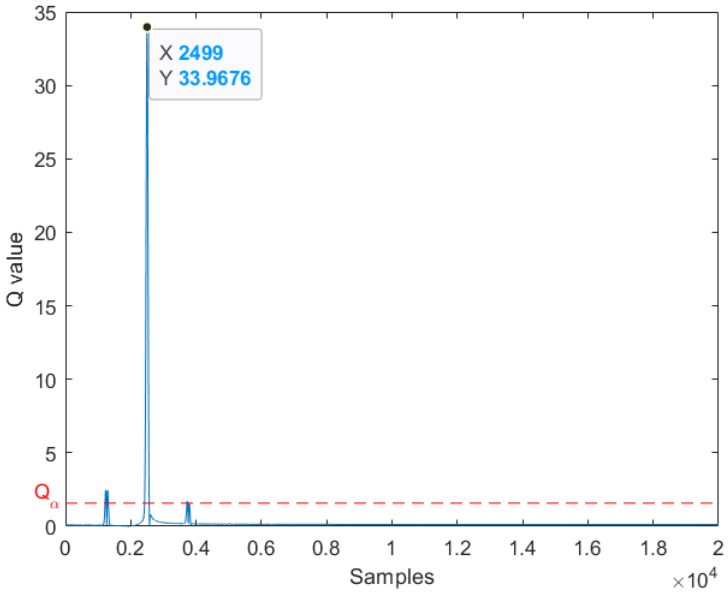


Figure 4.10 – Q chart of the new measured samples of the Y topology, first case, position P2

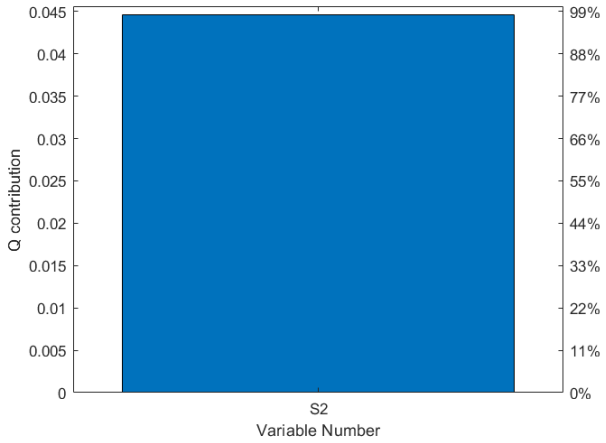


Figure 4.11 – Contribution plot of the sample 2499

4.3.2.2 Second Case: Not Equal Branch's length

For this second case, the branches length are different: $B_1 = 2m, B_2 = 1m, and B_3 = 2.3m$. Figure 4.12 shows the second topology and the different studied fault positions. Here, three fault positions are studied, by addressing one fault at a time. $X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3}]$ is the new measured data matrix where $\mathbf{x_j}$ describes the concatenated fault signature data vector for the sensor **Sj** at each fault position. E.g, for $j = 1$, $\mathbf{x_1} = [\mathbf{R_{S1@P1}} \quad \mathbf{R_{S1@P2}} \quad \mathbf{R_{S1@P3}}]^T$ where $\mathbf{R_{S1@P1}}$ describes the fault signature data vector measured by sensor **S1** at fault position P1.

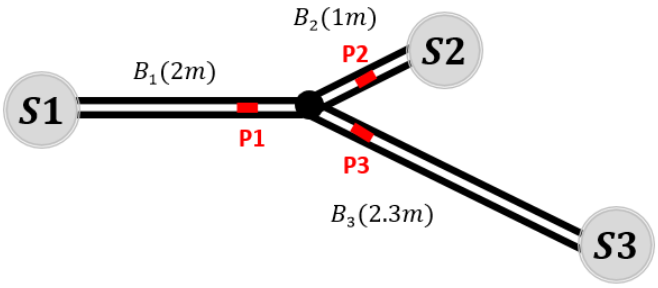


Figure 4.12 – Y topology: Second case

The first simulated fault is located at $1.5m$ from sensor **S1** in B_1. The second simulated fault is located at $0.5m$ from sensor **S2** in B_2. The last simulated fault is located at $0.5m$ from the junction in B_3. After calculating the Q values of these measurements (Figure 4.13), the contribution plots (Figure 4.14) of the highest peaks allow to select **S1**, **S2** or **S3** as the most relevant sensors for the fault at positions P1, P2 and P3, respectively. Hence, other sensors could be temporarily disabled. Any retained sensor could then be used to monitor the fault evolution for prognosis purposes based on its reflectometry measurements.

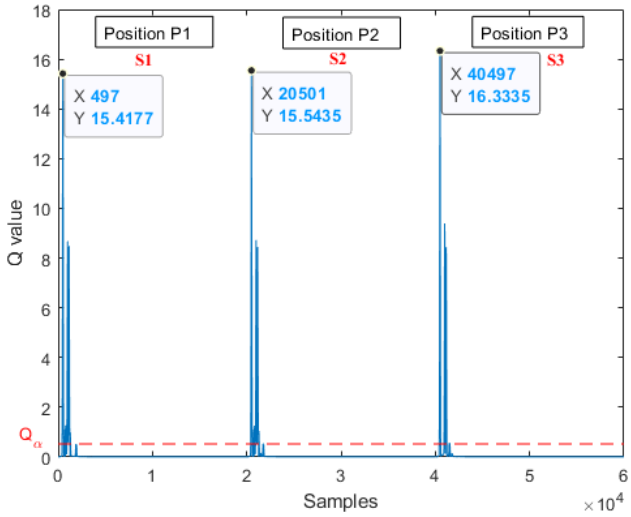
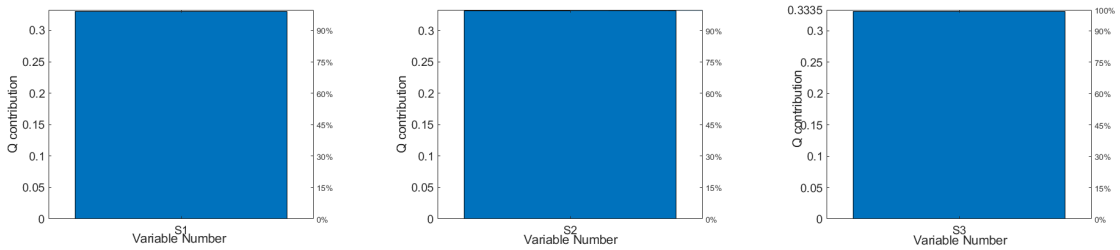


Figure 4.13 – Q chart of the new measured samples of the Y topology: Second case



(a) 497 (b) 20501 (c) 40497

Figure 4.14 – Contribution plot of the samples of the Y topology: Second case

4.3.3 . Obtained Results in the Case of CAN Bus Topology

The CAN bus is a common digital data network used in automotive, industrial, medical and scientific systems. The CAN bus is used for routing sensor data between pieces of equipment. The main advantages are high resilience to noise, reliability, low cost, simple wiring and ease of use. It is considered to be a complex wiring topology. Figure 4.15 represents the considered CAN bus topology. This network is composed of several sections, namely, B_1 to B_7. Their lengths are $B_1 = B_2 = B_5 = 2.5m$, $B_3 = B_6 = 5m$, and $B_4 = B_7 = 10m$. Six $1.5m$ cables, denoted by $B^{'}{}_1$ to $B^{'}{}_6$, are used to connect the Electronic Control Units (ECU) to the bus for accessing the network. The network consists of six sensors **Sj**, $j \in \{1, 2, \ldots, 6\}$ with the same characteristics (homogeneous network). These sensors are considered matched with the network cables where $Z_c = 100\Omega$.

First, in healthy fault-free operating conditions, each of the six sensors injects the test signal consecutively. TDR responses are then obtained and used to create the reference data matrix

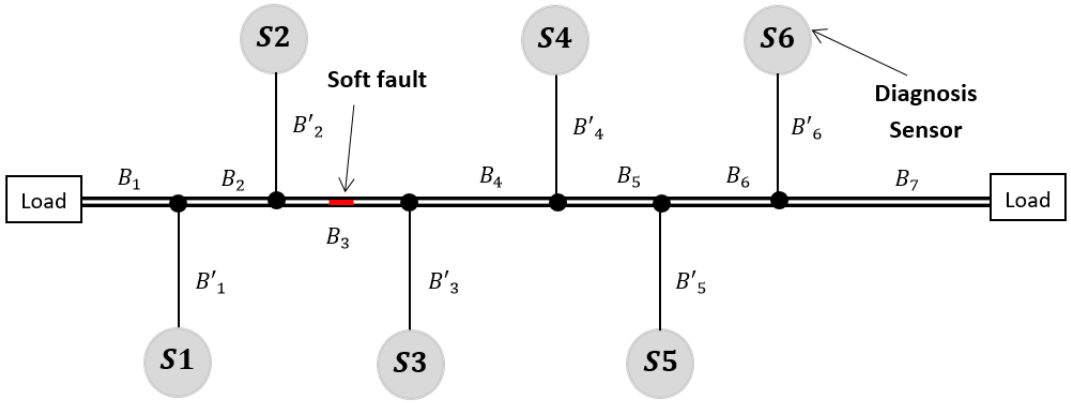


Figure 4.15 – CAN bus topology

X^* composed of six variables:

$$X^* = [\mathbf{x_1^*} \quad \mathbf{x_2^*} \quad \mathbf{x_3^*} \quad \mathbf{x_4^*} \quad \mathbf{x_5^*} \quad \mathbf{x_6^*}] \tag{4.4}$$

Each variable $\mathbf{x_j^*}$ corresponds to the CAN bus reference TDR response for the corresponding sensor **Sj** with $j = \{1, 2, \ldots, 6\}$.

The data matrix X^* is then used to obtain the PCA model; Table 4.1 indicates that the cumulative variance of two principal components. It is 92.28% for the first two ones, that is greater than 90%, which is the usual limit to be considered (see chapter 2 and [127]). This implies that the variables in X^* are highly correlated and that the data is supposed to be well described with the first two principal components.

Table 4.1 – PCA model variances

PC Number	Variance ratio %	Cumulated variance rate %
1	86.67	86.67
2	5.61	92.28
3	4.9	97.28
4	2.06	99.24
5	0.72	99.96
6	0.04	100

A 20% local variation of the characteristic impedance on the branch B_3 of the network is introduced to simulate a soft fault, i.e., $\Delta Z_c = 20\%$. TDR responses for each sensor are collected and are shown in Figure 4.16.

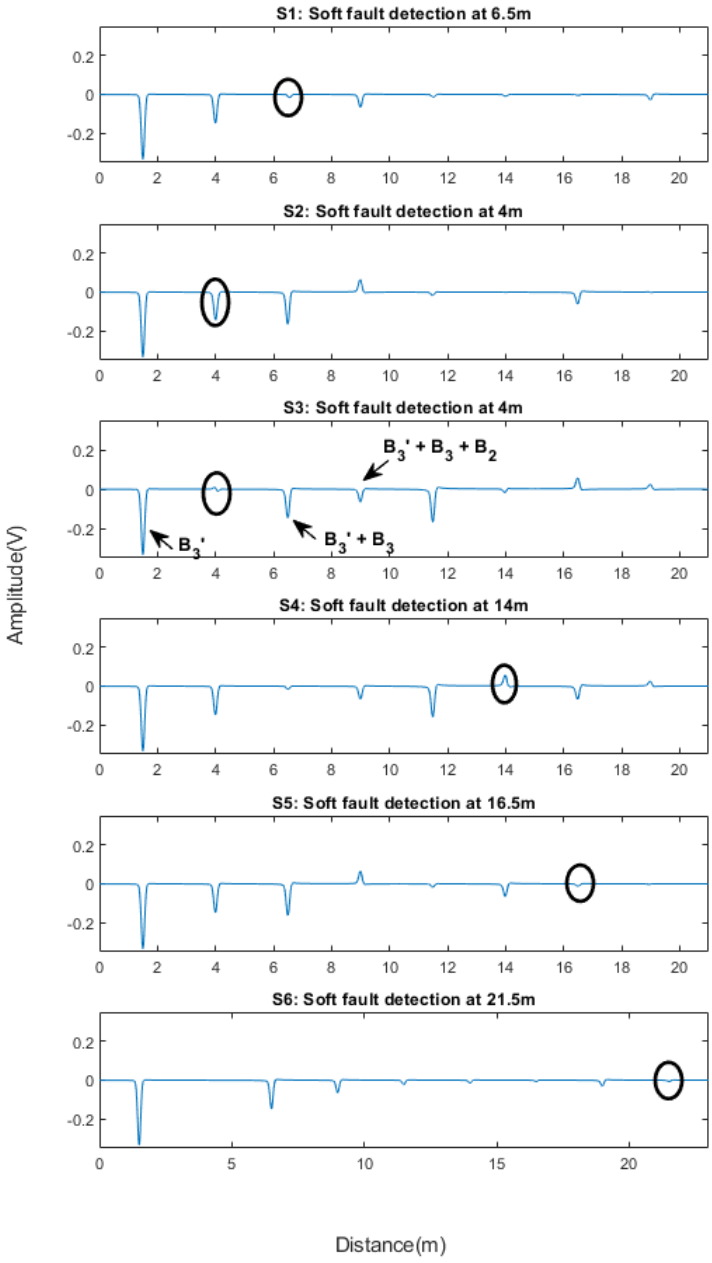


Figure 4.16 – TDR responses of the CAN bus topology

Considering the six sensors TDR responses, we can notice that the soft fault is located, respectively, at $6.5m$, $4m$, $4m$, $14m$, $16.5m$ and $21.5m$ from the corresponding sensors **S1** to **S6**. This soft fault with length $L_f = 0.05m$ is represented by a weak peak at those mentioned distances. Some peaks represent the ramifications on the network, the others represents the multiple echos on the network. As an example, for **S3**, the first peak obtained at $1.5m$ corresponds to the direct path to the junction (length of $B^{'}_3$, $l_{B^{'}_3} = 1.5m$). Using all these

faulty reflectometry data, the new measured data matrix X is created and defined as:

$$X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3} \quad \mathbf{x_4} \quad \mathbf{x_5} \quad \mathbf{x_6}] \tag{4.5}$$

where $\mathbf{x_j} = \mathbf{R_{Sj}}$ and $\mathbf{R_{Sj}}$ describes the faulty TDR response corresponding the sensor **Sj** with $j = \{1, 2, \ldots, 6\}$.

Now, the Q-statistic value for each new measured sample is calculated. Figure 4.17 represents the Q control chart, with the dashed red line representing the 95% confidence limit (Q_α). One can observe that some samples are above the threshold Q_α. Thus a fault occurrence has been detected. The first occurrence of the fault has the highest impedance variation in the reflectogram, and for each round-trip of this fault, its peak amplitude on the reflectogram decreases. Due to the presence of several round-trip peaks for the same fault, several abnormal samples are produced in the Q chart. On the other hand, the amplitude of the fault changes according to the sensor that we have used to do the measurements (due to the traveled distance between the sensor and the fault, and the existing junctions in between). Thus, each sensor will detect the fault peak, but its corresponding Q value will be different. Therefore, we consider the sample with the highest Q value as the abnormal one corresponding to the fault. Hence, in our example, sample number 3283 corresponds to the fault occurrence.

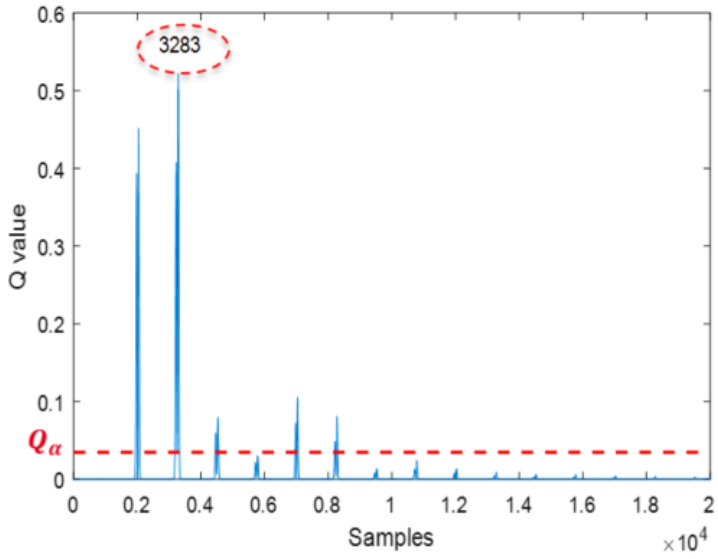


Figure 4.17 – Q chart of the new measured samples of the CAN bus topology

Plotting the contribution chart in Figure 4.18 of the indicated faulty sample permits to identify the variable $\mathbf{R_{Sj}}$ and the faulty sensor **Sj** that most influences this Q value. Thereby, the selection of the relevant sensor is performed by the Q control test. Each abnormal sample in the Q plot corresponds either to the fault peak with respect to one sensor or the fault's round trips. Hence, the contribution plot of this abnormal sample corresponds to the sensors **S2** and **S3** as they are equidistant from the fault (the peaks of the fault will coincide in the corresponding sample). So, it is incorrect to see the contribution plot for this sample as all the fault peaks are coincidental at the same sample. Therefore, we should retain from the

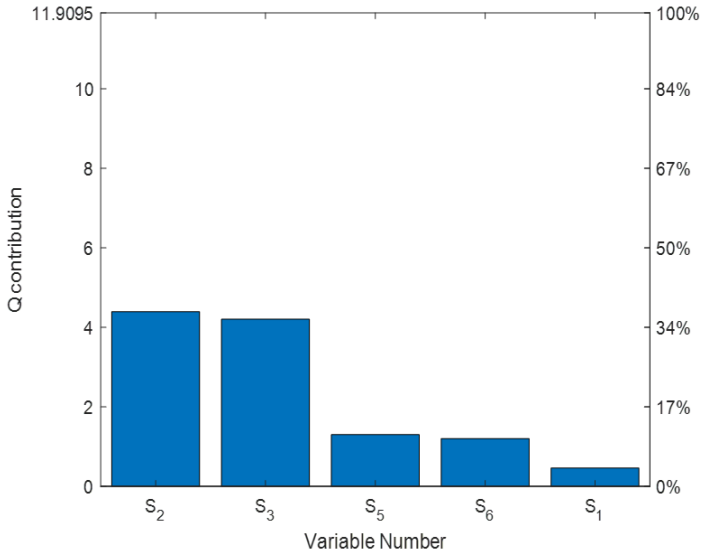


Figure 4.18 – Contribution plot for sample 3283

contribution plot just the highest contributions (in this case, **S2** and **S3**) and not the other contributions since they are unrelatable. For our application example, the contribution of the fourth variable corresponding to **S4** is almost neglected for the other variables ($< 1\%$). Hence, it does not appear in the contribution plot. Sensors **S2** and **S3** are then selected as the most relevant sensors, and other sensors could be temporarily disabled. Any retained sensor could then be used to monitor the fault evolution for prognosis purposes based on its reflectometry measurements.

4.4 . Experimental Validation

We propose to validate this methodology based on PCA for a fault detection and sensor selection on real cables using the experimental setup depicted in Figure 4.19.

4.4.1 . Test Bench Description

Figure 4.19 shows the considered system layout. It consists in multiple RG316 cables (modeled in chapter 3 section 3.2.1) with characteristic impedance 50Ω. In this study, B_1 to B_7 length's are $0.5m$, $0.5m$, $1m$, $1m$, $1m$, $0.5m$, and $0.5m$, respectively. The RG316 cables that ensure access to the network are denoted, respectively, $B^{'}{}_1$ to $B^{'}{}_6$ with lengths $1m$, $0.6m$, $2m$, $1.5m$, $1m$, and $1m$. The ends of lines B_1 and B_7 are matched using 50Ω resistors, where as the branches $B^{'}{}_1$ to $B^{'}{}_6$ are left open ended. Two soft faults with lengths $5mm$ and $10mm$ and shield cutaway sections $\theta_f = 180°$ are created in the middle of B_3. The reflected signals and the corresponding reflectograms are obtained using a VNA Agilent E5071C as shown in the setup above. The cable input terminal was attached with a 50Ω series RF coaxial cable connector for each measurement. The reflectometry responses were measured by the network analyzer separately at each end of the cable for the $30kHz$ to $4GHz$ band.

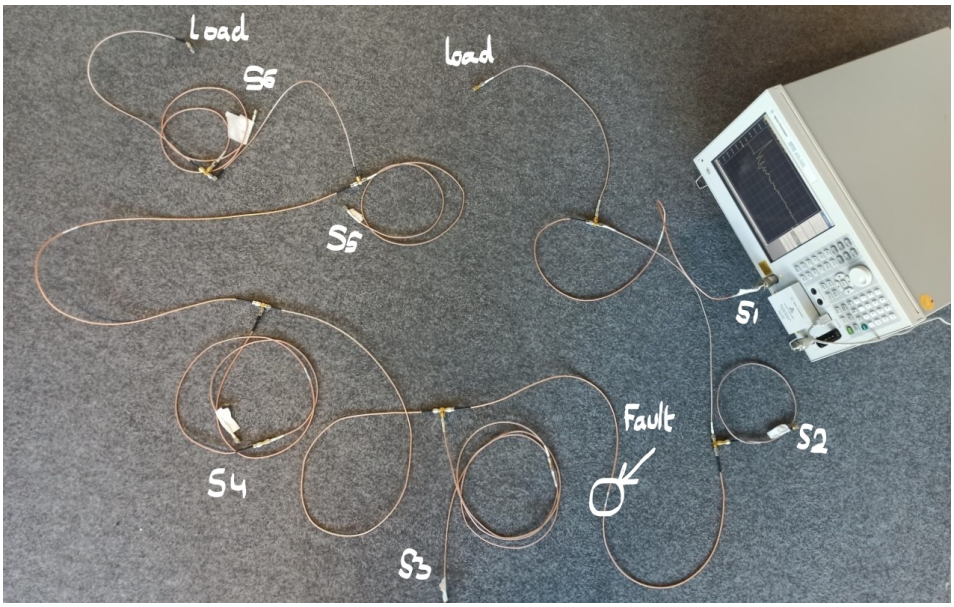


Figure 4.19 – Experimental Setup

4.4.2 . TDR-based Measurements

4.4.2.1 10mm Fault Case

Figure 4.20 shows the reflectogram obtained by **S1** in both, healthy and faulty networks. The peak at the input is due to a 50Ω N series RF coaxial cable connector attached to the cable's input terminal, it is called the mismatched impedance. The negative peaks correspond to the junctions located at $1m$ and $1.5m$ from the injection point, respectively. The soft fault is detected at $2m$ from sensor **S1**. Since the end of B'_2 is left open-ended, a positive peak appears at $2.1m$.

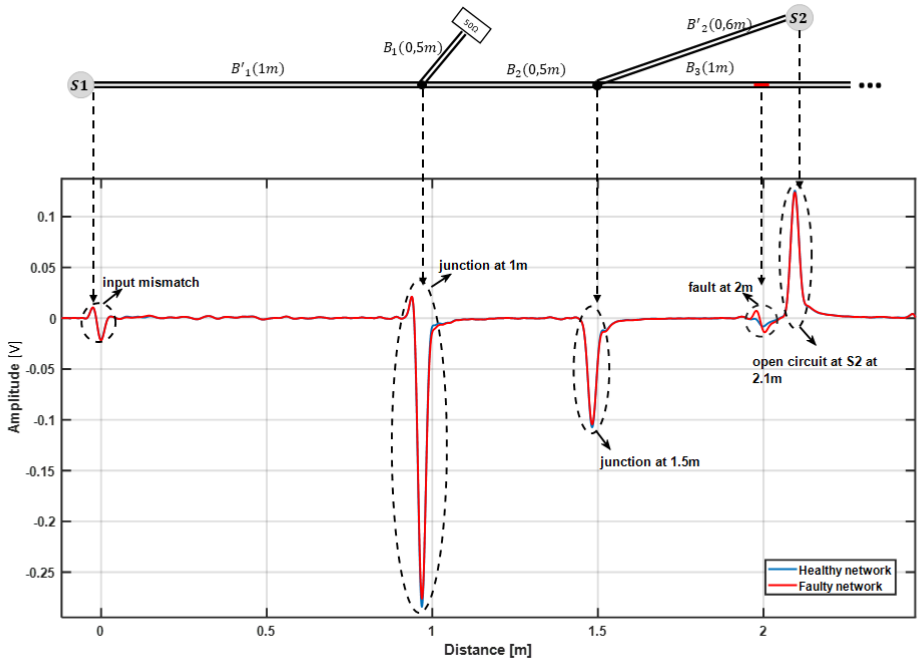


Figure 4.20 – Reflectogram of **S1**

Figure 4.21 shows the reflectogram obtained by **S2**. The negative peaks correspond to the junctions located at $0.6m$, $1.1m$, and $1.6m$ from the injection point. The soft fault is detected at $1.1m$ from **S2**, and it is masked by the junction at the same position.

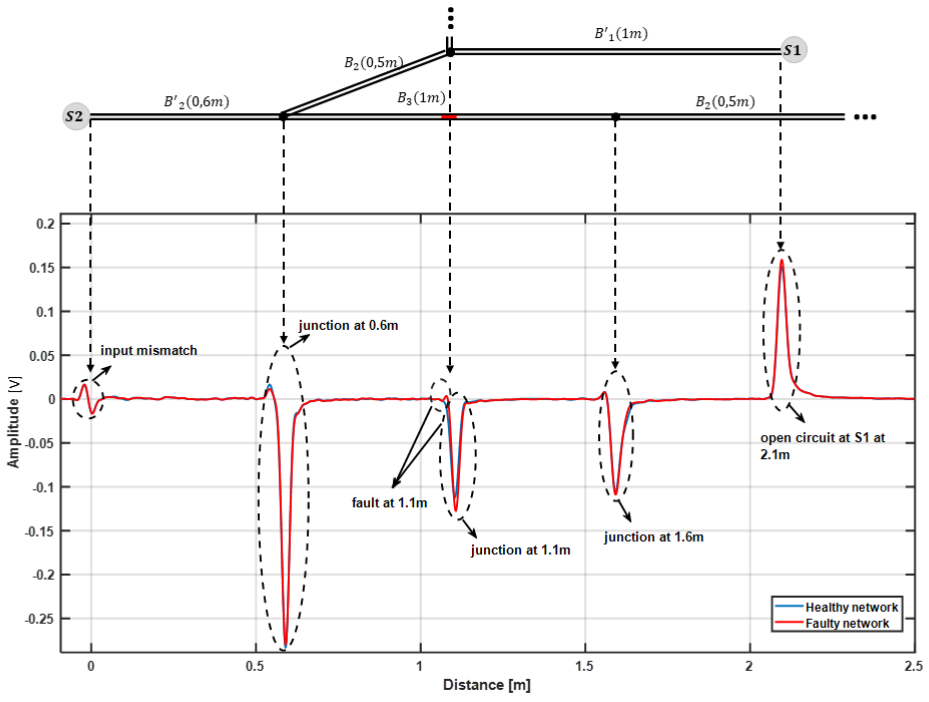


Figure 4.21 – Reflectogram of **S2**

Figure 4.22 shows the reflectogram obtained by **S3**. The second peak corresponds to the connector used to connect two $1m$ cables to form the $2m$ $B^{'}_{3}$. The negative peaks correspond to the junctions located at $2m$ and $3m$ from the injection point. The soft fault is detected at $2.5m$ from sensor **S3**.

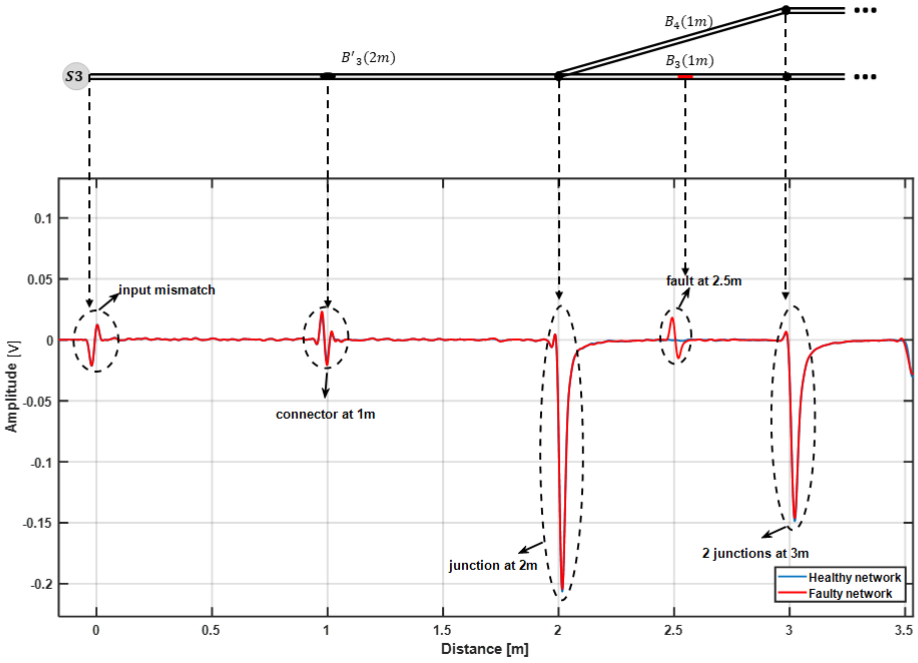


Figure 4.22 – Reflectogram of **S3**

Figure 4.23 shows the reflectogram obtained by **S4**. The second peak corresponds to the connector at $B^{'}_{4}$. The negative peaks correspond to the junctions at 1.5 and $2.5m$ from the injection point. The soft fault is detected at $3m$ from sensor **S4** and it is masked with the connector at the same position.

Figure 4.24 shows the reflectogram obtained by **S5**. The negative peaks correspond to the junctions at $1m$, $1.5m$, $2m$, and $3m$ from the injection point. The positive peak at $2.5m$ corresponds to an open circuit at the end of $B^{'}_{6}$. The soft fault is detected at $3.5m$ from **S5** and it is masked with the open circuit at the end of $B^{'}_{4}$.

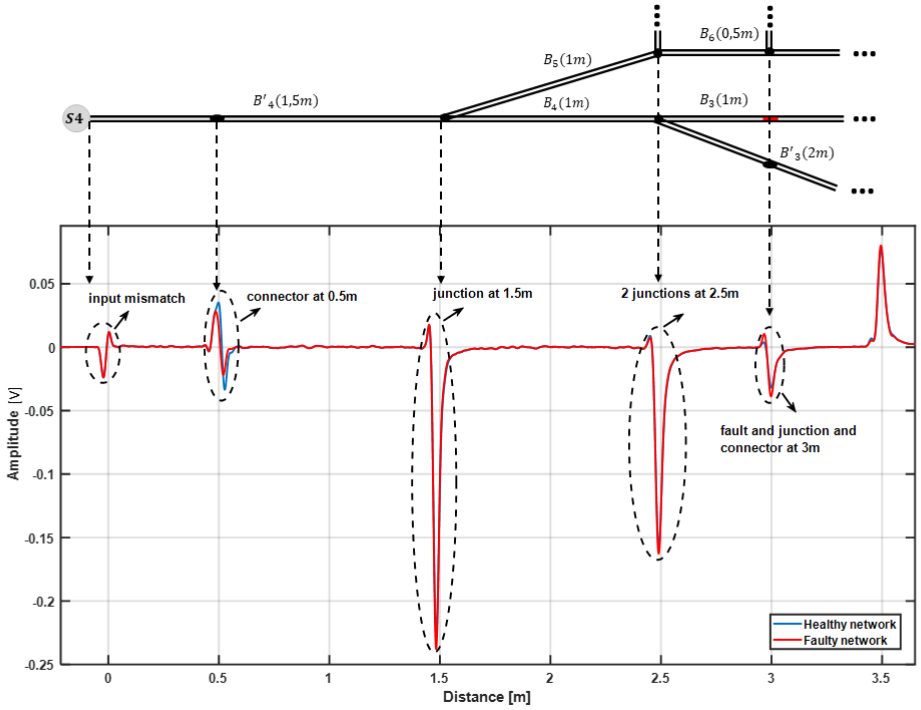


Figure 4.23 – Reflectogram of **S4**

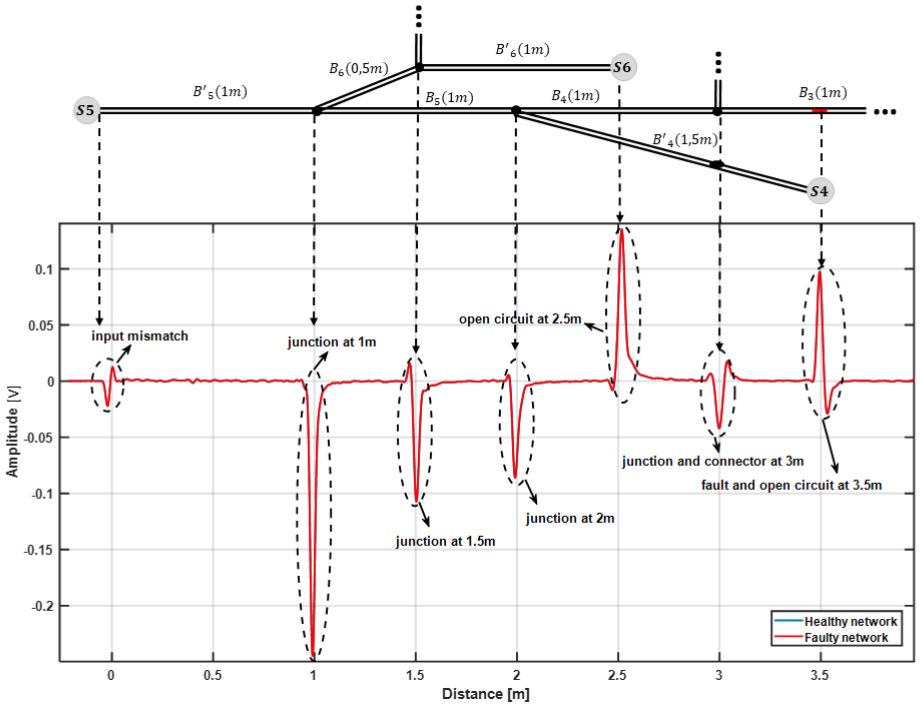


Figure 4.24 – Reflectogram of **S5**

Figure 4.25 shows the reflectogram obtained by **S6**. The negative peaks correspond to the junctions located at $1m$, $1.5m$, and $3.5m$ from the injection point. At $2m$ and $3m$, there are

multiple echoes from the junction. The peak at $2.5m$ corresponds to a junction and the open circuit at the end of $B^{'}_{6}$. The soft fault is detected at $4m$ from **S6** and it is masked with the open circuit at the end of $B^{'}_{5}$.

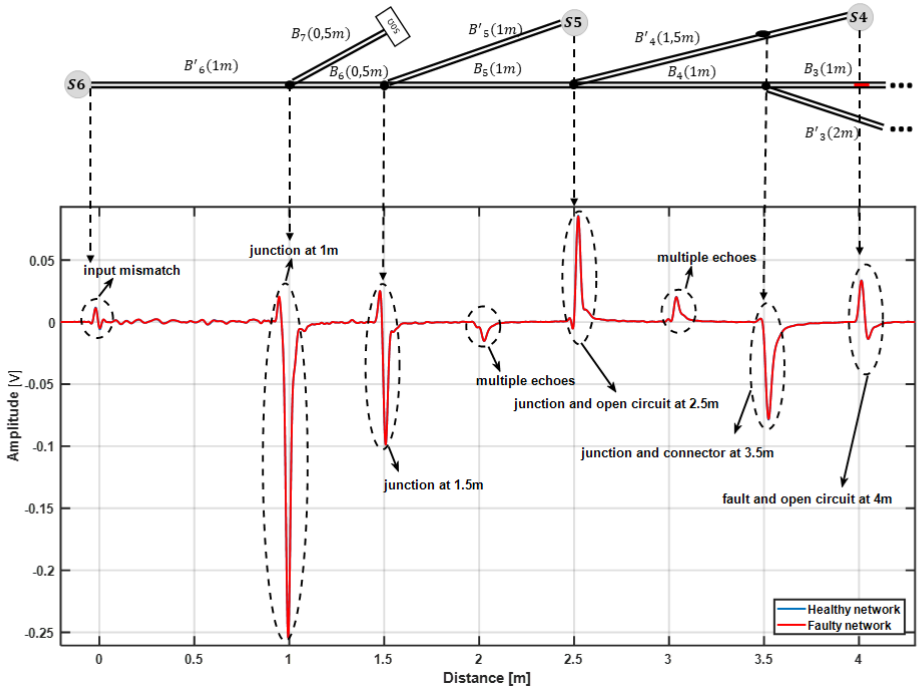


Figure 4.25 – Reflectogram of **S6**

For the above figures, it is obvious that the difficulty of analyzing the reflectometry response increases with the network's complexity involving the presence of one or several junctions, discontinuities, mismatched loads, etc. This leads to the degradation of the localization precision and consequently the quality of the diagnosis. In the case of soft faults, this problem will be more serious, and the fault may be hidden in various echoes of the signal as we will see also in the next section, for a smaller soft fault.

4.4.2.2 5mm Fault Case

Figure 4.26, Figure 4.27 and Figure 4.28 show the reflectogram obtained by **S1**, **S2** and **S3** respectively. It can be shown that the soft fault signature amplitude decreases significantly in comparison with the fault defined in case 1 (4.4.2.1). Thus, it is trickier to deal with as its signature could be more easily masked in various echoes of the signal as in Figure 4.27. This leads to the degradation of the localization precision and consequently the quality of the diagnosis.

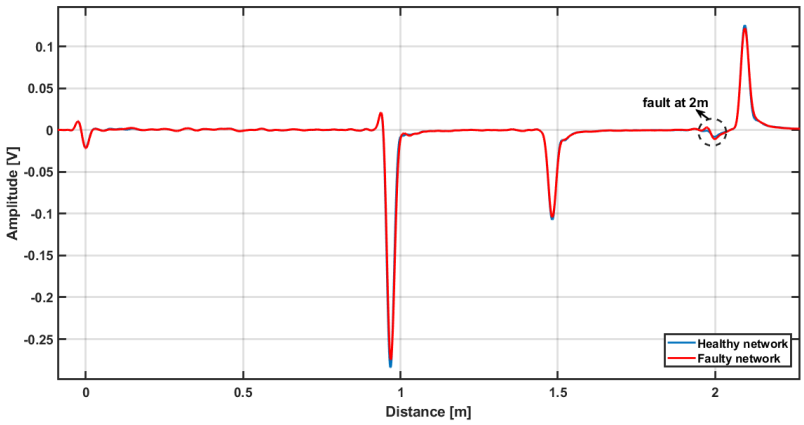


Figure 4.26 – Reflectogram of **S1**

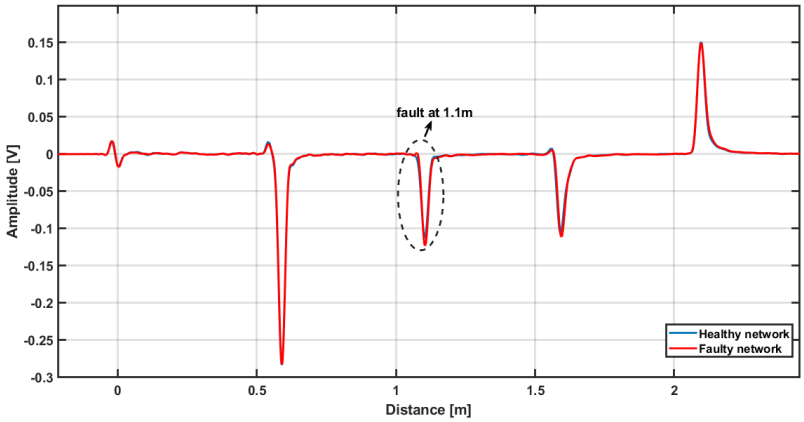


Figure 4.27 – Reflectogram of **S2**

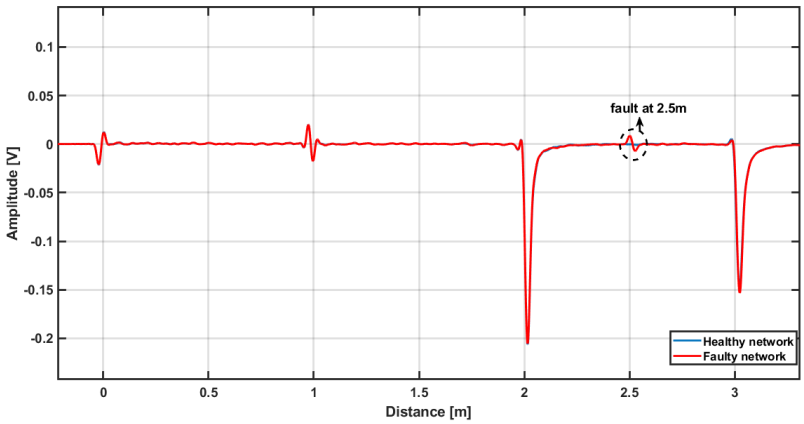


Figure 4.28 – Reflectogram of **S3**

The analysis of the reflectogram gets more and more complicated with the increasing num-

ber of branches in the presence of junctions, connectors and faults. Thus, the traditional way of analysis where the expert could detect the fault from analysing the reflectogram is not feasible. Hence, in the next section, PCA will be used first for fault detection then for the best sensor selection as explained in section 4.2.

4.4.3 . Sensor Selection based on Measurements

For this study, the data matrices are constructed as in section 4.3.3 where reflectometry measurements are done for both the reference and the faulty network for each of the 6 sensors.

The data matrix X^* is then used to build the PCA model. Table 4.2 displays that the cumulative variance of the principal components. The first component by itself explains less than 40% of the variance, so more components might be needed. It can be seen that the first five principal components explain more than 90% of the total variability in the standardized X^*, so that might be a reasonable way to reduce the dimensions. Thus, the number of the retained components, k, is equal to five.

Table 4.2 – PCA model variances

PC Number%	Variance ratio %	Cumulated variance rate
1	32.87	32.87
2	21.53	54.41
3	16.21	70.62
4	15.43	86.05
5	10.21	96.27
6	3.72	100

4.4.3.1 10mm Fault Case

Figure 4.29 represents the Q control chart, with the dashed red line representing the 95% confidence limit ($Q_\alpha = 4.39$). We can observe that some samples are above the threshold. Thus fault occurrence has been detected. We consider the sample with the highest Q value as the abnormal one. Therefore, in our example, the sample number 2806 corresponds to the fault occurrence.

Plotting the contribution chart (Figure 4.30) of the selected faulty sample permit to identify the variable $\mathbf{R_{Sj}}$ and the sensor **Sj** that most influences this Q value. Thereby, the selection of the relevant sensor is performed by the Q control test.

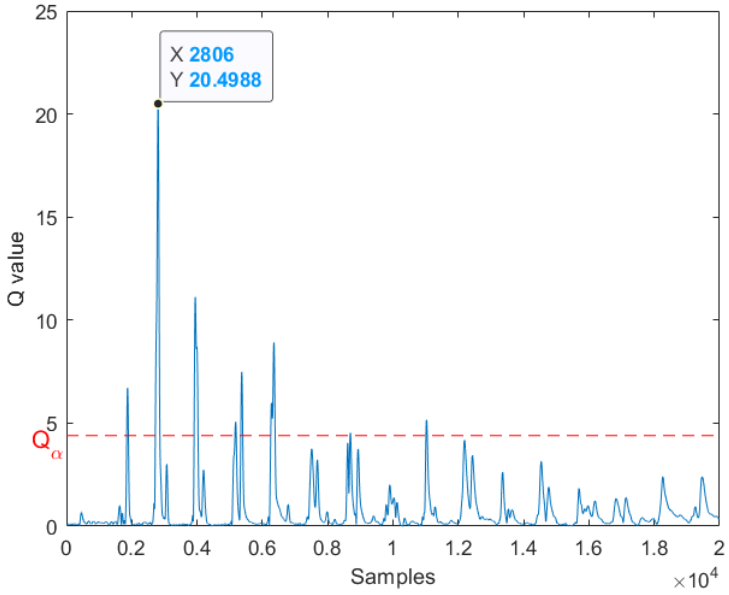


Figure 4.29 – Q chart of the new measured samples of the experimental measurements for the $10mm$ fault case

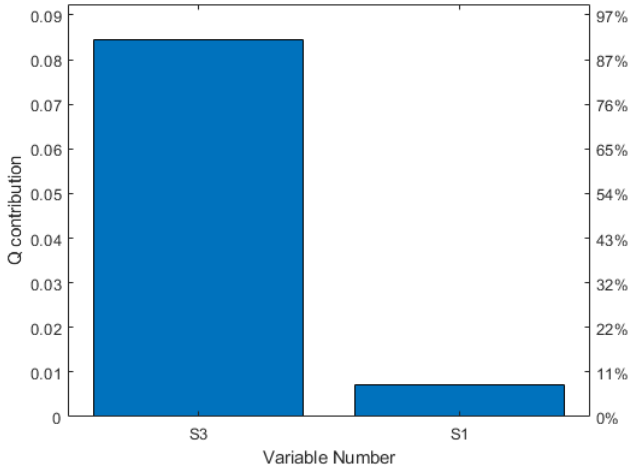


Figure 4.30 – Contribution plot for sample 2806

For our application example, sensor **S3** is selected as the most relevant one by the Q test. It is used to monitor the fault evolution for prognosis purpose based on its reflectometry measurements. As the signature of the soft fault in Figure 4.22 has the most significant amplitude, the choice of sensor **S3** is logical if we compare the traveled distance $(2.5m)$ and the number of the existing junctions (one junction) between this sensor and the fault. For **S2**, although it represents the lowest traveled distance from the fault $(1.1m)$, the fault is masked by the junction existing at the same position.

4.4.3.2 5mm Fault Case

Figure 4.31 represents the Q control chart, with the dashed red line representing the confidence limit. We can observe that for some samples the Q value is above the threshold. Thus fault occurrence has been detected. Sample number 2805 corresponds to the fault occurrence. The contribution chart of the indicated faulty sample is presented in Figure 4.32. For our application example, sensors **S3** is selected by the Q test, as the most relevant sensors and other sensors could be disabled.

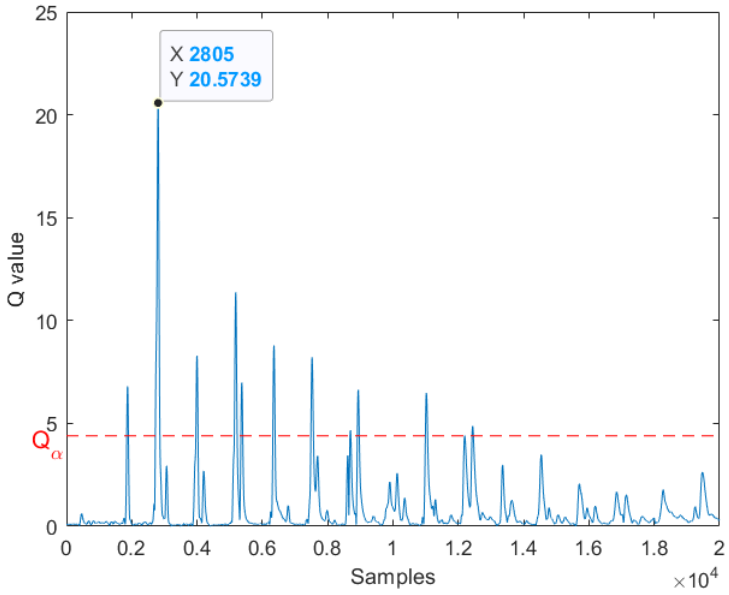


Figure 4.31 – Q chart of the new measured samples of the experimental measurements for the $5mm$ fault case

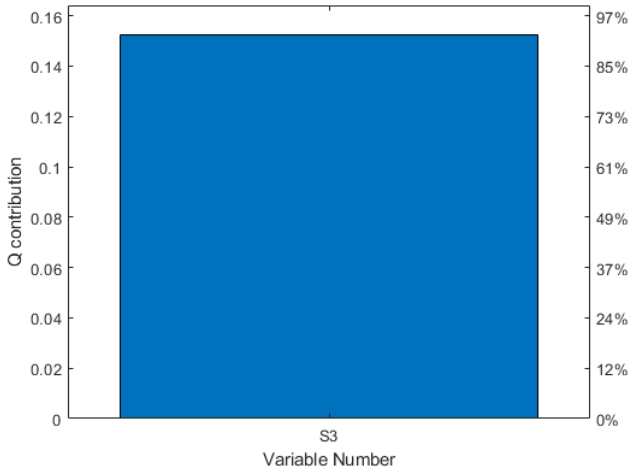


Figure 4.32 – Contribution plot for sample 2805

Based on these results, the sensor's number could be reduced, and the non-selected ones could be inactivated, reducing energy consumption, computing complexities, and sensor fusion problems and we can choose the most relevant sensor to monitor the evolution of this fault based on its reflectometry measurements. In what follows, the performance of this approach will be studied where it will be applied to a set of cables with different operating conditions.

4.5 . Performance Analysis in the presence of noise

The performance analysis flow chart of the sensor selection in the presence of noise that was represented in section 3.6 is used in this study. However, in this case, the selected parameter is the best sensor for monitoring and diagnosing the soft faults detected in the network in a noisy environment. This study is divided into three parts; first, the soft fault detection, where fault detection using PCA is applied and the probability of detection (P_d) is calculated in the presence of noise. Second, the false alarm evaluation, where the false alarm probability (P_{FA}) for each noise level is calculated. Finally, the best sensor selection where the sensor selection in the noisy environment is performed.

4.5.1 . Soft Fault Detection

4.5.1.1 Training Data Generation

The training data X^* is the data matrix that was defined in section 4.3.3 where the healthy fault-free CAN bus topology in Figure 4.15 is used.

4.5.1.2 Testing Data Generation: Noise introduction

To simulate different scenarios, two soft faults are studied. The first soft fault, F1, is the one introduced in 4.3.3 where a 20% local variation of the characteristic impedance on B_3 of the network is introduced, i.e., $\Delta Z_c = 20\%$. The position of this fault will be defined as x_{f1} at the middle of B_3 (at $2.5m$). The second soft fault, F2, has the same severity as the first fault but at a different position. The position of this fault will be defined as x_{f2}, located, respectively, at $21m$, $18.5m$, $13.5m$, $3.5m$, $2m$ and $7m$ from the corresponding sensors **S1** to **S6**.

For F2, TDR responses for each sensor are collected and are shown in Figure 4.33. This soft fault with length $L_f = 0.05m$ is represented by a weak peak. Some peaks represent the ramifications on the network, the others represent the multiple echos on the network.

The testing data X is the one defined in 4.3.3, however, $\mathbf{R_{Sj}}$ describes the concatenation of both faults TDR responses corresponding to the sensor **Sj** with $j = \{1, 2, \ldots, 6\}$. Now, the Q-statistic value of each new measured sample is calculated. The 95% confidence limit in this case is $Q_\alpha = 3.47$.

For this application example, **S2** is selected as the most relevant one for fault F1 and **S5** is

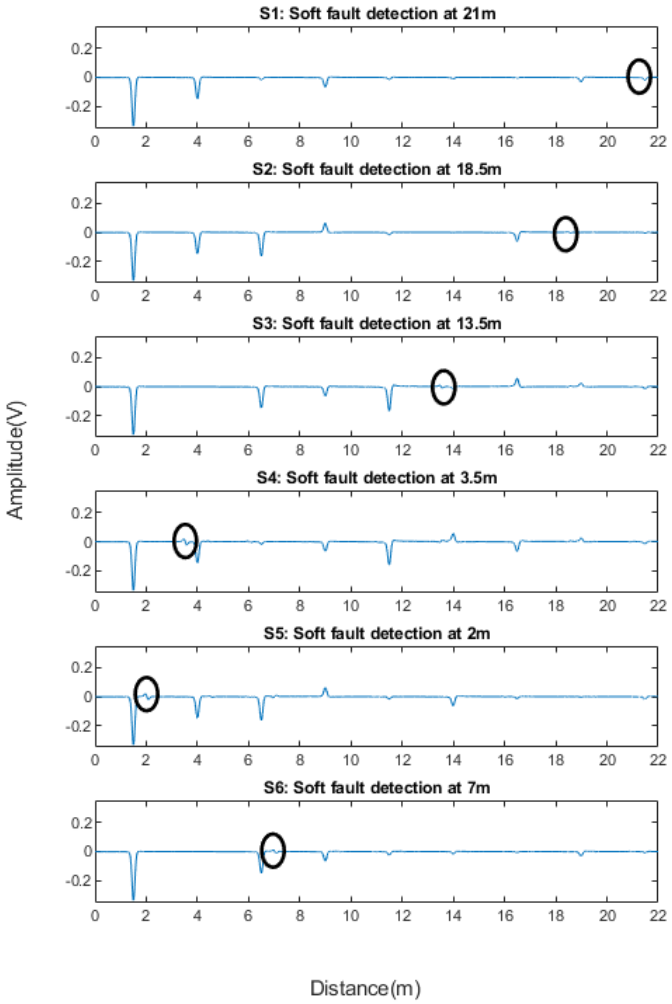


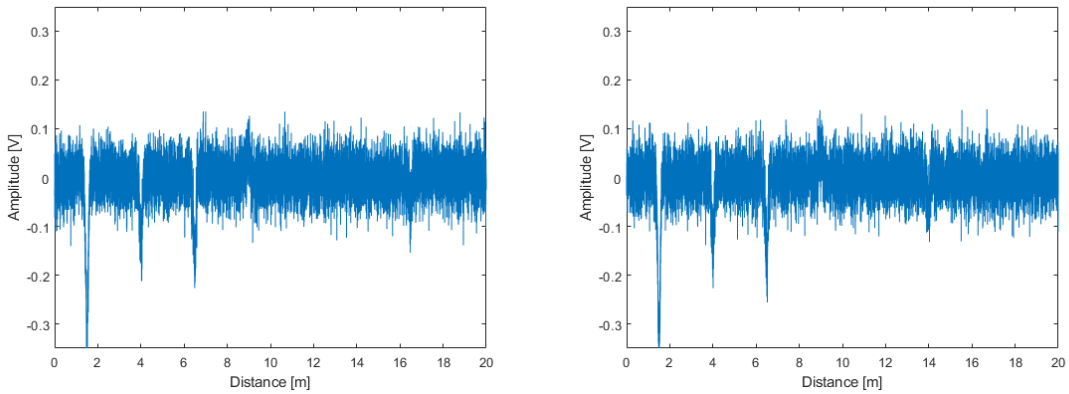
Figure 4.33 – TDR responses of the CAN bus topology in the presence of F2

selected as the most relevant one for fault F2; other sensors could be temporarily disabled.

Now, noises are added to the test data X, resulting in X_ν where the used SNR levels are $-5dB$, $0dB$, $5dB$, $10dB$ and $15dB$. $\mathbf{x_j}$ is a column vector of X taken for the jth variable. The noise is assumed to be Additive White Gaussian. The noise vector added to the variable $\mathbf{x_j}$ is $\nu_\mathbf{j} \sim N(0, p_\nu)$ ($\mathbf{x}_{\nu\mathbf{j}} = \mathbf{x_j} + \nu_\mathbf{j}$). Its power p_ν is related to the signal's power p_s by (4.6). 500 realizations are performed at each SNR level.

$$SNR = 10\log\left(\frac{p_s}{p_\nu}\right) \tag{4.6}$$

Figure 4.34 shows **S2** and **S5** TDR responses for F1 and F2, respectively, with an additional noise at SNR$= -5dB$.



(a) (b)

Figure 4.34 – TDR responses at SNR$= -5dB$ with (a) case F1, **S2** and (b) case F2, **S5**

4.5.1.3 Fault Detection

Sensor **S2** is studied for fault F1 and sensor **S5** is studied for fault F2. For each realization, we apply the Q test to find if the fault case is detected or not. Then, P_d is calculated for each fault case at each SNR, and if $P_d \geqslant \epsilon_d$, the fault is considered detected at this noise level. The two detected fault cases without noise, in the presence of noise, are evaluated with our method with a probability of detection equals to one. Here, ϵ_d is not considered as it depends on the application domain. As a hypothesis, we assume that the obtained detection probabilities are sufficient for the next steps in the performance analysis study.

Table 4.3 presents P_d of the fault cases F1 (middle of B_3) and F2 (at $2m$ on B_5) defined above for the different noise levels. It is shown that the same fault at different positions, in the presence of noise, is detected with a probability of detection equals to one. Hence, the Q test has excellent efficiency with 100% detection capability for the noise levels (SNR $\geq -5dB$), and its performances are not affected by the noise level for the two studied fault cases.

SNR [dB]	x_{f1}	x_{f2}
-5	1	1
0	1	1
5	1	1
10	1	1
15	1	1

Table 4.3 – P_d for different noise levels

4.5.2 . False Alarm Analysis

The probability of the false alarm P_{FA} value for the different noise levels is shown in Table 4.4. It is 0.012 for a noise level of $-5dB$ and 0.0014 for a noise level of $0dB$ and reduces

to 0 for the other SNR values.

SNR [dB]	P_{FA}
-5	0.012
0	0.0014
5	0
10	0
15	0

Table 4.4 – P_{FA} for different noise levels

The value of the false alarm threshold ϵ depends on the application domain. Hence, if $P_{FA} > \epsilon$, data X_ν needs to undergo a pre-possessing step; else, we proceed to the best sensor selection part. Here, ϵ is not considered as it depends on the application domain. As a hypothesis, we assume that the obtained false alarm probabilities are sufficient for the next steps in the performance analysis study.

4.5.3 . Sensor Selection

4.5.3.1 Robustness to Noise Evaluation

For each detected fault case, at a given SNR and for a specific realization, selecting the best sensor is attained using the contribution plot of the detected abnormal sample related to the fault case. Taking into account the 500 realizations, the selected sensor $\mathbf{Sj}_n$ is the sensor with the highest occurrence rate among the six used sensors.

For the case study above, Figure 4.35 represents the variation of the selected sensor with the noise level for different fault positions. It is noted that the chosen sensor to monitor the fault defined by the parameters ($\Delta Z_c = 20\%$, $L_f = 0.05m$) depends on its position and the current noise level.

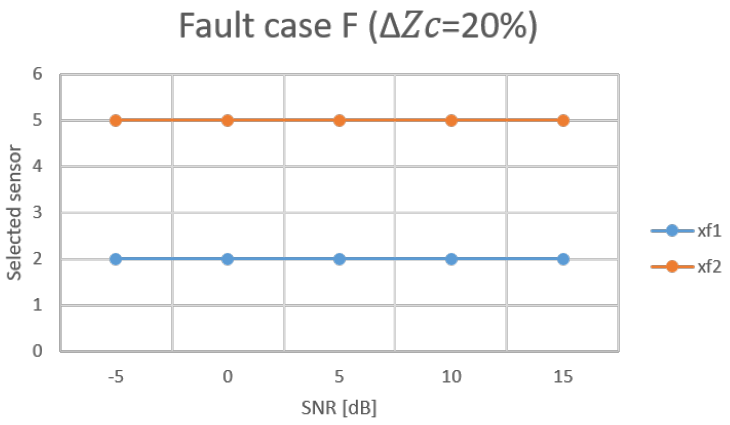
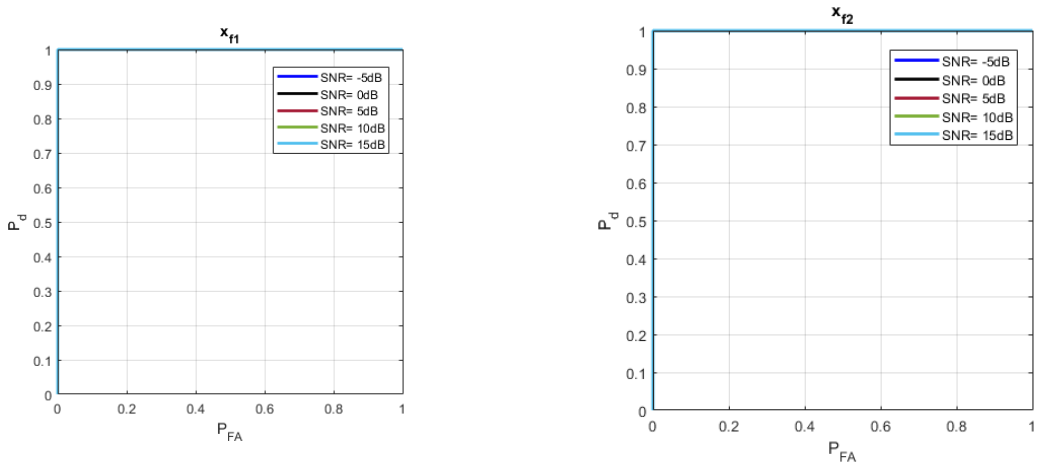


Figure 4.35 – Variation of the selected sensor with the noise level for two fault positions

4.5.3.2 ROC Curves Investigation

We consider the BUS network above for the fault case ($\Delta Z_c = 20\%$, $L_f = 0.05m$), at the two positions x_{f1} and x_{f2}. The sensor **S2** is investigated for the position x_{f1}, and the sensor **S5** is investigated for the position x_{f2}. The ROC curves for the different noise levels are represented in Figure 4.36 for the fault positions x_{f1} and x_{f2}, respectively.



(a) x_{f1} (b) x_{f2}

Figure 4.36 – ROC curves for CAN Bus with a $\Delta Z_c = 20\%$ and $L_f = 0.05m$ fault

It is noted that the Q test has excellent efficiency with 100% detection capability for the noise levels (SNR $\geq -5dB$) with a low false alarm probabilities for both faults' position. Moreover, the detection performances of Q test are not affected by the noise level. As the performance analysis of this method is investigated, it reveals excellent performances for soft faults detection. The probability of detection is equal to one for the used noise levels.

4.6 . Conclusion

The proposed method combines distributed Time Domain Reflectometry technique with Principal Component Analysis (PCA). For a given CAN bus topology, a distributed reflectometry approach is considered where sensors perform their reflectometry measurements consecutively. The TDR responses are the data that are collected and arranged into a database. With this database, a PCA model is developed and used to detect the existing soft faults. Coupled with statistical analysis based on squared prediction error Q, the most relevant sensors for monitoring and diagnosing the soft faults in the network are identified.

The efficiency of this approach was validated with several networks of different complexities,

starting with a simple topology, then Y topology, ending with the CAN bus topology in the presence of different fault scenarios.

Next, experimental validation was carried out. The experimental and the simulation results are coherent in terms of fault detection and sensor choice.

Finally, the performance analysis of this method is investigated. It reveals excellent performances for soft faults detection. The probability of detection is equal to one with a low false alarm probability for the different noise levels. It is noted that the chosen sensor to monitor a fault case depends on the fault parameters and the current noise level.

With this study, the sensor selection is obtained, whatever the fault location in the NUT. Therefore, the number of the sensors could be reduced, and the non-selected ones could be disabled (inactivated), reducing energy consumption, computing complexities, and mitigation of sensor fusion issues. It leads to saving time and objectivity in the decision-making. The selected sensors could be used to monitor the fault evolution for prognosis perspectives based on its reflectometry measurements.

General Conclusion and Perspectives

The research conducted in this thesis aims to propose new diagnosis techniques for multi-branched wired networks to detect and locate faults that manufacturers and users face today in electrical systems. The critical importance of soft faults has been emphasized and explained while showing the shortcomings of existing methods.

In the first chapter, we presented the context as well as the issues raised in this study. Indeed, the increased complexity of wired networks and the exposure of cables to various aggressive conditions favor the appearance of faults. Some faults can have serious consequences when cables are part of critical systems for which safety is an issue. Therefore, as an early-warning approach to ensure critical infrastructures' safe operation, it is essential to develop robust methods for identifying and locating faults. Within this context, two main groups for fault diagnosis have been highlighted: non-reflectometry methods and reflectometry-based ones. It has been observed that reflectometry-based methods are the most suitable. Ultimately, reflectometry-based-methods' performance evaluation has been elaborated. Although these methods exhibit good performance, they have several limitations. They shared an inability to provide efficient results with the network complexity increases. In the case of soft faults, the latter problem is even more worse; the fault may be easily hidden in various echoes of the testing signal. As a matter of fact, soft faults are often the premises of more severe faults which will eventually appear, and will be a potential source of incidents or accidents. Based on this, we have underlined our main focusing issues demonstrated by soft fault diagnosis using time-domain reflectometry.

For that, chapter two presents soft fault detection and localization problems in multi-branched networks using reflectometry methods, such as ambient noise, inhomogeneity of propagation, blind zones, and localization ambiguity. Several post-processing methods have been proposed in the literature. However, they remain prone to attenuation and dispersion inherent to wave propagation in transmission lines. Indeed, we have shown in this chapter that these phenomena significantly reduce the location accuracy when the propagation distance is long. Hence, the choice of the test signal bandwidth is critical and affects the diagnosis performance.
In the case of a multi-branched network, it is shown that distributed reflectometry solves the problem of localization ambiguities to the detriment of energy consumption, computing complexities, and sensor fusion problems. Thus, reducing the number of active sensors in the network is of great importance.
Finally, a throughout study about the existing enhanced methods in the literature for soft fault detection, localization, and characterization as well as their limitations is presented. This state of the art has made it possible to identify the main weaknesses of these methods, linked in particular to complex topologies and the detection of soft faults. Thus new tools used for fault detection and diagnosis are needed; for that, Principal Component Analysis (PCA) principle is presented.

The third chapter presents an efficient approach to select the best frequency bandwidth for

soft fault detection in wired networks based on a judicious combination of reflectometry and PCA. The proposed method allows configuring the VNA at different frequencies. It performs measurements at different frequencies for the reference case. After which, the PCA model is established. It performs the new measurements at different frequencies. If a difference is detected between the reference and the measured data, the contribution for each variable (i.e. frequencies) to this difference is calculated. The algorithm then chooses the most relevant frequency to monitor the soft fault. The advantages are thus time-saving and enable automatic computerized decision-making.
The three-dimensional modeling of a soft fault and its influence on the propagation in a cable has been studied with CST. The simulation results made it possible to determine the disturbance generated by the soft fault with different levels of degradation. These disturbances have been represented in terms of the reflection coefficient to determine the fault characteristics. To overcome the increase in the computational burden and the simulation time caused by the CST in the case of longer cables, the fault parameters are extracted using an $RLCG$ characterization method and are implemented in a Mathworks Matlab code to simulate different scenarios where the cable length l and the fault position x_f are changed. The simulation results of this method are compared to the output of the CST simulator with a high degree of success, and the Matlab code was tested for its ability to retrieve (estimate) the TDR response. Experimental measurements had been carried out to validate the numerical model in the case of shielding damage. Once validated, the developed model can be, hence, used later to validate the proposed methodology.
This method is then investigated for a set of cable lengths with different operating conditions and different fault types. The simulation results are in coherence with the rule of thumb inherent to short cables, where the higher the excitation frequency, the better is the fault detection resolution. In addition, as the fault reaches the cable end for longer cables, the selected frequency decreases due to higher attenuation and dispersion.
Experimental validations in the presence of soft faults were carried out. The experimental and simulation results are coherent in terms of fault detection and frequency choice. The performance analysis of this method has also been investigated. It has revealed good performance for soft fault detection. When the soft fault is near the injection point, the detection probability equals one even when as SNR values as low as $0\ dB$. As the fault position approaches the end of the cable, the performance is still good, but for lower fault severities, the detection is more tedious. In this study, it is noted that the selected frequency to monitor a fault case depends on several parameters: the cable type and characteristics, the fault severity and position, and the current noise level. Regarding the statistical tests, the Q criteria is more relevant for evaluating the faults.

In the fourth chapter, we proposed an efficient approach to select the best sensor for soft fault detection in multi-branched wired networks by combining distributed TDR technique with PCA. With this study, the sensor selection is obtained, whatever the fault location in the NUT. Therefore, the number of sensors could be reduced, and the non-selected ones could be temporarily disabled (inactivated), reducing energy consumption, computing complexities, and mitigation of sensor fusion issues. It leads to saving time and objectivity in the decision-making. The efficiency of this approach was validated with several networks of different complexities, starting with a simple topology, then Y-topology, ending with the CAN bus configuration in the presence of different fault scenarios.

For a given CAN bus configuration, a distributed reflectometry approach is considered where sensors perform their reflectometry measurements consecutively. The TDR responses are collected and arranged into a database. With this database, a PCA model is developed and used to detect soft faults. Coupled with statistical analysis based on squared prediction error, the most relevant sensors for monitoring and diagnosing the soft faults in the network are identified. This has been followed by a campaign of experimental validations. The experimental and the simulation results are coherent in terms of fault detection and sensor choice.
Finally, the performance analysis of this method is investigated. It reveals excellent performance for soft fault detection. The probability of detection is equal to unity with a low false alarm probability for the different noise levels. It is noted that the selected sensor to monitor a fault depends on the fault parameters and the current noise level.

Perspectives: We opted to work in this thesis with TDR and PCA to address the problems of frequency and sensor selection for soft fault diagnosis in multi-branched wired networks. As a matter of fact, efficient and significant results have been obtained with different soft fault severities and characteristics in different network structures and in the presence of noise. With this in mind, an extended study could be conducted to continue or handle other features that weren't addressed in this work and increase the sensitivity to other faults. Some of these perspectives are listed in the following:

For this study, we have considered an additive white Gaussian noise when calculating the SNR; however, depending on the system and its environment, several other types of noise more relevant in industrial applications can be considered, such as impulsive noise in vehicular networks, mechanical noise as in airplanes, etc. These require the establishment of models that shall be integrated with the analytical models of the fault detection so that to evaluate their effect. Post-processing approaches are then needed to mitigate their effect and enable a precise localization of the fault. The impact of different types of noise on performance needs to be studied.

The study carried out in this thesis resulted in a methodology for diagnosing soft faults representative of coaxial lines. The cable shielding was analyzed. Subsequently, it shall be interesting to address other types of state-of-art cables used in different applications. It is also necessary to extend the studies to other faults depending on their importance. For example, there exist wiring faults that occur commonly in aircraft and are referred to as intermittent arc faults (about 37%). These are the most frustrating, mysterious, and extremely difficult faults to detect and locate because they can appear in a few milliseconds due to vibrations, for example, and then disappear. In avionics, intermittent faults are very short-lived faults (around 1 microsecond) that can appear during flight, for example, but which are not easy to reproduce during aircraft maintenance. As an example, we can cite the partial discharge which is an intermittent phenomenon and in most cases invisible. This type of faults is no longer taken lightly by manufacturers, given the colossal damage they can cause in the long term. Thus, it will be interesting to see how we can apply the proposed approaches for these faults and improve the proposed detection methods.
On the other hand, some faults with smaller severities were not detected by PCA in chapter 3. Thus, further studies should be made to see how the reflectometry data could be pre-processed before applying the PCA method. In addition, the non-linearity character of soft faults in their earlier stages (incipient faults) could be the cause of the problem in this case, because their

variation is so small, so other techniques rather than the PCA could be considered.

The main line of research is the prognosis. Indeed, reflectometry methods generally make it possible to detect the presence of the fault after its occurrence. However, it would be more interesting to predict the appearance of the fault. Another line of research would be to use the selected sensors in chapter 4 to monitor the fault evolution for prognosis perspectives based on its reflectometry measurements.

It would be desirable to carry out a thorough study on the different fault types' characterization and identify their nature (mechanical, thermal, electrical). In the literature, the authors in [137] uses ANN to characterize the fault (faults' resistance). However, it is applicable only for hard faults. Authors in [113] characterize the fault (faults' impedance). However, it requires increasing the number of examples in the database and optimizing the number of neurons in the hidden layer. Authors in [138] use Random Forest to characterize the hard faults. Thus further studies should be made on the soft faults. Furthermore, traditional machine learning theories do not apply to the increasingly grown data because of the low generalization performance, reducing diagnosis accuracy. Therefore, it is necessary to investigate diagnosis models that can simultaneously extract features from raw collected data and automatically recognize the health states of cables, which would make it possible to reduce the calculation times and the resources required in the context of on-board electronics.

The feasibility of the frequency and sensor selection methods was verified by simulation results and validated by experimental results. The different functions of the method are not yet integrated into an electronic card to implement an autonomous diagnostic system. The integration of all the algorithms developed during our study is now based on engineering work to implement a reliable and optimized system capable of adapting to the constraints of on-board systems.
Integrating the diagnosis system enables diagnosis of the wired network while it is in use by one or more systems simultaneously. This diagnosis, known as "online diagnosis", allows continuous network monitoring. It is especially useful in the detection and localization of intermittent faults [9]. Online diagnosis provides the ability to perform diagnosis alongside native network operation. Thus, it makes it possible to be in the real conditions of the system to establish a more in-depth diagnosis, such as the characterization of intermittent faults.
It is thus interesting to study how the proposed approaches can fit into a system where other protocols or signals share the same environment. It is, therefore, necessary to implement an improvement on the method to take into account the constraints that come with the "online" mode, such as the robustness to noise since the diagnosis system can be subjected to all kinds of disturbances coming either from the useful signals of the system itself or from adjacent sources and it should not be intrusive in the native functioning of the system. Indeed, reflectometry signals must not interfere with useful signals present on the network. This interference can distort the response of the system. Therefore, it is necessary to implement an improvement in the proposed methods in this work to operate in "online" mode.

Another line of research would be the application of the proposed approaches to the diagnosis of faults in power grids. Faults like short circuit conditions in the power system network result in severe economic losses and reduce the electrical system's reliability. Thus, it is interesting to see the results of the proposed approaches for such networks.

Scientific Productions and Patents

Patents

[1] N. Taki, W. Ben Hassen, N. Ravot, C. Delpha, and D. Diallo, "Procédé de détection de défauts non francs dans un câble par analyse en composantes principales," Numéro d'enregistrement : FR1904582, 2019.

[2] N. Taki, W. Ben Hassen, N. Ravot, C. Delpha, and D. Diallo, "Procédé et système de surveillance d'un réseau de câbles par analyse en composantes principales," Numéro d'enregistrement : FR1908991, 2019.

International Conference Papers

[1] N. Taki, W. Ben Hassen, N. Ravot, C. Delpha, and D. Diallo, "Frequency selection for reflectometry-based soft fault detection using principal component analysis," in Prognostics and Systems Health Management Conference (PHM), Paris, France, IEEE, May 2-5 2019.

[2] N. Taki, W. Ben Hassen, N. Ravot, C. Delpha, and D. Diallo, "Sensors Selection for Distributed Reflectometry-based Soft Fault Detection using Principal Component Analysis," in IEEE AUTOTESTCON, Maryland, USA, Aug 26-29 2019.

Bibliography

[1] HVAC & Construction. Wires and cables market size, share & trends analysis report by voltage (low, medium, high, extra-high), by installation (overhead, underground), by end use, and segment forecasts. Technical Report GVR-2-68038-704-9, June 2019. Accessed: 2021-01.

[2] Nexans. http://www.nexans.fr/. Accessed: 2021-01.

[3] Gilles Millet, Serge Bruillot, Didier Dejardin, Nicolas Imbert, Fabrice Auzanneau, Luca Incarbone, Marc Olivas, Loïc Vincent, Alain Cremezi, and Sylvain Poignant. Aircraft electrical wiring monitoring system. In *Embedded Real Time Software and Systems (ERTS2014), Toulouse, France*, Feb 2014.

[4] Capecodofficeconsultants. https://www.capecodofficeconsultants.com/minimize-your-utility-bills-part-5-purchase-your-modem/coax-cable/. Accessed: 2021-01.

[5] Actelser. https://actelsershop.com/es/blog/diferencias-entre-la-fibra-optica-plastica-y-otros-cables-n5. Accessed: 2021-01.

[6] ENTSOE and Europacable. Feasibility and technical aspects of partial undergrounding of extra high voltage power transmission lines. *European Network of Transmission System Operators for Electricity (ENTSOE), Brussels, Belgium*, May 2010.

[7] Kevin R Wheeler, I Xander Twombly, Kai F Goebel, and Phil F Wysocki. Aging aircraft wiring fault detection survey. Technical report, NASA Ames Research Center, Moffett Field CA 94035, June 2007.

[8] Labinal Power Systems. Harness bite : Analyse des mécanismes de dégradations des défauts et leurs effets électriques. Technical Report interne D2.1R1, 2009.

[9] Fabrice Auzanneau. Wire troubleshooting and diagnosis: review and perspectives. *Progress In Electromagnetics Research B*, 49:253 – 279, 2013.

[10] Josy Cohen and Nicolas Gregis. Method of determining lineal parameters of a transmission line, August 18 2020. US Patent 10,746,771.

[11] Wafa Ben Hassen. *Étude de stratégies de diagnostic embarqué des réseaux filaires complexes*. PhD thesis, Institut National Polytechnique de Toulouse, France, October 2014.

[12] Lance Allen Griffiths, Rohit Parakh, Cynthia Furse, and Brittany Baker. The invisible fray: A critical analysis of the use of reflectometry for fray location. *IEEE Sensors Journal*, 6(3):697–706, 2006.

[13] Wafa Ben Hassen, M Gallego Roman, B Charnier, N Ravot, A Dupret, A Zanchetta, and F Morel. Embedded omtdr sensor for small soft fault location on aging aircraft wiring systems. *Procedia Engineering*, 168:1698–1701, 2016.

[14] Wafa Ben Hassen, Fabrice Auzanneau, Luca Incarbone, François Pérès, and Ayeley P Tchangani. Distributed sensor fusion for wire fault location using sensor clustering strategy. *International Journal of Distributed Sensor Networks*, 11(4):538643, 2015.

[15] Wafa Ben Hassen and Moussa Kafal. Shielding damage characterization in twisted pair cables using omtdr-based reflectometry and inverse problems. In *2019 PhotonIcs & Electromagnetics Research Symposium-Spring (PIERS-Spring)*, pages 3093–3101. IEEE, 2019.

[16] Brett Portwood. Aircraft wiring service history. *Small*, 1500:80, 2003.

[17] Fabrice Auzanneau and Nicolas Ravot. Détection et localisation de défauts dans des réseaux filaires de topologie complexe. In *Annales des télécommunications*, volume 62, pages 193–213. Springer, 2007.

[18] J Schonfeld, O Greulich, A Patterson-Hine, L Lee, J Cockrell, and L Hofland. Wiring integrity research (WIRe) Pilot Study. *Doc. No. A0SP-0001-XB1*, August 2000.

[19] You Chung Chung, Nirmal N Amarnath, Cynthia M Furse, and John Mahoney. Capacitance and inductance sensors for location of open and short circuited wires. *IEEE Transaction on Instrumentation and Measurement*, 58(8), 2009.

[20] Hashem M Hashemian. State-of-the-art predictive maintenance techniques. *IEEE Transactions on Instrumentation and measurement*, 60(1):226–236, 2010.

[21] Fabrice Auzanneau. Diagnostic filaire-détection, localisation et caractérisation de défauts dans des réseaux filaires complexes. *Recherche et Innovation des Techniques de l'Ingénieur*, pages 1–11, May 2010.

[22] Cynthia Furse and Randy Haupt. Down to the wire : Aging, brittle wiring within aircraft poses a hidden hazard that emerging technologies aim to address. *IEEE spectrum*, 38(2):34–39, February 2001.

[23] Cynthia Furse, You Chung Chung, Rakesh Dangol, Marc Nielsen, Glen Mabey, and Raymond Woodward. Frequency-domain reflectometry for on-board testing of aging aircraft wiring. *IEEE Transactions on Electromagnetic Compatibility*, 45(2):306–315, 2003.

[24] Concettina Buccella, Mauro Feliziani, and Giuliano Manzi. Accurate detection of low entity cable faults by wavelet transform. In *2004 International Symposium on Electromagnetic Compatibility (IEEE Cat. No. 04CH37559)*, volume 3, pages 936–941. IEEE, 2004.

[25] G Cerri, R De Leo, L Della Nebbia, S Pennesi, V Mariani Primiani, and P Russo. Fault location on shielded cables: Electromagnetic modelling and improved measurement data processing. *IEE Proceedings-Science, Measurement and Technology*, 152(5):217–226, September 2005.

[26] Yong June Shin. *Theory and application of time-frequency analysis to transient phenomena in electric power and other physical systems*. The University of Texas at Austin, 2004.

[27] Maud Franchet, Nicolas Ravot, and O. Picon. Soft fault detection in cables using the cluster time-frequency domain reflectometry. *IEEE Electromagnetic Compatibility Magazine*, 2(1):54–69, May 2013.

[28] Soumaya Sallem and Nicolas Ravot. Self-adaptive correlation method for soft defect detection in cable by reflectometry. In *SENSORS, 2014 IEEE*, pages 2114–2117. IEEE, November 2014.

[29] Soumaya Sallem and Nicolas Ravot. Soft defects localization by signature magnification with selective windowing. In *2015 IEEE SENSORS*, pages 1–4. IEEE, November 2015.

[30] Mohamed Oumri, Qinghua Zhang, and Michel Sorine. A reduced model of reflectometry for wired electric networks. In *19th International Symposium on Mathematical Theory of Networks and Systems (MTNS), Budapest, Hungary*, July 2010.

[31] Leon Cohen. Wave propagation with dispersion and damping. In *Advanced Signal Processing Algorithms, Architectures, and Implementations XIV*, volume 5559, pages 201–220. International Society for Optics and Photonics, 2004.

[32] C Chen and L Roemer. Attenuation and dispersion compensation via cepstral processing. *IEEE transactions on acoustics, speech, and signal processing*, 29(5):1080–1084, 1981.

[33] Nicolas Ravot. *Analyse et diagnostic de réseaux filaires complexes par réflectométrie*. PhD thesis, Faculté des Sciences d'Orsay, France, Juin 2007.

[34] Wafa Ben Hassen, Fabrice Auzanneau, Francois Peres, and Ayeley Tchangani. A distributed diagnosis strategy using bayesian network for complex wiring networks. *IFAC Proceedings Volumes*, 45(31):42–47, 2012.

[35] Wafa Ben Hassen, Fabrice Auzanneau, François Pérès, and AP Tchangani. Ambiguity cancellation for wire fault location based on cable life profile. *IFAC Proceedings Volumes*, 47(3):9593–9598, 2014.

[36] Rosani Penha and J Wesley Hines. Using principal component analysis modeling to monitor temperature sensors in a nuclear research reactor. In *Proceedings of the 2001 maintenance and reliability conference (MARCON 2001), Knoxville, TN*, 2001.

[37] Mostafa-Kamel Smaïl, Yann Le Bihan, and Lionel Pichon. Fast diagnosis of transmission lines using neural networks and principal component analysis. *International Journal of Applied Electromagnetics and Mechanics*, 39(1-4):435–441, 2012.

[38] Cynthia Furse, Paul Smith, Chet Lo, You Chung Chung, Praveen Pendayala, and Kedarnath Nagoti. Spread spectrum sensors for critical fault location on live wire networks. *Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures*, 12(3-4):257–267, 2005.

[39] Adrien Lelong, M Olivas-Carrion, Virginie Degardin, and Martine Lienard. On line wire diagnosis by modified spread spectrum time domain reflectometry. In *Progress in Electromagnetics Research Symposium, PIERS 2008*, pages 182–186. _, 2008.

[40] Erkki Lakervi and Edward J Holmes. *Electricity distribution network design*. Peter Peregrins Ltd. on behalf of The Institution of Electrical Engineers, 2nd ed.. edition, 2003.

[41] B.J. McPartland, J.F. McPartland, and F.P. Hartwell. *McGraw-Hill National Electrical Code 2008 Handbook, 26th Ed.* McGraw-Hill Education, 2008.

[42] Anthony Manet. *Etude de défauts non francs sur des câbles en vue du diagnostic*. PhD thesis, Université Paris-Est, 2016.

[43] RK Rao Yarlagadda. *Analog and digital signals and systems*, volume 1. Springer, 2010.

[44] Sing-Yui King and N Halfter. Underground power cables. *S. Y. King and N. A. Halfter, Book, 1982, Many figs. and tables, 411, Longman Group Limited, Longman House, Burnt Mill, Harlow, Essex, 30 Pounds*, 1982.

[45] Louis Rochester. *Coaxial Cables Application Data*. Technical Publishers Incorporated, 1969.

[46] Thomas J Siekierka and Robert David Kenny. Twisted pair cable, March 31 1998. US Patent 5,734,126.

[47] A. H. Affel et G. Estill. Coaxial conductor system, May 1932. US Patent 1,856,204.

[48] Mark J. Wilson H. Ward Silver. *Analog and digital signals and systems*, volume 1. The American Radio Relay League, 2010.

[49] Martin J Van Der Burgt. Coaxial cables and applications. *Belden Electronics Division*, 2003.

[50] Alexander Graham Bell. Telephone-circuit, July 19 1881. US Patent 244,426.

[51] C Buccella, M Feliziani, G Manzi, and F Maradei. Prediction of voltage and current propagation in twisted wire pairs (TWPs) by a circuit model. In *2005 International Symposium on Electromagnetic Compatibility, 2005. EMC 2005.*, volume 1, pages 51–55. IEEE, 2005.

[52] R Pappas. Aging systems research program. In *DER Recurrent Seminar*, volume 6, pages 53–64, 2000.

[53] Mark Strickland Jim Blanche. Wiring integrity technology assessment. Technical report, Marshall Space Flight Center, NASA, December 2004.

[54] Flightglobal. https://www.flightglobal.com/, 2006. Accessed: 2021-01.

[55] C Teal and C Satterlee. Managed aircraft wiring health directly relates to improved avionics performance. In *19th DASC. 19th Digital Avionics Systems Conference. Proceedings (Cat. No. 00CH37126)*, volume 1, pages 3B6–1. IEEE, 2000.

[56] Concettina Buccella, Mauro Feliziani, and Giuliano Manzi. Identification and localization of defects in shielded cables by a numerical/experimental procedure. In *2003 IEEE Symposium on Electromagnetic Compatibility. Symposium Record (Cat. No. 03CH37446)*, volume 1, pages 213–218. IEEE, 2003.

[57] Stanley P Cygan and Javaid R Laghari. Effects of multistress aging (radiation, thermal, electrical) on polypropylene. *IEEE transactions on nuclear science*, 38(3):906–912, 1991.

[58] Yacine Mecheri, L Boukezzi, A Boubakeur, and M Lallouani. Dielectric and mechanical behavior of cross-linked polyethylene under thermal aging. In *2000 Annual report conference on electrical insulation and dielectric phenomena (Cat. No. 00CH37132)*, volume 2, pages 560–563. IEEE, 2000.

[59] Cynthia Furse, You Chung Chung, Chet Lo, and Praveen Pendayala. A critical comparison of reflectometry methods for location of wiring faults. *Smart Structures and Systems*, 2(1):25–46, 2006.

[60] Paul Smith, Paul Kuhn, and Cynthia Furse. Intermittent fault location on live electrical wiring systems. *SAE International Journal of Aerospace*, 1(1), 2009.

[61] Paul Smith, Cynthia Furse, and Jacob Gunther. Analysis of spread spectrum time domain reflectometry for wire fault location. *IEEE sensors journal*, 5(6):1469–1478, 2005.

[62] Florent Loëte, Sophie Noël, René Meyer, M Olivas, F Auzanneau, and D Chandon. Feasability of the detection of vibration induced faults in connectors by reflectometry. In *24th International Conference on Electrical Contacts*, pages 440–443. Saint-Malo, 2008.

[63] Huaibin Tang and Qinghua Zhang. An inverse scattering approach to soft fault diagnosis in lossy electric transmission lines. *IEEE transactions on antennas and propagation*, 59(10):3730–3737, 2011.

[64] William H Hayt and H William. Buck, engineering electromagnetics, 1983.

[65] Clayton R Paul. *Analysis of multiconductor transmission lines*. John Wiley & Sons, 2007.

[66] Robert E Collin. *Foundations for microwave engineering*. John Wiley & Sons, 2007.

[67] William H Hayt Jr, John A Buck, and M Jaleel Akhtar. *Engineering Electromagnetics*/*(SIE)*. McGraw-Hill Education, 2020.

[68] Florent Loete, Qinghua Zhang, and Michel Sorine. Inverse scattering experiments for electrical cable soft fault diagnosis and connector location. In *PIERS proceedings, Kuala Lumpur, MALAYSIA*, March 2012.

[69] Ruslan Papazyan, Per Pettersson, Hans Edin, Roland Eriksson, and U Gafvert. Extraction of high frequency power cable characteristics from s-parameter measurements. *IEEE Transactions on Dielectrics and Electrical Insulation*, 11(3):461–470, December 2004.

[70] Qinghai Shi, Uwe Tröltzsch, and Olfa Kanoun. Analysis of the parameters of a lossy coaxial cable for cable fault location. In *Eighth International Multi-Conference on Systems, Signals & Devices, Tunisia*, pages 1–6. IEEE, November 2011.

[71] James Baker-Jarvis, Eric J Vanzura, and William A Kissick. Improved technique for determining complex permittivity with the transmission/reflection method. *IEEE Transactions on microwave theory and techniques*, 38(8):1096–1103, 1990.

[72] Dean A Frickey. Conversions between s, z, y, h, abcd, and t parameters which are valid for complex source and load impedances. *IEEE Transactions on microwave theory and techniques*, 42(2):205–211, 1994.

[73] Qinghai Shi and Olfa Kanoun. System simulation of network analysis for a lossy cable system. In *International Multi-Conference on Systems, Signals & Devices*, pages 1–5. IEEE, 2012.

[74] Levent Sevgi and C Uluisik. A matlab-based transmission-line virtual tool: finite-difference time-domain reflectometer. *IEEE Antennas and Propagation Magazine*, 48(1):141–145, 2006.

[75] Kane Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on antennas and propagation*, 14(3):302–307, 1966.

[76] Allen Taflove, Susan C Hagness, and Melinda Piket-May. Computational electromagnetics: the finite-difference time-domain method. *The Electrical Engineering Handbook*, 3, 2005.

[77] Mostafa Kamel Smail, Lionel Pichon, Marc Olivas, Fabrice Auzanneau, and Marc Lambert. Detection of defects in wiring networks using time domain reflectometry. *IEEE Transactions on Magnetics*, 46(8):2998–3001, 2010.

[78] David M Pozar. Microwave engineering usa: John wiley & sons. 2009.

[79] Michael Bernard Steer. *Microwave and RF design: a systems approach*. SciTech Pub., 2010.

[80] Pedro LD Peres, Carlos R De Souza, and Ivanil S Bonatti. Abcd matrix: a unique tool for linear two-wire transmission line modelling. *International Journal of Electrical Engineering Education*, 40(3):220–229, 2003.

[81] Scott Huss. A mathematical and lumped-element model for multiple cascaded lossy transmission lines with arbitrary impedances and discontinuities. In *Proceedings of ISCAS'95-International Symposium on Circuits and Systems*, volume 3, pages 1844–1847. IEEE, 1995.

[82] Christophe Guiffaut, Alain Reineix, and Bernard Pecqueux. New oblique thin wire formalism in the fdtd method with multiwire junctions. *IEEE Transactions on antennas and propagation*, 60(3):1458–1466, 2011.

[83] Pierre Bonnet. *Resolution des equations de maxwell instationnaires et harmoniques par une technique de volumes finis. Application a des problemes de comptabilite electromagnetique*. PhD thesis, Clermont-Ferrand 2, 1998.

[84] Kevin R Wheeler, Tolga Kurtoglu, and Scott D Poll. A survey of health management user objectives related to diagnostic and prognostic metrics. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 48999, pages 1287–1298, 2009.

[85] N Kamdor and Cynthia Furse. An inexpensive distance measuring system for location of robotic vehicles. In *IEEE Antennas and Propagation Society International Symposium. 1999 Digest. Held in conjunction with: USNC/URSI National Radio Science Meeting (Cat. No. 99CH37010)*, volume 3, pages 1498–1501. IEEE, 1999.

[86] You Chung Chung, Cynthia Furse, and Jeremy Pruitt. Application of phase detection frequency domain reflectometry for locating faults in an f-18 flight control harness. *IEEE Transactions on Electromagnetic Compatibility*, 47(2):327–334, 2005.

[87] Adrien Lelong, Marc Olivas Carrion, Virginie Degardin, and Martine Lienard. Characterization of electromagnetic radiation caused by on line wire diagnosis. In *URSI Conference, Chicago, USA*. Citeseer, August 2008.

[88] V Taylor, M Faulkner, A Kalam, and J Haydon. Digital simulation of fault location on ehv lines using wideband spread spectrum techniques. *IEE Proceedings-Generation, Transmission and Distribution*, 142(1):73–80, 1995.

[89] V Taylor and M Faulkner. Line monitoring and fault location using spread spectrum on power line carrier. *IEE Proceedings-Generation, Transmission and Distribution*, 143(5):427–434, 1996.

[90] Ousama Osman. *Méthodes de diagnostic en ligne, embarqué et distribué dans les réseaux filaires complexes*. PhD thesis, Université Clermont Auvergne(2017-2020), 2020.

[91] N Ravot, F Auzanneau, Y Bonhomme, M Olivas, and F Bouillault. Distributed reflectometry-based diagnosis for complex wired networks. *EMC: Safety, Reliability and Security of Communication and Transportation Systems, Paris*, 2007.

[92] Adrien Lelong. *Méthodes de diagnostic filaire embarqué pour des réseaux complexes*. PhD thesis, Université des sciences et Technologies de Lille, France, December 2010.

[93] Hang Xu, Bingjie Wang, Jingxia Li, Anbang Wang, and Yuncai Wang. Location of wire faults using chaotic signal generated by an improved colpitts oscillator. *International Journal of Bifurcation and Chaos*, 24(04):1450053, 2014.

[94] Jian Guo Zhang, Hang Xu, Bing Jie Wang, Li Liu, Peng Cheng Su, and Jing Xia Li. Wiring fault detection with boolean-chaos time-domain reflectometry. *Nonlinear Dynamics*, 80(1):553–559, 2015.

[95] Fabrice Auzanneau. Binary time domain reflectometry: A simpler and more efficient way of diagnosing defects in wired networks. In *2018 IEEE AUTOTESTCON*, pages 1–8. IEEE, 2018.

[96] Fabrice Auzanneau. Natural amplification of soft defects signatures in cables using binary time domain reflectometry. *IEEE Sensors Journal*, 21(2):937–944, 2019.

[97] Lola El Sahmarany. *Méthodes d'amélioration pour le diagnostic de câble par réflectométrie*. PhD thesis, Université Blaise Pascal-Clermont-Ferrand II, France, 2013.

[98] Maud Franchet. *Réflectométrie appliquée à la détection de défauts non francs dans les torons de câbles*. PhD thesis, Université Paris-Est, 2012.

[99] Yong June Shin, Roger Dougal, and EJ Powers. Joint time-frequency domain reflectometry for diagnostics of coaxial cables. In *8th Joint NASA/FAA/DoD Conference on Aging Aircraft*, 2005.

[100] Philip Crapse, Jingjiang Wang, Yong-June Shin, Roger Dougal, Trang Mai, Joseph Molnar, and Lan Tran. Design of optimized reference signal for joint time-frequency domain reflectometry-based wiring diagnostics. In *2008 IEEE AUTOTESTCON*, pages 195–201. IEEE, 2008.

[101] J Wang, PEC Stone, Y-J Shin, and RA Dougal. Application of joint time–frequency domain reflectometry for electric power cable diagnostics. *IET signal processing*, 4(4):395–405, 2010.

[102] Jingjiang Wang, Philip Crapse, John Abrams, Yong-June Shin, Roger Dougal, Trang Mai, Lan Tran, and Joseph Molnar. Diagnostics and prognostics of wiring integrity via joint time-frequency domain reflectometry. In *10th Joint FAA/DoD/NASA Conference on Aging Aircraft*, 2007.

[103] Sandrine Roblot. *Caractérisation des couplages électromagnétiques dans les réseaux filaires cuivre en vue d'optimiser les transmissions à haut débit*. PhD thesis, Université de Limoges, 2007.

[104] Robert H Flake. Part i. theory. signal propagation without distortion on lossy transmission lines having frequency dependent parameters. In *Proceedings. 9th IEEE Workshop on Signal Propagation on Interconnects, 2005.*, pages 43–45. IEEE, 2005.

[105] Ousama Osman, Soumaya Sallem, Laurent Sommervogel, Marc Olivas Carrion, Arnaud Peltier, Pierre Bonnet, and Françoise Paladian. Method to improve fault location accuracy against cables dispersion effect. *Progress In Electromagnetics Research Letters*, 83:29–35, 2019.

[106] Mostafa Kamel Smail. *Développement d'une méthodologie dédiée à la réflectométrie en vue du diagnostic filaire*. PhD thesis, 2010.

[107] O Osman, S Sallem, M Olivas, L Sommervogel, P Bonnet, and F Paladian. Wired network distributed diagnosis and sensors fusion by multi-carrier time domain reflectometry and graph theory. In *2019 PhotonIcs & Electromagnetics Research Symposium-Spring (PIERS-Spring)*, pages 3086–3092. IEEE, 2019.

[108] Mostafa Kamel Smail, Lionel Pichon, Marc Olivas, Fabrice Auzanneau, and Marc Lambert. Reconstruction of faulty wiring networks using reflectometry response and genetic algorithms. *International Journal of Applied Electromagnetics and Mechanics*, 35(1):39–55, 2011.

[109] Ousama Osman, Soumaya Sallem, Laurent Sommervogel, Marc Olivas Carrion, Pierre Bonnet, and Françoise Paladian. Distributed reflectometry for soft fault identification in wired networks using neural network and genetic algorithm. *IEEE Sensors Journal*, 20(9):4850–4858, 2020.

[110] Pratibha Bajpai and Manoj Kumar. Genetic algorithm–an approach to solve global optimization problems. *Indian Journal of computer science and engineering*, 1(3):199–206, 2010.

[111] Moussa Kafal, Jaume Benoit, and Christophe Layer. A joint reflectometry-optimization algorithm for mapping the topology of an unknown wire network. In *2017 IEEE SENSORS*, pages 1–3. IEEE, 2017.

[112] Mostafa-Kamel Smaïl, Tarik Hacib, Lionel Pichon, and Florent Loete. Detection and location of defects in wiring networks using time-domain reflectometry and neural networks. *IEEE Transactions on Magnetics*, 47(5):1502–1505, 2011.

[113] Ousama Osman, Soumaya Sallem, Laurent Sommervogel, Marc Olivas, Arnaud Peltier, Pierre Bonnet, and Françoise Paladian. Distributed sensor diagnosis in complex wired networks for soft fault detection using reflectometry and neural network. In *2019 IEEE AUTOTESTCON*, pages 1–6. IEEE, 2019.

[114] Yifeng Zhou, Juergen Hahn, and M Sam Mannan. Fault detection and classification in chemical processes based on neural networks with feature extraction. *ISA transactions*, 42(4):651–664, 2003.

[115] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

[116] Yagang Zhang, Qian Ma, Jinfang Zhang, Jing Ma, and Zengping Wang. Fault diagnosis based on graph theory and linear discriminant principle in electric power network. *Wireless Sensor Network*, 2(1):62, 2010.

[117] DongKai Fan and Ping Shi. Improvement of dijkstra's algorithm and its application in route planning. In *2010 seventh international conference on fuzzy systems and knowledge discovery*, volume 4, pages 1901–1904. IEEE, 2010.

[118] Jing Wang, Jingdong Wang, Gang Zeng, Zhuowen Tu, Rui Gan, and Shipeng Li. Scalable k-nn graph construction for visual descriptors. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1106–1113. IEEE, 2012.

[119] Theodora Kourti. Process analysis and abnormal situation detection: from theory to practice. *IEEE control systems magazine*, 22(5):10–25, 2002.

[120] Thomas Mc Avoy. Intelligent "control" applications in the process industries. *Annual Reviews in Control*, 26(1):75–86, 2002.

[121] Dražen Slišković, Ratko Grbić, and Željko Hocenski. Multivariate statistical process monitoring. *Tehnicki Vjesnik-Technical Gazette*, 19(1):33–41, 2012.

[122] Jinane Harmouche, Claude Delpha, and Demba Diallo. Incipient fault detection and diagnosis based on kullback–leibler divergence using principal component analysis: Part i. *Signal processing*, 94:278–287, 2014.

[123] Raymond B Cattell. The scree test for the number of factors. *Multivariate behavioral research*, 1(2):245–276, 1966.

[124] JE Jackson and A Edward. User's guide to principal components. john willey sons. *Inc., New York*, page 40, 1991.

[125] Giancarlo Diana and Chiara Tommasi. Cross-validation methods in principal component analysis: a comparison. *Statistical Methods and Applications*, 11(1):71–82, 2002.

[126] Louis Guttman. Some necessary conditions for common-factor analysis. *Psychometrika*, 19(2):149–161, 1954.

[127] John L Horn. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185, 1965.

[128] Scott B Franklin, David J Gibson, Philip A Robertson, John T Pohlmann, and James S Fralish. Parallel analysis: a method for determining significant principal components. *Journal of Vegetation Science*, 6(1):99–106, 1995.

[129] David Zumoffen and Marta Basualdo. From large chemical plant data to fault diagnosis integrated to decentralized fault-tolerant control: pulp mill process application. *Industrial & Engineering Chemistry Research*, 47(4):1201–1220, 2008.

[130] S Joe Qin. Statistical process monitoring: basics and beyond. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 17(8-9):480–502, 2003.

[131] Tao He, Wei-Rong Xie, Qing-Hua Wu, and Tie-Lin Shi. Process fault detection and diagnosis based on principal component analysis. In *2006 International Conference on Machine Learning and Cybernetics*, pages 3551–3556. IEEE, 2006.

[132] Haiqing Wang, Zhihuan Song, and Ping Li. Fault detection behavior and performance analysis of principal component analysis based process monitoring methods. *Industrial & Engineering Chemistry Research*, 41(10):2455–2464, 2002.

[133] Fred Gardiol. *Electromagnétisme*, volume 3. PPUR presses polytechniques, 2002.

[134] Cynthia Furse. Reflectometry for structural health monitoring. In *New Developments in Sensing Technology for Structural Health Monitoring*, pages 159–185. Springer, 2011.

[135] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58, 2009.

[136] Luca Incarbone and Miguel GALLEGO ROMAN. Method for detecting soft faults in a cable, which method is based on the integral of a reflectogram, June 23 2020. US Patent 10,690,712.

[137] A Laib, M Melit, B Nekhoul, K El Khamlichi Drissi, and K Kerroum. Soft fault identification in electrical network using time domain reflectometry and neural network. In *International Conference on Electrical Engineering and Control Applications*, pages 365–376. Springer, 2017.

[138] MK Smail, Y Sellami, HREH Bouchekara, and A Boubezoul. Wiring diagnosis using time domain reflectometry and random forest. In *2019 22nd International Conference on the Computation of Electromagnetic Fields (COMPUMAG)*, pages 1–4. IEEE, 2019.

A - Appendix Chapter 3

A.1 . Long Cable Analysis

Table A.1 represents the Q value for each detected fault case in section 3.4.2.3.

Table A.1 – Q values of new measurement samples

Fault label	Sample number	Q value
$F1$	723	0.014
$F2$	4965	0.265
$F3$	9110	0.346
$F4$	13306	0.053
$F5$	17551	0.978
$F6$	21694	1.291
$F7$	25894	0.163
$F8$	30139	2.894
$F9$	34282	3.861
$F11$	44213	0.088
$F12$	48283	0.117
$F13$	56798	0.331
$F14$	60863	0.44
$F15$	65066	0.057
$F16$	69387	1.021
$F17$	72037	1.181
$F18$	73451	0.698
$F20$	82963	0.044
$F21$	86984	0.058
$F23$	95548	0.165
$F24$	99561	0.22
$F25$	103765	0.029
$F26$	108137	0.52
$F27$	112148	0.694
$F30$	125684	0.032
$F33$	138259	0.114
$F35$	142480	0.019
$F36$	150862	0.38

A.2 . Performance of the selected frequency f in the presence of noise

Comparison and Analysis For the cases in section 3.6.3.2 where the selected frequency in the noisy environment is different from that in the noise-free environment $f_n \neq f$, the data at frequency f_n is considered. First, the noise will be added like before, where 500 realizations are obtained. Then, using the fusion algorithm above, the probability of detection P_{nd} for each fault case is calculated. Table A.2, Table A.3 and Table A.4 compare the probabilities of detection, P_d calculated at the frequency f, and P_{nd} calculated at f_n for the three fault cases $F25$, $F30$ and $F35$, respectively. It is shown that the probability of detection calculated at f is better than that calculated at f_n, as the fault severity become very low, the corresponding fault peak will be of very low amplitude and this increase as the fault position is at the cable end, the noise can mask the fault leading to a low probability of detections.

Table A.2 – Probability of detection comparison for fault case $F25$

	$F25$	
SNR [dB]	P_{nd}	P_d
0	0.356	0.4977

Table A.3 – Probability of detection comparison for fault case $F30$

	$F30$	
SNR [dB]	P_{nd}	P_d
0	0.248	0.2719
5	0.261	0.2856
10	0.339	0.4781

Table A.4 – Probability of detection comparison for fault case $F35$

	$F35$	
SNR [dB]	P_{nd}	P_d
0	0.0301	0.0306

B - Résumé en Français

B.1 . Contexte et Motivations

Les câbles électriques sont utilisés dans tous les secteurs industriels pour transférer de l'énergie ou de l'information. Au cours des dernières décennies, il y a eu une électrification croissante de plusieurs fonctions dans les systèmes industriels. Cette augmentation de l'utilisation de composants électriques s'est accompagnée d'une augmentation de la demande de câbles électriques, par exemple dans l'industrie du transport, et les systèmes industriels. Ainsi, la longueur et la complexité des réseaux filaires ont également augmenté. Par exemple, la longueur cumulée des câbles électriques dans une voiture moderne, est de plus de 4 km comparée à quelques centaines de mètres il y a 30 ans. Elle est supérieure à 40 km dans les avions de combat modernes, est près de 200 km dans les trains à grande vitesse, et 400 km dans les avions civils récents. Dans un pays avec plus de 40000 km de voies ferrées, près de 1 million km de câbles électriques sont utilisés pour les infrastructures. Par conséquent, les câbles électriques jouent un rôle essentiel dans les réseaux qui deviennent des sous-systèmes fondamentaux dont le fonctionnement est crucial.

Au cours de sa vie, un câble peut montrer à un moment donné des signes de faiblesse dus soit à des causes externes, telles qu'une contamination chimique, des vibrations mécaniques, une pénétration d'humidité, etc., soit à des causes internes telles qu'une fabrication défectueuse ou un échauffement local. Ainsi, des défauts plus ou moins sévères peuvent apparaître, allant d'une simple fissure dans la gaine à une rupture totale du câble. Parmi les défauts les plus fréquents figurent les courts-circuits et les circuits ouverts, généralement appelés défauts francs, caractérisés par la perte de la fonctionnalité du câble. En revanche, toute altération mineure affectant un câble est classée en défaut non-franc, à savoir des défauts d'isolement, des fissures, etc. Ces derniers peuvent être de natures très différente et sont plus difficiles à détecter. À première vue, ils peuvent ne pas engendrer de conséquences importantes pour le système. Cependant, le vieillissement des câbles, les contraintes mécaniques et les environnements hostiles peuvent faire évoluer ces défauts non-francs en défauts francs.

Dans le cas des câbles électriques dédiés aux opérations de sécurité et de contrôle, toute baisse de leurs performances due à un défaut peut être extrêmement tragique d'un point su vue humain, écologique, financier, etc. L'enquête du National Transportation Safety Board (NTSB) a révélé que la catastrophe du Boeing 747 TWA Flight 800 en 1996 et la catastrophe du vol Swissair MD-11 en 1998 qui ont fait des centaines de morts ont été, toutes deux, causées par des défauts de câblage électrique. De nombreux autres incidents ont été constatés, qui n'ont pas entraîné d'accidents tragiques, mais ont été attribués à des défaillances de câble comme dans le cas du Boeing 757 de AA (2008) et celui de l'Airbus340 de VA (2009) [16]. D'après les données recueillies par l'AFSA [17], 29% des accidents d'avion sont dus à des défauts de câbles.

Ainsi, pour garantir une utilisation fiable des câbles, il est nécessaire de détecter les défauts

qui pourraient mettre en péril l'ensemble du système. Différentes méthodes ont été développées pour améliorer la fiabilité des réseaux filaires afin de détecter et localiser certains types de défauts dans les câbles. Parmi ces méthodes, on peut distinguer les techniques classiques d'inspection visuelle d'utilisation, de rayons X, de mesures capacitives et inductives [18, 19], et de réflectométrie qui sont largement utilisées et facilement embarquées. Même si plusieurs méthodes de diagnostic des câbles électriques et non électriques ont été étudiées et développées au cours des dernières décennies [20,21], les techniques basées sur la réflectométrie sont toujours au centre de la recherche et des applications industrielles. Leur concept général repose sur la propagation d'une onde électromagnétique dans le réseau filaire à tester, suivie de l'analyse des signaux réfléchis pour détecter la présence, la position et la nature d'une discontinuité d'impédance éventuellement causée par un défaut [21].

Selon le domaine d'analyse de l'onde réfléchie, les méthodes de réflectométrie peuvent être classées en deux grandes familles : la réflectométrie dans le domaine temporel (TDR) [22] et la réflectométrie dans le domaine fréquentiel (FDR) [23]. Certes les méthodes basées sur la réflectométrie ont montré leurs efficacité dans le diagnostic des défauts francs en raison de leurs coefficients de réflexion élevés, mais elles ont montré de moins bonnes performances chaque fois que des défauts non-francs sont traités [24, 25]. En effet, les défauts non-francs qui sont généralement caractérisés par de faibles coefficients de réflexion, produisent des échos faibles par rapport à ceux provoqués par exemple par des jonctions au sein d'un réseau. De plus, l'énergie du signal de test peut être considérablement atténuée en raison de la présence d'inhomogénéité du câble, de jonctions, de couplage, d'épissures, etc., ce qui rend plus complexe la détection de défauts non-francs (c'est-à-dire frottement, rayon de courbure, pincement, etc.). De plus, la détection des défauts non-francs est perturbée par des conditions environnementales telles que les vibrations, les températures élevées, la diaphonie, etc. Dans le cas d'un réseau multi-branche, la détection des défauts non-francs devient plus difficile à cause de la présence de jonctions, etc., qui se traduit par plusieurs échos sur le réflectogramme.

Des développements supplémentaires sont donc nécessaires pour rendre la méthode de réflectométrie suffisamment sensible afin de détecter et localiser efficacement les défauts non-francs. Dans ce contexte, plusieurs méthodes de post-traitement ont été proposées telles que [26–29]. Cependant, ces méthodes sont sensibles aux phénomènes d'atténuation et de dispersion du signal [30]. En effet, les phénomènes d'atténuation et de dispersion réduisent considérablement la précision de la localisation lorsque la distance de propagation est importante [9]. Par conséquent, le choix de la bande passante du signal de test est critique et affecte les performances de diagnostic. De plus, la fréquence maximale du signal de réflectométrie est un paramètre critique pour détecter et localiser un défaut dans un câble efficacement. En effet, plus cette fréquence est élevée, meilleure est la résolution du réflectogramme et la précision de localisation des petits défauts. Cependant, en cas de détection de défaut sur de longs câbles, l'augmentation de la fréquence du signal n'est pas recommandée car elle introduit une dispersion et augmente l'atténuation du signal [31, 32]. En pratique, l'expert configure et calibre l'analyseur du réseau (VNA) à une fréquence donnée et enregistre la mesure dans le cas d'un câble sain. Des mesures à la même fréquence sont alors effectuées sur un câble supposé défectueux. L'analyse des mesures est faite à cette fréquence. Si le défaut n'est pas détecté, cette opération doit être répétée avec une nouvelle fréquence. Par conséquent, il y a une perte d'information et de temps en plus de la subjectivité dans la prise de décision. Par conséquent, une technique plus précise et automatisée doit être proposée.

Les problèmes liés au diagnostic, décrits précédemment, sont aggravés dans le cas d'un réseau multi-branche [33]. Dans de tels réseaux, l'utilisation d'un seul capteur peut ne plus être suffisante pour surveiller l'ensemble du réseau en raison de l'atténuation du signal et de la complexité de la connectique. Pour y remédier, la réflectométrie distribuée est proposée dans la littérature pour lever les ambiguïtés liées à la position du défaut et maximiser la couverture du réseau. Cette méthode consiste à effectuer des mesures de réflectométrie à différentes extrémités du réseau sous test (NUT). Cependant, l'injection de signaux multiples est confrontée à une plus grande complexités de calcul et à des problèmes de fusion de capteurs [11]. De plus, la consommation d'énergie est un inconvénient majeur au regard des contraintes environnementales. L'étude relative à la réduction du nombre de capteurs dans les réseaux complexes et son impact sur la qualité du diagnostic est présentée dans [34]. Cependant, cette étude a révélé d'autres défis liés à l'allocation de la bande passante, au protocole de communication et à l'atténuation des interférences entre les capteurs. Ainsi, dans [35], le profil de durée de vie du câble est inclus (comme les contraintes environnementales (température, vibration, humidité, etc.), le type de câble, l'âge du câble, le bruit du canal et la méthode de diagnostic. Elle permet de réduire le coût du diagnostic en évitant l'utilisation d'un trop grand nombre de capteurs dans le réseau. Avec cette solution, la fiabilité des capteurs en émission et en réception est prise en compte dans les statistiques obtenues. Cette fiabilité diffère d'un capteur à l'autre et impacte la localisation du défaut [34]. En effet, la fiabilité du réflectomètre en injection ou en réception impacte la caractérisation du défaut. Par exemple, dans le cas d'un réflectomètre peu fiable, un signal de test erroné peut être injecté. Ensuite, une mauvaise interprétation du réflectogramme est faite pouvant conduire à une intervention inutile. Par conséquent, les solutions pour résoudre ces problèmes sont des sujets d'actualité dans le domaine du diagnostic filaire.

Cette thèse se propose d'aborder ces deux problèmes principaux dans le diagnostic filaire des défauts non-francs par réflectométrie. L'approche proposée s'appuie sur la combinaison de la réflectométrie avec l'analyse statistique, notamment l'analyse en composantes principales (PCA) pour améliorer la détection des défauts non-francs.
La PCA est une technique de modélisation statistique multivariée basée sur les données [36]. Elle utilise la redondance des informations dans un espace d'entrée corrélé de haute dimension pour projeter l'ensemble des données d'origine dans un sous-espace de dimension inférieure défini par les composantes principales (PC).
Bien que la combinaison de la réflectométrie et de la PCA dans la littérature ait été proposée pour une réduction efficace de l'espace de données de réflectométrie [37], l'application de la PCA sur les données basées sur la réflectométrie n'a pas été utilisée, jusqu'à présent, pour le diagnostic des défauts dans les réseaux filaires. Pour la détection des défauts, le modèle PCA du système est développé dans notre étude à partir des données d'un câble sain, puis utilisé pour tester de nouvelles données obtenues dans le cas d'un défaut.

Ce travail vise à améliorer le diagnostic des défauts non-francs dans les réseaux multi-branches en utilisant cette méthode de réflectométrie basée sur la PCA. Deux problématiques de recherche principales sont abordées. La première est liée à la sélection de la fréquence la plus adaptée à utiliser, et la seconde est liée à la sélection des capteurs dans les réseaux de câbles complexes en utilisant la réflectométrie distribuée. Pour les deux problématiques, les performances et les limites des solutions proposées sont étudiées. L'influence du bruit perturbateur est également analysée. L'approche est évaluée sur des données de simulation

puis validée sur un banc expérimental de laboratoire.

B.2 . Détermination de la Meilleure Fréquence par la PCA pour le Diagnostic des Défauts Non-Francs basé avec la Réflectométrie

Notre objectif est de proposer une méthode de sélection de la meilleure fréquence de signal de test pour la détection des défauts non-francs. Le modèle PCA est établi à partir des réflectogrammes mesurés pour le câble sain à plusieurs fréquences considérées comme les variables. De nouvelles mesures réalisées sur le câble testé aux mêmes fréquences, sont alors projetées dans l'espace défini par le modèle de la PCA. Si une dissemblance est détectée entre les données originales et les nouvelles données, la contribution de chaque variable (c'est-à-dire les fréquences) à cette variation est calculée. L'algorithme permet alors de choisir la fréquence la plus pertinente pour détecter le défaut non-franc. Les avantages de la proposition sont ainsi un gain de temps et une prise de décision automatisée.

Deux méthodes statistiques sont utilisées ici ; les statistiques Q (ou SPE) et le test de Hotelling (T^2) [121]. Si un défaut est détecté, la meilleure fréquence correspondant à celles ayant la plus forte contribution à la détection est alors sélectionnée. La Figure B.1 montre l'organigramme de la méthode proposée. Cette méthode comprend trois phases:

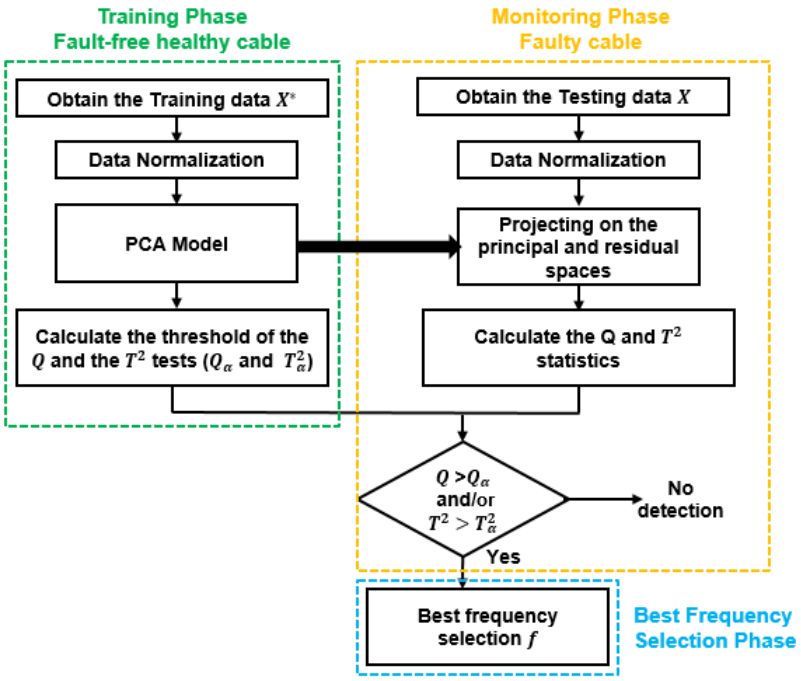


Figure B.1 – Organigramme de l'approche proposée

(1) Phase d'apprentissage, où les données X^* sont collectées pendant un fonctionnement sans défaut, et le modèle de la PCA est développé. Après avoir développé un modèle à

l'aide de données saines (d'entraînement), le modèle à dimension réduite peut être utilisé pour détecter et diagnostiquer des anomalies.

(2) La phase de surveillance, c'est-à-dire la détection du défaut, est gérée à l'aide des statistiques de surveillance T^2 et Q. Pour la détection des défauts, le modèle de la PCA développé basé sur les données du système sain est utilisé pour projeter les nouvelles données de mesure X et comparer la projection obtenue avec celle des données sains. Si la valeur des indicateurs Q ou T^2 est au dessus du seuil de confiance, alors il y a une anomalie.

(3) La phase de sélection de la meilleure fréquence, où le diagnostic des défauts sera géré à travers des tracés de contribution. À partir de la matrice de référence, X^* est construit pour que chaque variable corresponde à une fréquence spécifique ; alors on peut évaluer les variables (les fréquences dans ce cas) qui influencent le plus la détection de la valeur anormale de l'échantillon Q ou T^2 en observant la contribution de chaque variable. Celle ayant la valeur la plus élevée sera retenue comme la fréquence la plus pertinente pour détecter le défaut.

Le processus de simulation des données est divisé en deux parties où le modèle CST est utilisé pour les câbles courts et le modèle $RLCG$ sera utilisé pour les câbles plus longs. Nous simulons des cas de défaut caractérisés par trois paramètres à quatre fréquences différentes $f = [1GHz,\ 2GHz,\ 3GHz$ et $4GHz]$: la position x_f, la longueur $L_f = [5mm,10mm,20mm]$ et la largeur $\theta_f = [45°,90°,180°]$. Habituellement, l'emplacement du défaut est inconnu ; la longueur du défaut est également inconnue. Il est nécessaire d'analyser la sensibilité de la méthode proposée pour différentes valeurs de longueur de défaut pour évaluer sa robustesse.

Pour montrer l'effet de la position du défaut sur la fréquence retenue pour les câbles plus longs, les simulations pour un câble de $100m$ du cas du défaut ($L_f = 5mm$,$\theta_f = 180°$) sont effectuées. La position du défaut varie avec un pas de $2m$. Sur la Figure B.2, on remarque que la fréquence sélectionnée diminue lorsque la position du défaut se rapproche de l'extrémité du câble.

La méthodologie de détection des défauts à l'aide de la PCA et la sélection de la fréquence sont maintenant évaluées sur des câbles réels à l'aide de la configuration représentée sur la Figure B.3. La borne d'entrée du câble est reliée à un connecteur de câble coaxial RF série N 50Ω. L'autre extrémité est laissée libre. Les signaux réfléchis et les réflectogrammes correspondants sont obtenus à l'aide d'un VNA Agilent E5071C, avec calibration pour acquérir des mesures très précises. Un signal TDR est considéré sur une bande passante totale définie du continu à une fréquence maximale f_{max} ($1GHz$, $2GHz$, $3GHz$ et $4GHz$). Les réflectogrammes du câble sain et du câble avec le défaut de blindage situé à $0.5m$ sur un câble à $1m$ sont mesurés. La nouvelle mesure est définie par la matrice X, telle que, chaque vecteur $\mathbf{x_j}$ contient la signature défectueuse du défaut de blindage défini ci-dessus à la fréquence j. La Figure B.4 montre les résultats du test Q de cette expérience où le défaut a été détecté. Le tracé de la contribution de l'échantillon anormal est représenté sur la Figure B.5: on note que la variable correspondant à la fréquence $4GHz$ est celle qui contribue le plus à cet échantillon. Par conséquent, $f = 4GHz$ est la fréquence sélectionnée. Les résultats expérimentaux et de simulation sont cohérents en termes de détection de défaut et de choix de fréquence.

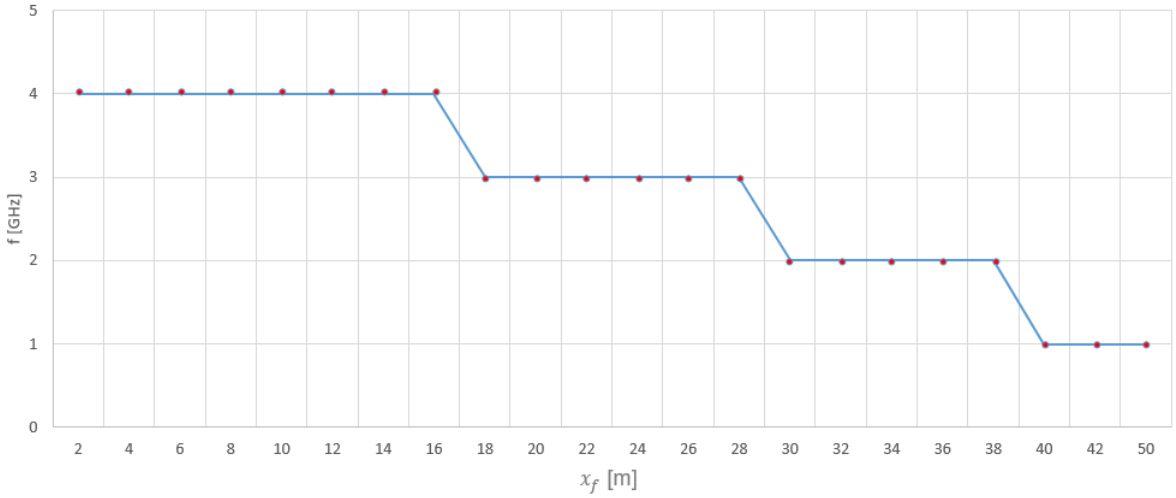


Figure B.2 – Variation de la fréquence sélectionnée avec la position du défaut pour un câble de $100m$ pour le cas de défaut ($L_f = 5mm, \theta_f = 180°$)

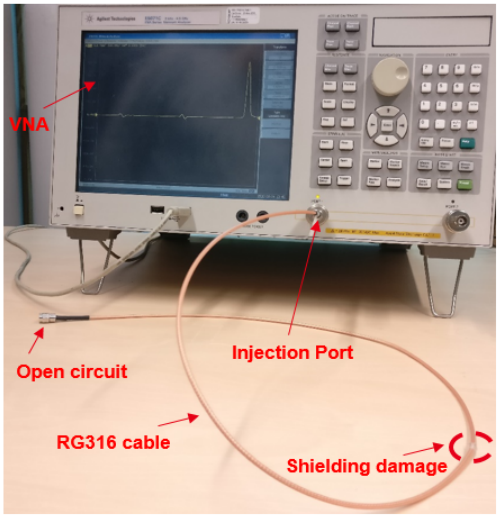


Figure B.3 – Montage expérimental

Les performances du test Q en présence de bruit sont étudiées. La Figure B.6 représente l'organigramme de la méthodologie utilisée. Cette étude est divisée en trois parties; Tout d'abord, la détection des défauts non-francs, où la détection des défauts utilisant la PCA est appliquée et la probabilité de détection du défaut P_d est calculée en présence de bruit. Si $P_d \geqslant \epsilon_d$, le défaut est considéré comme détecté à ce niveau de bruit spécifié avec la valeur du SNR. ϵ_d est un seuil spécifique au domaine. Sa valeur est déterminée par le domaine d'application. Ensuite, la probabilité des fausses alarmes, P_{FA} pour chaque SNR est calculée. La valeur du seuil de fausse alarme ϵ dépend du domaine d'application (idéalement, on souhaiterais avoir $\epsilon = 0$ et $\epsilon_d = 0$). Par conséquent, si $P_{FA} > \epsilon$, les données X_ν doivent subir une étape de pré-traitement; sinon, nous procédons à la sélection de la meilleure fréquence dans l'environnement bruité.

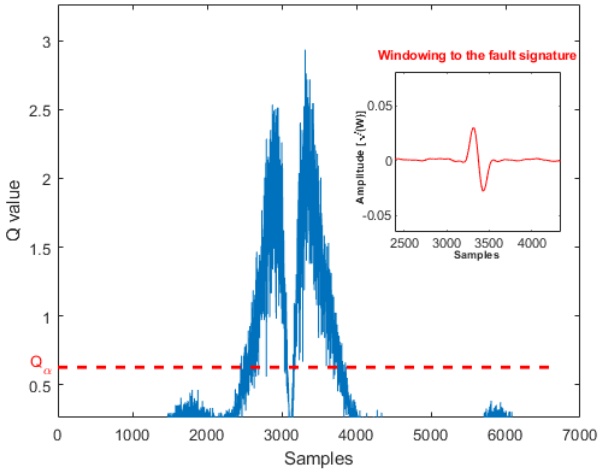


Figure B.4 – Valeurs Q des échantillons expérimentaux

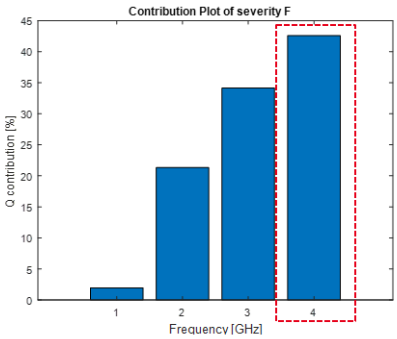


Figure B.5 – Contribution des variables à l'échantillon anormal dans les données expérimentales correspondant à un défaut de blindage de $5\ mm$ de long, $180°$ de large

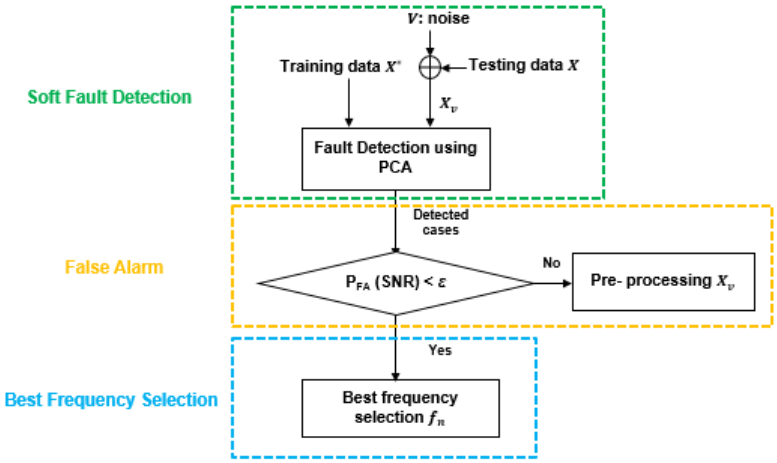


Figure B.6 – Organigramme de l'analyse des performances du test Q

Comme cas d'étude, soit le défaut F correspondant au cas de défaut avec les paramètres suivants $L_f = 5mm$ et $\theta_f = 180°$. On considère un câble de $10m$ pour lequel la position du défaut varie on considère les cas ou elle est à 20%, 50%, 70% ou à 90% par rapport à la longueur du câble.

Le tableau B.1 indique la probabilité de détection P_d pour les différents SNR. On montre que le défaut F, en présence de bruit, est détecté avec une $P_d \geqslant 0.85$. Ici, ϵ_d n'est pas pris en compte car il dépend du domaine d'application. Par hypothèse, nous supposons que les probabilités de détection obtenues sont suffisantes pour les prochaines étapes d'analyse des performances.

Table B.1 – P_d pour différents niveaux de bruit

SNR [dB]	$x_f = 20\%$	$x_f = 50\%$	$x_f = 70\%$	$x_f = 90\%$
-5	1	1	0.853	0.859
0	1	1	1	0.862
5	1	1	1	0.896
10	1	1	1	0.983
15	1	1	1	1

Maintenant, la valeur de P_{FA} pour les différents SNR est indiquée dans le tableau B.2. Elle est de 0.034 pour un SNR de $-5dB$ et de 0.018 à $0dB$ et 0 pour les autres valeurs de

Table B.2 – P_{FA} pour différents niveaux de bruit

SNR [dB]	P_{FA}
-5	0.034
0	0.018
5	0
10	0
15	0

SNR. Ici, le seuil ϵ n'est pas pris en compte car il dépend du domaine d'application. Comme hypothèse, nous supposons que les probabilités de fausse alarme obtenues sont suffisantes pour les prochaines étapes d'analyse des performances.

Pour l'étude de cas ci-dessus, la Figure B.7 représente la variation de la fréquence sélectionnée avec le SNR pour différentes positions du défaut. Il est à noter que la fréquence choisie pour surveiller F dépend de sa position et du SNR. Les deux positions du défaut, $x_f = 20\%$ et $x_f = 50\%$ ont exactement la même variation. Lorsque le défaut atteint l'extrémité du câble, la fréquence sélectionnée diminue. Ce changement est prévisible en raison de l'atténuation tout au long du câble et est conforme à la règle sur des câbles longs: plus la fréquence est basse, meilleure est la détection de défaut. Maintenant, pour un autre étude de cas, en considérant un câble de $100m$, la position du défaut F varie le long de la longueur du câble de telle sorte que, $x_f = 10\%$, 20%, 30% et 40%. Nous ne prenons que la moitié de la longueur du câble ($50m$); pour l'autre moitié, les mesures pourraient être faites à partir de l'autre extrémité du câble à partir d'un autre capteur, rendant la méthode proposée plus efficace. La Figure B.8

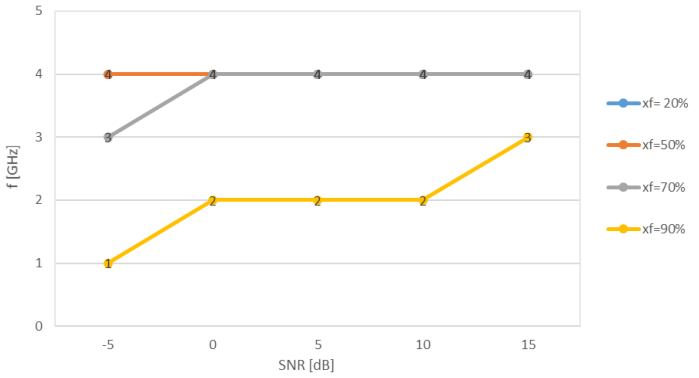


Figure B.7 – Variation de fréquence avec le SNR pour un câble de $10m$ d'un défaut de $5mm$ et $180°$

représente la variation de la fréquence sélectionnée avec le SNR pour différentes positions de défaut. A chaque SNR, pour chaque réalisation, les valeurs Q sont divisées en deux classes :

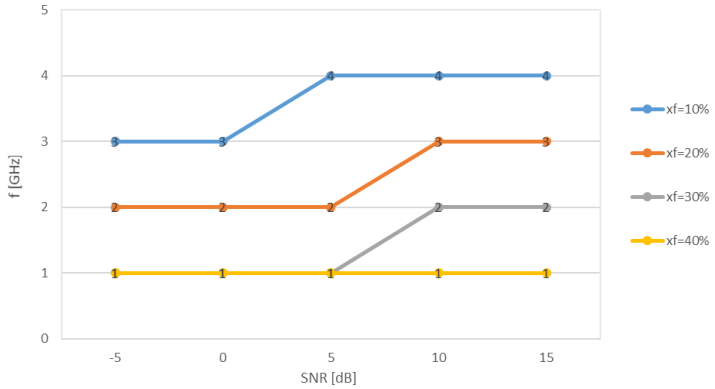
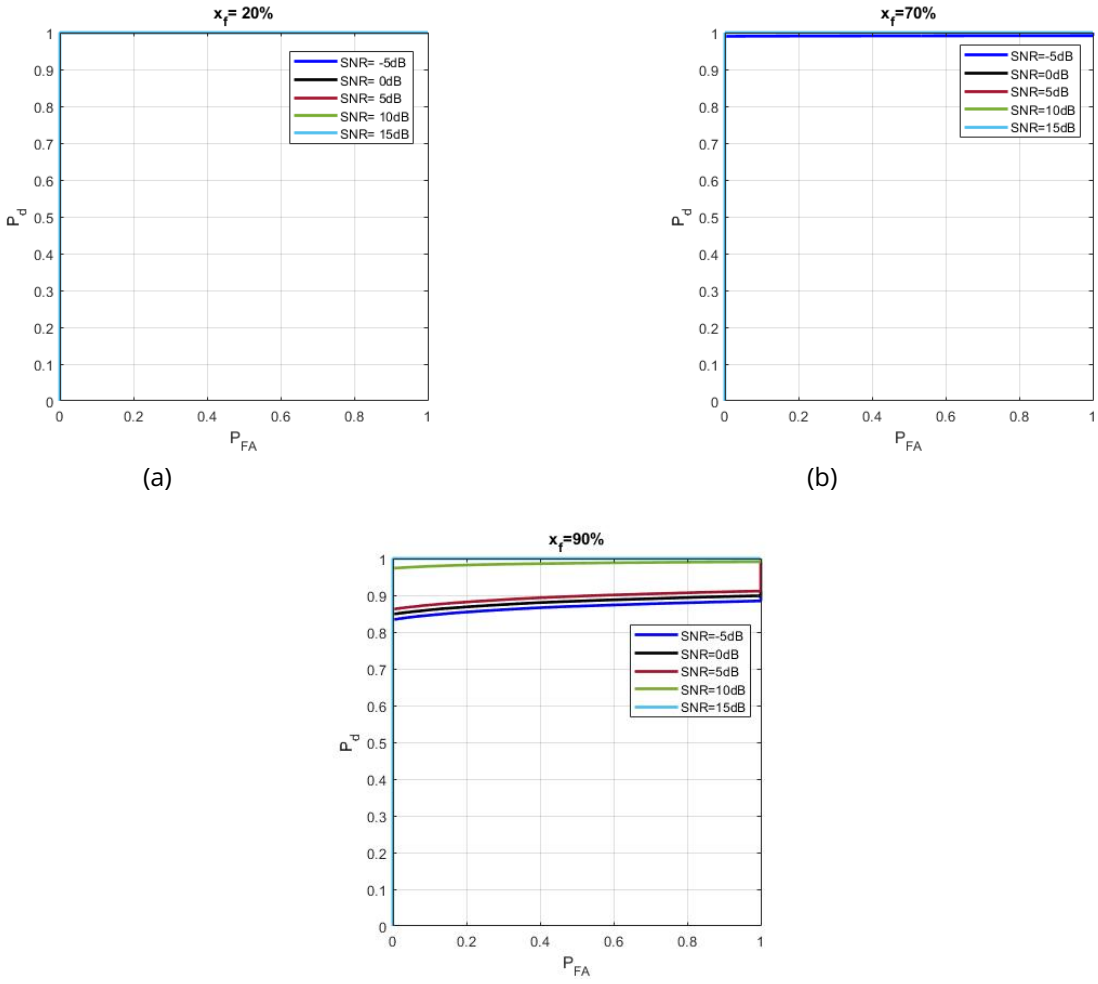


Figure B.8 – Variation de fréquence avec le SNR pour un câble de $100m$ d'un défaut de $5mm$ et $180°$

Saine et Défectueuse. Ensuite pour les 500 réalisations, les données sont concaténées dans les classes Saine et Défectueuse et les courbes ROC sont calculées. L'étude comparative se fait d'abord en réglant le SNR à $15dB$ et en faisant varier la position du défaut F ($5mm$,$180°$) pour un câble de $10m$. Les résultats des performances de détection utilisant le test Q pour différents niveaux de SNR sont représentés sur la Figure B.9, pour $x_f = 20\%$, $x_f = 70\%$ et $x_f = 90\%$, respectivement.

Grâce aux résultats de simulation d'un câble coaxial RG316 de $10m$ de long avec un blindage endommagé, étudié à quatre positions différentes x_f =[20%, 50%, 70%, 90%]. Il est noté que :

- Le test Q a une excellente efficacité avec une capacité de détection de 100% pour les niveaux de bruit (SNR $\geq 0dB$) avec une faible probabilité de fausse alarme pour $x_f = 20\%$.



(a) (b)

(c)

Figure B.9 – Courbes ROC pour un câble de $10m$ d'un défaut de $5mm$ et $180°$ pour (a) $x_f = 20\%$, (b) $x_f = 70\%$ et (c) $x_f = 90\%$

- Pour les niveaux de bruit inférieurs (SNR $< 0dB$), les performances de détection du test Q sont affectées par le bruit pour $x_f = 70\%$ et ce résultat est de 100% pour $x_f = 20\%$.

- Le test Q a une efficacité élevée avec une capacité de détection de 85% avec une faible probabilité de fausse alarme pour $x_f = 90\%$. Les performances de détection des défauts diminuent avec l'augmentation du niveau de bruit passant de 100% à $15dB$ à 85.9% à $-5dB$.

Le test Q est alors efficace pour détecter les défauts, mais ses performances sont affectées par le SNR lorsque la position du défaut est proche de l'extrémité du câble. Au fur et à mesure que l'analyse des performances de cette méthode est étudiée, elle révèle d'excellentes performances pour la détection des défauts non-francs. La capacité de détection est égale à 89.6% même lorsque SNR$= 5dB$. À mesure que la position du défaut approche de l'extrémité du câble, les performances restent toujours bonnes et sont égales à 85.9% à SNR$= -5dB$.

B.3 . Sélection de Capteurs pour le Diagnostic des Défauts Non-Francs basé sur la Réflectométrie Distribuée à l'aide de la PCA

Cette section présente la nouvelle approche de sélection des capteurs pertinents pour surveiller et diagnostiquer les défauts non-francs dans les réseaux câblés complexes. Cette approche combine la réflectométrie distribuée TDR avec la PCA. En effet, pour un NUT donné, une approche de réflectométrie distribuée est envisagée où les capteurs effectuent leurs mesures de réflectogramme. Ces données collectées sont utilisées pour établir un modèle PCA couplé à des tests statistiques pour évaluer l'état des nouvelles mesures. Chaque fois qu'un défaut est détecté, les capteurs pertinents pour la surveillance et le diagnostic sont identifiés avec une grande précision. Sur la base de ces résultats, le nombre de capteurs pourrait être réduit et ceux non sélectionnés pourraient être désactivés temporairement, réduisant ainsi la consommation d'énergie, la charge de calcul et les problèmes de fusion des capteurs. Les performances et les limites de la technique proposée sont étudiées en environnement bruité et validées expérimentalement.

La Figure B.10 décrit le principe de la nouvelle approche combinant les mesures de réflectométrie distribuée TDR avec la PCA. Elle est composée de trois phases principales: l'apprentissage, la surveillance et la sélection des capteurs. Premièrement, les données sont collectées lorsque le réseau est considéré comme sans défaut (étape 1). Dans ce cas, chaque capteur du réseau distribué collecte ses mesures TDR qui sont utilisées pour créer la base de données (étape 2). Ensuite, une matrice X^* est construite de telle sorte que chacune de ses colonnes (représentant les variables) correspond à une réponse TDR du capteur (étape 3). Un modèle PCA est développé sur la base de cette base de données (étape 4). Ce modèle est utilisé dans la deuxième étape pour évaluer de nouvelles données mesurées afin de détecter une anomalie [121, 122].
Deuxièmement, pour le NUT, X, la nouvelle matrice de données mesurées est construite de la même manière que la matrice de données de référence X^*. Elle est ensuite projetée dans le référentiel PCA obtenu lors de la phase d'apprentissage. Les données projetées (appelées scores) T_{new} et le $\tilde{T_{new}}$ résiduel sont alors calculées (étapes 5, 6 et 7).
Les statistiques de surveillance sont ensuite utilisées pour la détection des défauts, c'est-à-dire

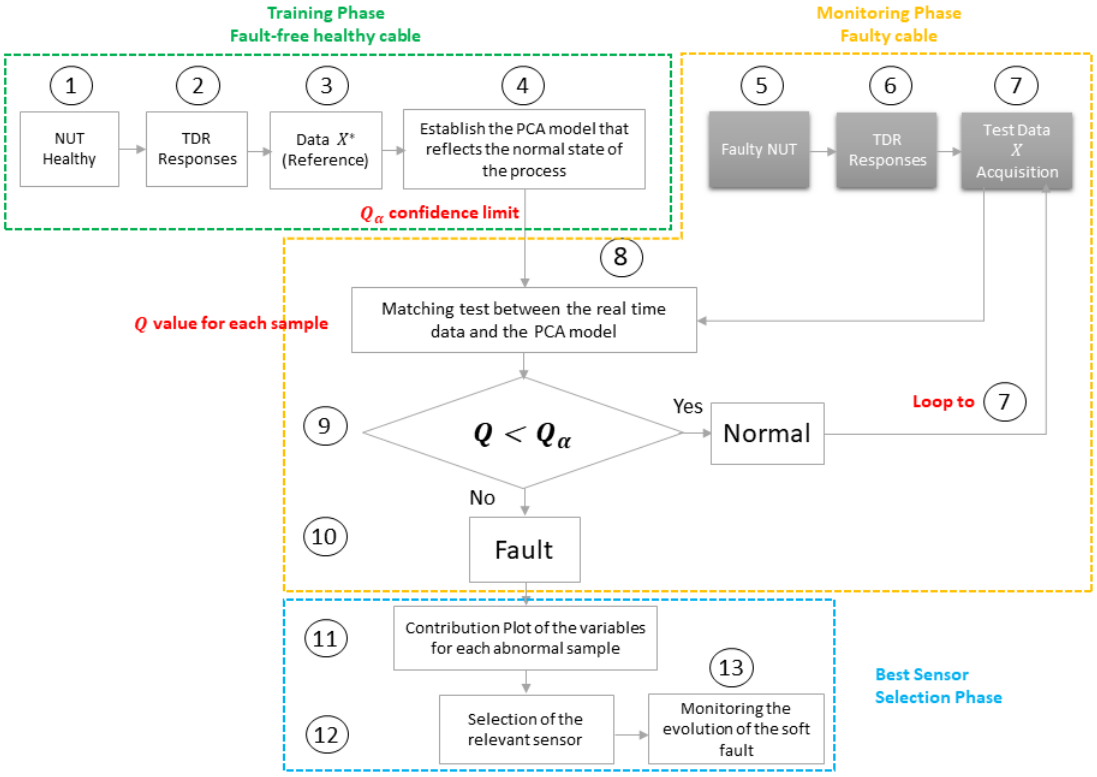


Figure B.10 – Méthodologie de l'approche par réflectométrie distribuée TDR basée sur la PCA

pour déterminer si un défaut s'est produit ou non. Pour cela, le test statistique Q est utilisé [121]. Ce test permet d'évaluer la présence du défaut (étape 8). Le seuil de confiance Q_α pour ce test est calculé en utilisant les données de référence X^* qui sont utilisées pour construire le modèle PCA dans l'étape d'apprentissage. Si la valeur de Q est en dehors de le seuil de confiance pour un échantillon spécifique, alors un défaut est détecté (étapes 9 et 10). Enfin, chaque fois qu'un défaut est détecté, on passe à l'étape d'analyse de sélection des capteurs. Pour chaque échantillon défectueux détecté, l'analyse commence par tracer la contribution des variables constituant la nouvelle matrice de données mesurées X (c'est-à-dire les réponses TDR des capteurs) (étape 11). Ensuite, nous pouvons inspecter les variables qui influencent fortement cette valeur statistique de l'échantillon. On peut donc choisir le capteur le plus pertinent pour suivre l'évolution de ce défaut et désactiver temporairement tous les autres capteurs (étapes 12 et 13).

Pour valider la méthodologie proposée, nous considérons le diagnostic des défauts non-francs dans un réseau point à point, un réseau en forme de Y et un bus CAN (Controller Area Network). Ces configurations présentent des topologies de complexité croissante en termes de branches, jonctions, échos, atténuation du signal, etc. La Figure B.11 représente la configuration de bus CAN considérée. Ce réseau est composé de plusieurs sections, soit B_1 à B_7. Leurs longueurs sont $B_1 = B_2 = B_5 = 2.5m$, $B_3 = B_6 = 5m$ et $B_4 = B_7 = 10m$. Six câbles de $1.5m$, notés $B^{'}{}_1$ à $B^{'}{}_6$, sont utilisés pour connecter les unités de contrôle électronique (ECU) au bus

pour accéder au réseau. Le réseau est constitué de six capteurs **Sj**, $j \in \{1, 2, \ldots, 6\}$ avec les mêmes caractéristiques (réseau homogène). Ces capteurs sont considérés comme adaptés avec les câbles réseau où $Z_c = 100\Omega$.

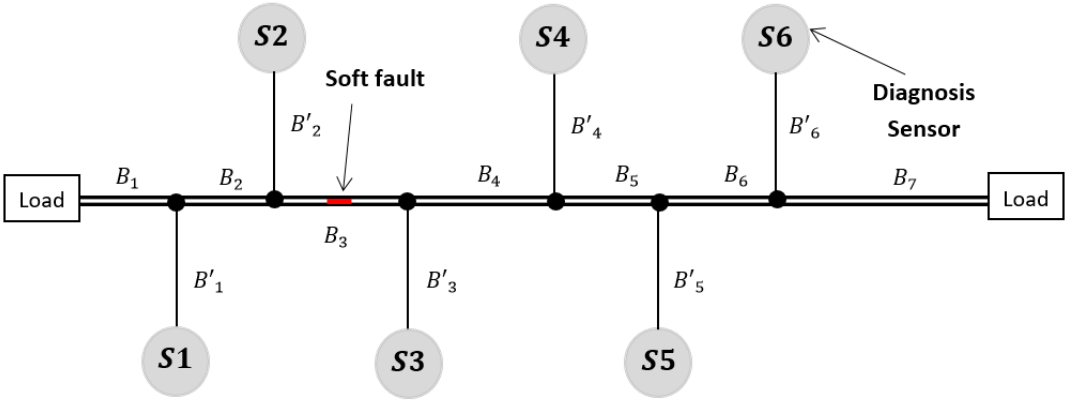


Figure B.11 – Topologie de bus CAN

Tout d'abord, dans des conditions de fonctionnement sans défaut, chacun des six capteurs injecte le signal de test consécutivement. Les réponses TDR sont ensuite obtenues et utilisées pour créer la matrice de données de référence X^* composée de six variables. Chaque variable $\mathbf{x_j^*}$ correspond à la réponse TDR de référence du bus CAN pour le capteur correspondant **Sj** avec $j = \{1, 2, \ldots, 6\}$.
La matrice de données X^* est ensuite utilisée pour obtenir le modèle de la PCA. Une variation locale de 20% de l'impédance caractéristique sur la branche B_3 du réseau est introduite pour simuler un défaut non-franc, soit $\Delta Z_c = 20\%$. Les réponses TDR pour chaque capteur sont collectées. En utilisant toutes ces nouvelles données de réflectométrie, la nouvelle matrice de données mesurées X est créée, où chaque variable $\mathbf{x_j} = \mathbf{R_{Sj}}$ et $\mathbf{R_{Sj}}$ décrit la réponse TDR correspondant au capteur **Sj** avec $j = \{1, 2, \ldots, 6\}$.

Maintenant, la valeur statistique Q pour chaque nouvel échantillon mesuré est calculée. La Figure B.12 représente la carte de contrôle Q, avec la ligne en pointillé rouge représentant le seuil de confiance de 95% (Q_α). On peut observer que certains échantillons sont au-dessus du seuil Q_α indiquant l'occurrence d'un défaut. Le graphique de la Figure B.13 représentant la contribution de l'échantillon défectueux indiqué permet d'identifier le capteur **Sj** qui influence le plus cette valeur Q. Ainsi, la sélection du capteur concerné est effectuée par le test de contrôle Q. Aussi, le tracé de contribution de cet échantillon anormal correspond aux capteurs **S2** et **S3** car ils sont équidistants du défaut.

Nous proposons de valider cette méthodologie sur des câbles réels en utilisant le montage expérimental représenté sur la Figure B.14. Il se compose de plusieurs câbles RG316 avec une impédance caractéristique de 50Ω. Dans cette étude, les longueurs de B_1 à B_7 sont de $0.5m$, $0.5m$, $1m$, $1m$, $1m$, $0.5m$ et $0.5m$, respectivement. Les câbles RG316 qui assurent l'accès au réseau sont notés respectivement B'_1 à B'_6 avec des longueurs $1m$, $0.6m$, $2m$, $1.5m$, $1m$, et $1m$. Les extrémités des lignes B_1 et B_7 sont appariées à l'aide de résistances 50Ω, tandis que les branches B'_1 à B'_6 sont laissées ouvertes. Deux défauts non-francs avec des longueurs

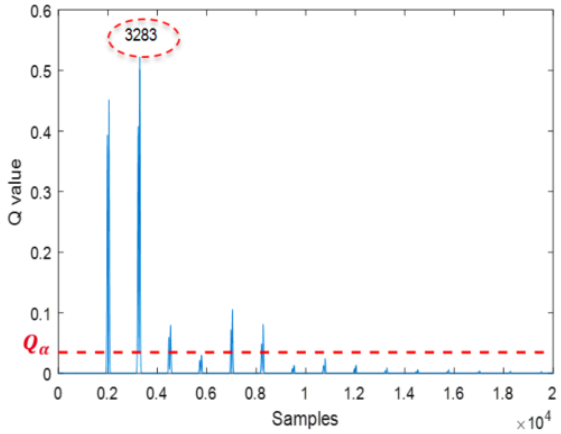


Figure B.12 – Graphique Q des nouveaux échantillons mesurés de la topologie du bus CAN

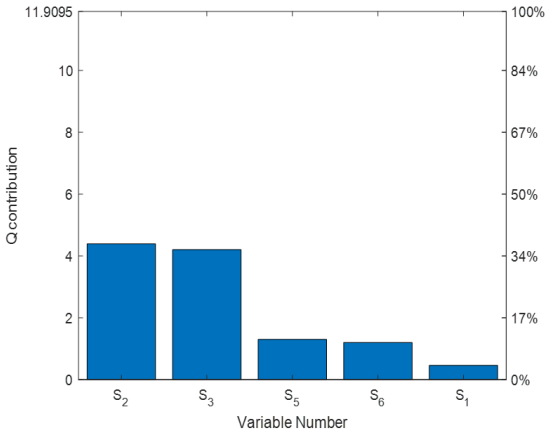


Figure B.13 – Diagramme de contribution pour l'échantillon 3283

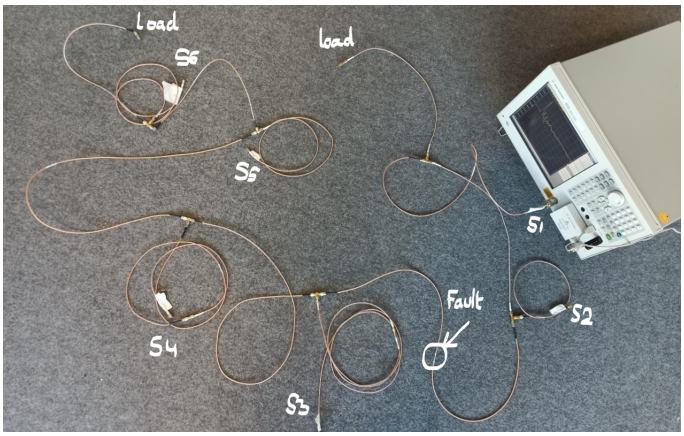


Figure B.14 – Montage Expérimental

de $5mm$ et $10mm$ et de larguer $\theta_f = 180°$ sont créées au milieu de B_3. Les signaux réfléchis et les réflectogrammes correspondants sont obtenus à l'aide d'un VNA Agilent E5071C comme indiqué dans la configuration ci-dessus. La borne d'entrée du câble est reliée à un connecteur

de câble coaxial RF série 50Ω pour chaque mesure. Les réponses de réflectométrie ont été mesurées par l'analyseur de réseau séparément à chaque extrémité du câble pour la bande de $30kHz$ à $4GHz$.

La Figure B.15 représente la valeur du test Q en fonction des échantillons. La ligne en pointillé rouge représentant le seuil de confiance 95% ($Q_\alpha = 4.39$). Nous pouvons observer que certains échantillons sont au-dessus du seuil. Ainsi, l'occurrence d'un défaut a été détectée. Nous considérons l'échantillon avec la valeur Q la plus élevée comme anormal. Par conséquent, dans notre exemple, le numéro d'échantillon 2806 correspond à l'occurrence du défaut. Le

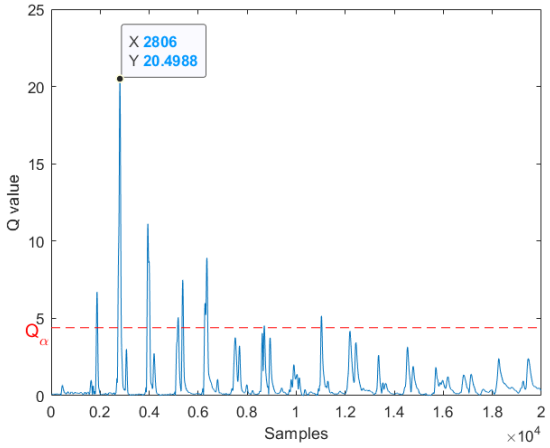


Figure B.15 – Graphique Q des nouveaux échantillons mesurés pour le cas de défaut de $10mm$

tracé du graphique de contribution (la Figure B.16) pour l'échantillon défectueux sélectionné permet d'identifier la variable $\mathbf{R_{Sj}}$ et le capteur **Sj** qui influence le plus cette valeur Q. Ainsi, la sélection du capteur concerné est effectuée par le test Q.

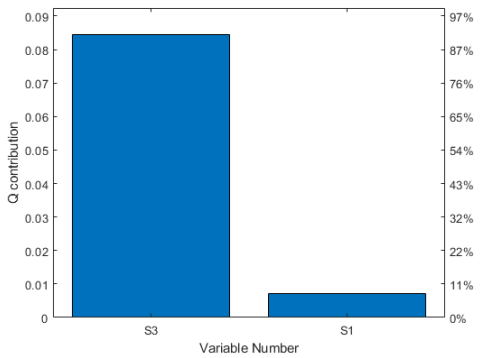


Figure B.16 – Diagramme de contribution pour l'échantillon 2806

Pour notre exemple d'application, le capteur **S3** est sélectionné comme le plus pertinent par le test Q. Il permet de suivre l'évolution des défauts à des fins de pronostic sur la base de ses mesures de réflectométrie. Comme la signature du défaut non-franc dans le réflectogramme **S3** a l'amplitude la plus importante, le choix du capteur **S3** est logique si l'on compare la distance

parcourue ($2.5m$) et le nombre de jonctions existantes (une jonction) entre ce capteur et le défaut. Pour **S2**, bien qu'il représente la plus faible distance parcourue depuis le défaut ($1.1m$), le défaut est masqué par la jonction existant à la même position.

L'organigramme de l'analyse des performances de la sélection du capteur en présence de bruit a été présenté dans la section B.2. Cependant, dans ce cas, le paramètre sélectionné est le meilleur capteur pour surveiller et diagnostiquer les défauts non-francs détectés dans le réseau en environnement bruité.

Pour simuler différents scénarios, deux défauts non-francs sont étudiés. Le premier, F1, est celui introduit ci-dessus, où une variation locale 20% de l'impédance caractéristique sur B_3 du réseau est introduite, c'est-à-dire $\Delta Z_c = 20\%$. La position de ce défaut sera définie comme x_{f1} au milieu de B_3 (à $2.5m$). Le second F2, a la même sévérité que le premier défaut mais à une position différente x_{f2}, située respectivement à $21m$, $18.5m$, $13.5m$, $3.5m$, $2m$ et $7m$ des capteurs correspondants **S1** à **S6**.

Maintenant, les bruits sont ajoutés aux données de test X, résultant en X_ν où les SNR utilisés sont $-5dB$, $0dB$, $5dB$, $10dB$ et $15dB$. $\mathbf{x_j}$ est un vecteur colonne de X pris pour la $j$ième variable. Le bruit est supposé être Additif blanc et Gaussien.

Le capteur **S2** est étudié pour le défaut F1 et le capteur **S5** est étudié pour le défaut F2. Pour chaque réalisation, nous appliquons le test Q pour trouver si le cas en défaut est détecté ou non. Ensuite, la probabilité de détection P_d est calculée pour chaque cas en défaut à chaque SNR, et si $P_d \geqslant \epsilon_d$, le défaut est considéré comme détecté à ce niveau de bruit. Les deux défauts qui ont été détectés sans bruit, sont évalués en présence de bruit avec notre méthode avec une probabilité de détection égale à un.

Le tableau B.3 présente P_d des cas de défaut F1 (milieu de B_3) et F2 (à $2m$ sur B_5) définis ci-dessus pour les différents niveaux de bruit. On montre qu'un même défaut à différentes positions, en présence de bruit, est détecté avec une probabilité de détection égale à un. Par conséquent, le test Q a une excellente efficacité avec une capacité de détection de 100% pour les niveaux de bruit (SNR $\geq -5dB$), et ses performances ne sont pas affectées par le niveau de bruit pour les deux cas de défaut étudiés.

Table B.3 – P_d pour différents niveaux de bruit

SNR [dB]	x_{f1}	x_{f2}
-5	1	1
0	1	1
5	1	1
10	1	1
15	1	1

La probabilité de fausse alarme P_{FA} pour les différents niveaux de bruit est indiquée dans le tableau B.4. Elle est de 0.012 pour un niveau de bruit de -$5dB$ et de 0.0014 pour un niveau de bruit de $0dB$. Elle se réduit à 0 pour les autres valeurs de SNR ($5dB$, $10dB$ et $15dB$). Comme hypothèse, nous supposons que les probabilités de fausse alarme obtenues sont suffisantes pour les prochaines étapes de l'étude d'analyse de performance.

Table B.4 – P_{FA} pour différents niveaux de bruit

SNR [dB]	P_{FA}
-5	0.012
0	0.0014
5	0
10	0
15	0

La Figure B.17 représente la variation du capteur sélectionné avec le niveau de bruit pour différentes positions de défaut. On note que le capteur choisi pour surveiller le défaut défini par les paramètres ($\Delta Z_c = 20\%$, $L_f = 0.05m$) dépend de sa position et du niveau de bruit actuel.

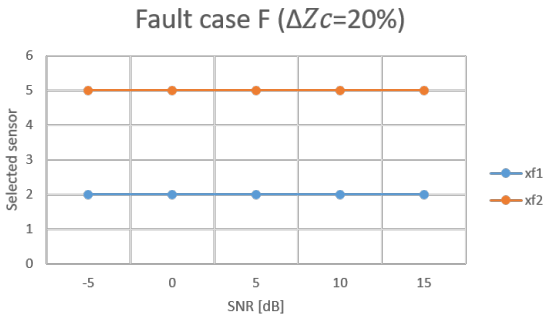
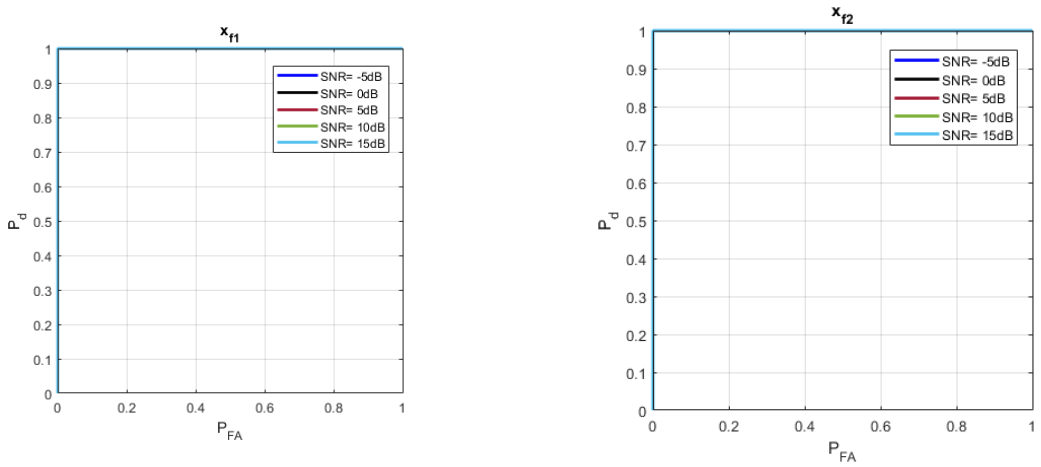


Figure B.17 – Variation du capteur sélectionné avec le niveau de bruit pour deux positions de défaut

On considère le réseau BUS ci-dessus pour le cas de défaut ($\Delta Z_c = 20\%$, $L_f = 0.05m$), aux deux positions x_{f1} et x_{f2}. Le capteur **S2** est étudié pour la position x_{f1}, et le capteur **S5** est étudié pour la position x_{f2}. Les courbes ROC pour les différents niveaux de bruit sont représentées dans la Figure B.18 pour les positions de défaut x_{f1} et x_{f2}, respectivement. Il est à noter que le test Q a une excellente efficacité avec une capacité de détection de 100% pour les niveaux de bruit (SNR $\geq -5dB$) avec de faibles probabilités de fausses alarmes pour les deux positions des défauts. De plus, les performances de détection du test Q ne sont pas affectées par le niveau de bruit. Alors que nous étudions l'analyse des performances de cette méthode, elle révèle d'excellentes performances pour la détection des défauts non-francs. La probabilité de détection est égale à un pour les niveaux de bruit utilisés.

B.4 . Conclusion et Perspectives

Les recherches menées dans cette thèse visent à proposer de nouvelles techniques de diagnostic des réseaux filaires complexes pour détecter et localiser les défauts auxquels les fabricants et les utilisateurs sont aujourd'hui confrontés dans les systèmes électriques. L'importance critique des défauts non-francs a été soulignée et expliquée tout en montrant les limitations des méthodes existantes. Ainsi, de nouveaux outils utilisés pour le diagnostic des défauts sont



(a) x_{f1} (b) x_{f2}

Figure B.18 – Courbes ROC pour un Bus CAN avec un défaut $\Delta Z_c = 20\%$ et $L_f = 0.05m$

nécessaires ; pour cela, le principe de l'Analyse en Composantes Principales (PCA) est présenté grâce à son utilisation harmonieuse avec la réflectométrie, plusieurs obstacles scientifiques son levés, permettant ainsi la détection et caractérisation efficace de défauts non-francs.

La deuxième section présente une approche permettant de sélectionner la meilleure bande passante pour la détection de défauts non-francs dans les réseaux filaires basée sur une combinaison judicieuse de la réflectométrie et de la PCA. La méthode proposée permet de configurer le VNA à différentes fréquences. Des mesures à différentes fréquences pour le cas de référence sont faites pour mettre en place le modèle de la PCA. Ensuite, de nouvelles mesures à différentes fréquences sont réalisées. Si une différence est détectée entre la référence et les données mesurées, la contribution de chaque variable (c'est-à-dire les fréquences) à cette différence est calculée. L'algorithme choisit alors la fréquence la plus pertinente pour surveiller le défaut non-franc. Les avantages sont ainsi un gain de temps et permettent une prise de décision automatisée.
La modélisation tridimensionnelle d'un défaut non-franc et son influence sur la propagation dans un câble a été étudiée avec CST. Les résultats de la simulation ont permis de déterminer la perturbation générée par le défaut non-franc avec différents niveaux de dégradation. Ces perturbations ont été représentées en termes de coefficient de réflexion pour déterminer les caractéristiques du défaut. Pour pallier à l'augmentation de la charge de calcul et du temps de simulation causés par le CST dans le cas de câbles plus longs, les paramètres de défaut sont extraits à l'aide d'une méthode de caractérisation $RLCG$ et sont implémentés dans un code Matlab pour simuler différents scénarios où la longueur l et la position du défaut x_f sont modifiées. Les résultats de simulation de cette méthode sont comparés à la sortie du simulateur CST avec un degré élevé de succès, et le code Matlab a été testé pour sa capacité à récupérer (estimer) la réponse TDR. Des mesures expérimentales ont permis de valider le modèle numérique pour un cas d'endommagement du blindage. Une fois validé, le modèle développé peut donc

être utilisé ultérieurement pour évaluer la méthodologie de diagnostic proposée.
Cette méthode est ensuite étudiée pour un ensemble de longueurs de câble avec différentes conditions de fonctionnement et différents sévérité de défauts. Les résultats de la simulation sont en cohérence avec la règle empirique inhérente aux câbles courts, où plus la fréquence d'excitation est élevée, meilleure est la résolution de détection du défaut. De plus, lorsque le défaut atteint l'extrémité du câble, pour les câbles plus longs, la fréquence sélectionnée diminue en raison d'une atténuation et d'une dispersion plus élevées.
Des validations expérimentales en présence de défauts non-francs ont été réalisées. Les résultats expérimentaux et de simulation sont cohérents en termes de détection de défaut et de choix de fréquence. L'analyse des performances de cette méthode a également été étudiée. Elle a révélé de bonnes performances pour la détection des défauts non-francs. Lorsque le défaut non-franc est proche du point d'injection, la probabilité de détection est égale à 1 même lorsque les valeurs de SNR sont aussi faibles que $0\ dB$. Lorsque la position du défaut approche de l'extrémité du câble, les performances sont toujours bonnes, mais pour des sévérités de défaut plus faibles, la détection est plus délicate. Dans cette étude, il est noté que la fréquence choisie pour surveiller un cas de défaut dépend de plusieurs paramètres : le type et les caractéristiques du câble, la sévérité et la position du défaut, et le niveau de bruit actuel. Concernant les tests statistiques, le critère Q est plus pertinent pour évaluer les défauts.

Dans la troisième section, nous avons proposé une approche efficace pour sélectionner le meilleur capteur pour la détection des défauts non-francs dans les réseaux filaires complexes en combinant la technique TDR distribuée avec la PCA. Avec cette étude, la sélection du capteur est obtenue, quelle que soit la localisation du défaut dans le NUT. Par conséquent, le nombre de capteurs pourrait être réduit et ceux non sélectionnés pourraient être temporairement désactivés, réduisant ainsi la consommation d'énergie, la charge de communication sur le réseaux et les problèmes de fusion de capteurs. Elle conduit également à un gain de temps et d'objectivité dans la prise de décision.
L'efficacité de cette approche a été validée avec plusieurs réseaux de complexités différentes, en commençant par une topologie simple, puis en Y, pour finir par la configuration du bus CAN en présence de différents scénarios de défaut.
Pour une configuration de bus CAN donnée, une approche de réflectométrie distribuée est envisagée dans laquelle les capteurs effectuent leurs mesures de réflectométrie consécutivement. Les réponses TDR sont collectées et organisées dans une base de données. Avec cette base de données, un modèle de la PCA est développé et utilisé pour détecter les défauts non-francs. Couplé à une analyse statistique basée sur l'erreur de prédiction au carré, les capteurs les plus pertinents pour surveiller et diagnostiquer les défauts non-francs dans le réseau sont identifiés. L'approche proposée a été testée lors d'une campagne de validations expérimentales. Les résultats expérimentaux et de simulation sont cohérents en termes de détection de défaut et de choix des capteurs.
Enfin, l'analyse des performances de cette méthode est étudiée. Il révèle d'excellentes performances pour la détection des défauts non-francs malgré la présence de bruit. La probabilité de détection est égale à l'unité avec une faible probabilité de fausse alarme pour les différents niveaux de bruit considérés. Il est à noter que le capteur sélectionné pour surveiller un défaut dépend des paramètres du défaut et du niveau de bruit actuel.

Perspectives: Nous avons choisi de travailler dans cette thèse avec TDR et la PCA pour aborder les problèmes de choix de fréquence et de sélection de capteurs pour le diagnostic des

défauts non-francs dans les réseaux filaires complexes basé sur la réflectométrie. Des résultats significatifs ont été obtenus avec différentes sévérités et caractéristiques de défauts non-francs pour différentes structures de réseau et en présence de bruit. Dans cette optique, une étude approfondie pourrait être menée pour poursuivre ou gérer d'autres fonctionnalités qui n'ont pas été abordées dans ce travail et augmenter la sensibilité à d'autres défauts. Certaines de ces perspectives sont énumérées dans ce qui suit :

Pour cette étude, nous avons considéré un bruit additif blanc et Gaussien; cependant, selon le système et son environnement, plusieurs autres types de bruits plus pertinents dans les applications industrielles peuvent être envisagés, comme les bruits impulsifs dans les réseaux véhiculaires, les bruits mécaniques comme dans les avions, etc. Ceux-ci nécessitent l'établissement de modèles qui doivent être intégrés avec les modèles analytiques de la détection des défauts afin d'évaluer leur effet. Des approches de post-traitement sont alors nécessaires pour atténuer leur effets et permettre une localisation précise du défaut. L'impact des différents types de bruit sur les performances doit être étudié.

L'étude menée dans cette thèse a abouti à une méthodologie de diagnostic des défauts non-francs représentatifs des lignes coaxiales. Le blindage du câble a été analysé. Par la suite, il sera intéressant d'aborder d'autres types de câbles de l'état de l'art utilisés dans différentes applications. Il est également nécessaire d'étendre les études à d'autres défauts en fonction de leur importance. Par exemple, il existe des défauts de câblage qui se produisent couramment dans les aéronefs et sont appelés défauts intermittents tels que les arcs électriques (environ 37%). Ce sont les défauts les plus frustrants, mystérieux et extrêmement difficiles à détecter et à localiser car ils peuvent apparaître en quelques millisecondes à cause des vibrations par exemple, puis disparaître. En avionique, les défauts intermittents sont des défauts de très courte durée (environ 1 microsecondes) qui peuvent apparaître en vol par exemple, mais qui ne sont pas faciles à reproduire lors de la maintenance de l'avion. A titre d'exemple, on peut citer la décharge partielle qui est un phénomène intermittent et dans la plupart des cas invisible. Ce type de défaut n'est plus pris à la légère par les constructeurs, compte tenu des dégâts colossaux qu'ils peuvent occasionner à long terme. Ainsi, il sera intéressant de voir comment nous pouvons appliquer les approches proposées pour ces défauts et améliorer les méthodes de détection proposées.
En revanche, certains défauts de moindre sévérité n'ont pas été détectés par la PCA au chapitre 3. Ainsi, des études complémentaires devraient être menées pour voir comment les données de réflectométrie pourraient être pré-traitées avant d'appliquer la méthode PCA. De plus, le caractère non-linéaire des défauts non-francs à leurs stades antérieurs (défauts naissantes) pourrait être la cause du problème. Dans ce cas, on pourrait enviager d'appliquer d'autres méthodes que la PCA.

L'axe de recherche principal suivant serait le pronostic. En effet, les méthodes de réflectométrie permettent généralement de détecter la présence du défaut après son apparition. Cependant, il serait plus intéressant de prédire l'apparition du défaut. Un autre axe de recherche serait d'utiliser les capteurs sélectionnés dans la section 3 pour suivre l'évolution du défaut et faire un pronostic.

La faisabilité des méthodes de sélection de fréquence et de capteur a été vérifiée par des résultats de simulation et validée par des résultats expérimentaux. Les différentes fonctions de

la méthode ne sont pas encore intégrées dans une carte électronique pour mettre en œuvre un système de diagnostic autonome. L'intégration de tous les algorithmes développés au cours de notre étude repose désormais sur un travail d'ingénierie pour mettre en œuvre un système fiable et optimisé capable de s'adapter aux contraintes des systèmes embarqués.
L'intégration du système de diagnostic permet de d'évaluer le réseau filaire pendant son utilisation par un ou plusieurs systèmes simultanément. Ce diagnostic, appelé "diagnostic en ligne", permet une surveillance continue du réseau. Il est particulièrement utile dans la détection et la localisation de défauts intermittents [9]. Le diagnostic en ligne offre la possibilité d'effectuer un diagnostic parallèlement au fonctionnement du réseau natif. Ainsi, il permet d'être dans les conditions réelles du système pour établir un diagnostic plus approfondi, comme la caractérisation des défauts intermittents.
Il est donc intéressant d'étudier comment les approches proposées peuvent s'intégrer dans un système où d'autres protocoles voire signaux, partagent le même environnement. Il est donc nécessaire de mettre en œuvre une amélioration de la méthode pour prendre en compte les contraintes liées au mode "en ligne", telles que la robustesse au bruit puisque le système de diagnostic peut être soumis à toutes sortes de perturbations provenant soit des signaux utiles du système lui-même ou de sources adjacentes. De plus, il ne doit pas être intrusif dans le fonctionnement natif du système. En effet, les signaux de réflectométrie ne doivent pas interférer avec les signaux utiles présents sur le réseau. Cette interférence peut fausser la réponse du système. Par conséquent, il est nécessaire de mettre en œuvre une amélioration des méthodes proposées dans ce travail pour fonctionner en mode "en ligne".

Un autre axe de recherche serait l'application des approches proposées au diagnostic des défauts dans les réseaux d'énergie. Dans ce cas, les défauts peuvent entraîner de graves pertes économiques et réduisent la fiabilité du système électrique. Ainsi, il est intéressant de voir les résultats des approches proposées pour de tels réseaux.