Soft fault diagnosis in wiring networks using reflectometry and Principal Component Analysis

Nour Taki, Claude Delpha, Demba Diallo, Wafa Ben Hassen, Nicolas Ravot

Version of Record: https://www.sciencedirect.com/science/article/pii/S0263224122006133
Manuscript_8c524fcf1efc5ad62e1e5cf146f0262e

Soft Fault Diagnosis in Wiring Networks Using Reflectometry and Principal Component Analysis

Nour TAKI[a,b], Claude DELPHA[b], Demba DIALLO[c], Wafa BEN HASSEN[a], Nicolas RAVOT[a]

[a] *Commissariat à l'énergie atomique et aux énergies alternatives, Laboratoire Fiabilité et Intégration Capteurs, Paris-Saclay Campus - Nano-INNOV, Bat. 862-PC 172, Gif-sur-Yvette, F-91192, France*
[b] *Université Paris Saclay, CNRS, CentraleSupelec, Laboratoire des Signaux et Systèmes, 3, Rue Joliot Curie, Gif Sur Yvette, F-91192, France*
[c] *Université Paris Saclay, CNRS, CentraleSupelec, Group of Electrical Engineering of Paris, 11, Rue Joliot Curie, Gif Sur Yvette, F-91192, France*

Abstract

During operation, 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. 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 shown to be less effective for soft faults. Indeed, soft faults are characterized by a small impedance variation, resulting in a low amplitude signature in the reflectograms. Accordingly, the detection of these faults depends strongly on the test signal bandwidth. Although the increase of the maximal frequency of the test signal enhances the soft fault's "spatial" resolution, the performance is limited by signal attenuation and dispersion. This study proposes a method to select the best maximal frequency for soft fault detection. It is based on a combination of reflectometry with Principal Component Analysis (PCA), and the analysis of the Squared Prediction Error (SPE). Experimental validation is carried out, and performance analysis in the presence of noise is investigated. The results for shielding damage show that when the soft fault is near the injection point, the detection probability equals to one even for SNR values as low as $0dB$. As the fault position approaches the end of the cable, the performance is still acceptable, but for lower fault severities, the detection is almost impossible. The results also show that the selected frequency depends on the fault severity, the fault position, and the noise level.

Keywords: Wire diagnosis, Soft fault detection, Time domain reflectometry, Principal component analysis, Squared Prediction Error.

1. Introduction

The replacement of conventional pneumatic, hydraulic and mechanical systems in modern systems by electrical ones has resulted in more complex electrical systems as well as kilometers of wires (i.e., up to $530km$ in an Airbus 380), linking electrical systems, computers, displays, avionics, sensors, actuators, and cabin entertainment. As an example, different types of cables are used but coaxial ones are remaining the most common ones for communication, and accurate transmission purpose. The electrical wiring networks are considered as a critical component since it is subjected to aggressive environmental conditions such as vibration, heating, moisture, etc. This can lead to the appearance of faults in wires consisting in hard faults (i.e. open circuit and short circuit) or soft faults (i.e., chafing, pinching,

bending radius, etc.). Soft faults do not interrupt the propagation of the signal or the energy down the cable but they can disturb it. Moreover, soft fault can evolve into hard one leading to system failure.

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 with the fault detection and location in cables [1, 2, 3, 4, 39]. Among these methods, we can distinguish the classical visual inspection, X-ray, capacitive and inductive measurements [5, 6], and reflectometry techniques [39, 7, 8, 9, 3], 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 [10, 4, 41, 42], reflectometry-based techniques are still in the center stage of research and industrial applications. 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.

Depending on the analysis domain of the reflected wave, reflectometry methods can be categorized into two main families: Time Domain Reflectometry (TDR) [11, 40] and Frequency Domain Reflectometry (FDR) [12, 39, 43]. The considerable development and improvement of electronic components during the last decade have simplified the implementation of TDR-based methods. High-resolution TDR provides precision Distance-to-Fault measurements supplemented with useful frequency-domain information. When fast, accurate localization of cable and connector faults are needed, TDR is usually a better choice than FDR. In the case of complex topologies (multi-branched, connectors, splices, junctions) the efficiency of TDR is proved. Besides, over the last decade, a massive improvement in electronic components (FPGA[1] and DSP[2], high bandwidth converters, etc.) has been observed. It has simplified the implementation of TDR based methods. For this reason, methods derived from TDR are generally preferred for cable diagnosis, especially when on-board. Consequently, in this work, we will focus on the TDR methods.

Reflectometry methods have provided effective results with hard faults due to their high reflection coefficients, but they have shown poorer performance whenever soft faults are addressed [14, 13, 39]. Indeed, 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 more complex. Moreover, soft fault detection is disturbed by environmental conditions such as vibration, high temperature, crosstalk, noise disturbance, etc.

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 [15, 16, 17, 18]. However, these methods are prone to test signal attenuation and dispersion phenomena [19]. Indeed, these phenomena significantly reduce the location accuracy when the propagation distance is important [20]. Hence, the choice of

[1]Field-Programmable Gate Array
[2]Digital Signal Processor

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: the higher is the frequency, the better is 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 [21, 22]. In practice, the expert configures and calibrates the Vector Network Analyzer (VNA) at a given frequency and records the healthy cable measurement. Then measurements at the same frequency are then done on the cable under test, and compared to the previous ones. If no fault is detected, this operation should be repeated for several frequencies. 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 and select the best frequency in a given set.

Paper Contribution. With the knowledge that different methods have been used in the literature for fault detection and diagnosis in other domains than wire diagnosis using data-driven approaches [44, 45], our main goal in this paper is to propose an efficient solution for improving soft fault detection based on the benefit of the reflectometry and then cope with the frequency selection problem.

1. We propose a new approach combining reflectometry with statistical analysis using the principal component analysis (PCA) for soft fault detection and then the frequency selection among a given set. PCA is a multivariate data-driven statistical modelling technique [23] that uses information redundancy in a high-dimensional correlated input space to project 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 [24], the application of PCA on reflectometry-based data has not been used, so far, for fault diagnosis in wired networks.
2. This work is an extension of the study done in [25] in which the proposal permits to configure and calibrate the VNA at different frequencies: first, measurements at different frequencies are done for the healthy cable to design the PCA model. New measurements at the same frequencies are projected in the PCA reference frame and compared to the healthy data. If a dissimilarity is detected, 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 advantages of the proposal are thus time-saving and enabling automatic computerized decision-making.

The rest of the paper is organized as follows. In section 2, 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. Section 3 will be devoted to describing the frequency selection method and simulation and experimental results. Section 4 will analyze the fault detection performance in the presence of noise. Section 5 will evaluate the frequency selection robustness to environmental nuisance. The conclusion in section 6 will close the paper.

2. 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. The FAA[3], NAVAIR[4], and NASA[5] have all identified wire chafing as the most significant factor contributing to electrical wiring and interconnect system failures in aging aeronautic systems [26]. Coaxial cables are used for the numerical model to validate the proposed method and in the conducted experimental measurements.

2.1. Coaxial Cable Modeling and Simulation

A 3D EM model of a coaxial cable is developed using CST as shown in Figure 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 1.

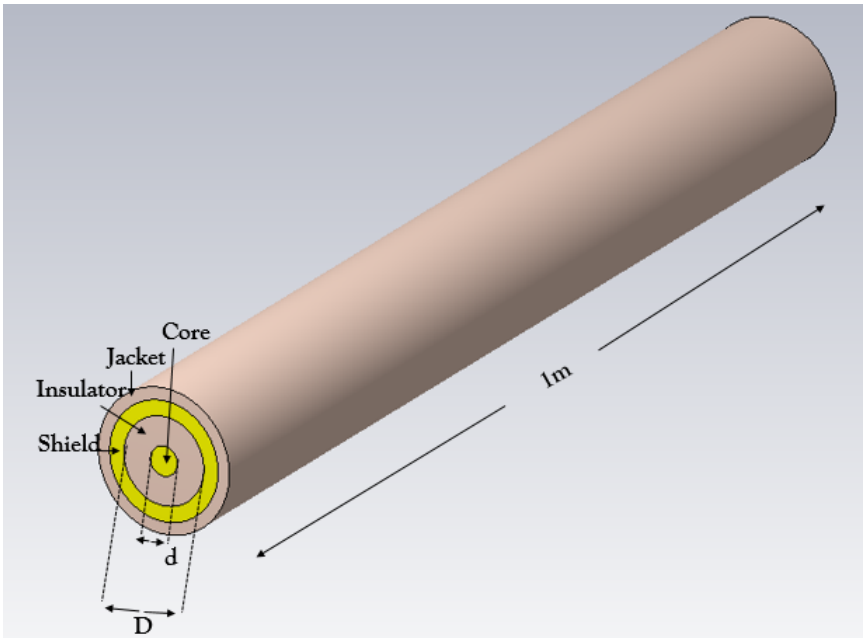


Figure 1: The developed 3D EM model of a shielded coaxial

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

[3]Federal Aviation Agency
[4]Naval Systems Air Command
[5]National Aeronautics and Space Administration

Table 1: RG-316 Coaxial cable construction specifications

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

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

$$Z_c = \frac{138 \times \log_{10}\left(\frac{D}{d}\right)}{\sqrt{\epsilon_r}} \tag{1}$$

A discrete port with an impedance $Z_0 = 50\Omega$ connects the two conductors through the center point to inject a test signal and receive the reflected one. The transmission line is supposed to be open-circuited at its end, and the power supply is a Gaussian excitation source located at x= 0 with a maximum frequency of $4GHz$. For the simulations, a Gaussian pulse is injected into the wire and the reflected signals are recorded.

Figure 2 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). 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. In case of an open circuit, $Z_l = \infty$ and $\Gamma_l = 1$, 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 $0.836V$, compared to the amplitude of the Gaussian pulse $1V$.

Simulation results for longer cables with different fault parameters are inducted. However, in CST, the accuracy of the solution is strongly related to the number of meshes. Consequently, for long cables, the number of meshes is very high, which increases the computational burden and the simulation time. Therefore, at first, we will use the CST modeled coaxial cable in [25], where the R2RLCG method [27] is applied to extract the $RLCG$ parameters. These parameters will then be used for soft fault evaluation using the characterization method in [28]. This method extracts the fault parameters, which are, then, implemented in a Matlab©language code to simulate different fault scenarios where the cable length l and the fault position x_f are changed.

2.2. RLCG-based Model

For the previous cable, the R2RLCG method explained in [27] is applied to extract the $RLCG$ parameters displayed in Figure 3. R2RLCG method [27] 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

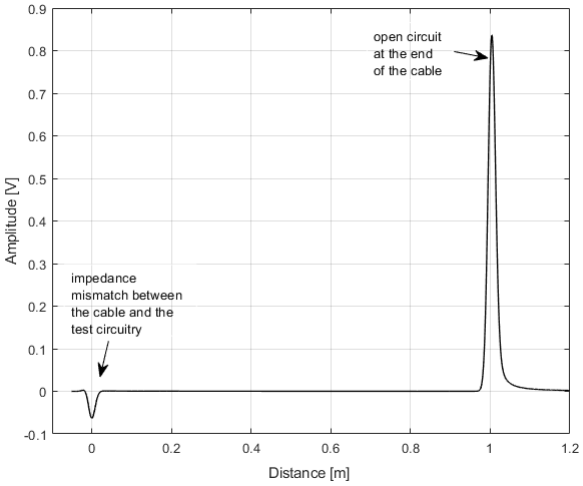


Figure 2: TDR response for the $1m$ modeled coaxial cable with an open circuit at its end

methods described by some articles [35, 36, 37, 38] 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:

- 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 [27]. 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 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 assumed constant over the frequency band.

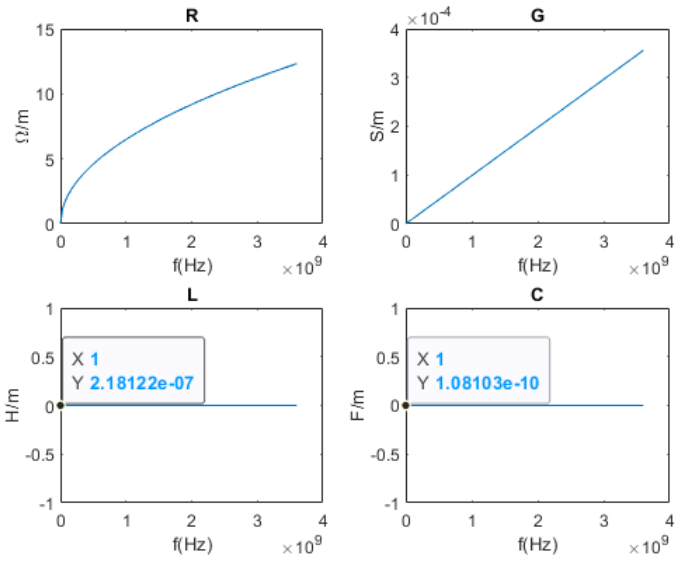


Figure 3: Variation of estimated cable $RLCG$ parameters as a function of the frequency

Figure 4 shows the TDR reflectograms of the $RLCG$-based model and the 3D EM model. For these simulations, the bandwidth ranges from DC to $4GHz$, and the width of the Gaussian pulse is $250ps$. The correlation coefficient used to evaluate the similarity between the two models is equal to 0.9584. It is considered that the data is strongly correlated if the correlation coefficient is between 0.9 and 1 [28]. The model with distributed constants is validated. Therefore, the computed R, L, C, and G parameters of the healthy cable will be used for soft fault characterization in the next section.

2.3. Soft Fault Modeling and Simulation

We will focus on the impact of the degradation of the cable shielding, taking into account the three-dimensional characteristics of the fault. 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 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).

The coaxial cable modeled in section 2.1 is degraded (a portion of the shield is removed) as shown in Figure 5. For this work, only the shield of the cable is damaged. As highlighted in Figure 5, the core is not affected by the cable degradation. The fault is characterized by three parameters: the fault position x_f, the fault length L_f and the angular cutaways θ_f. We will seek to understand the influence of each parameter on the shape of the reflectogram.

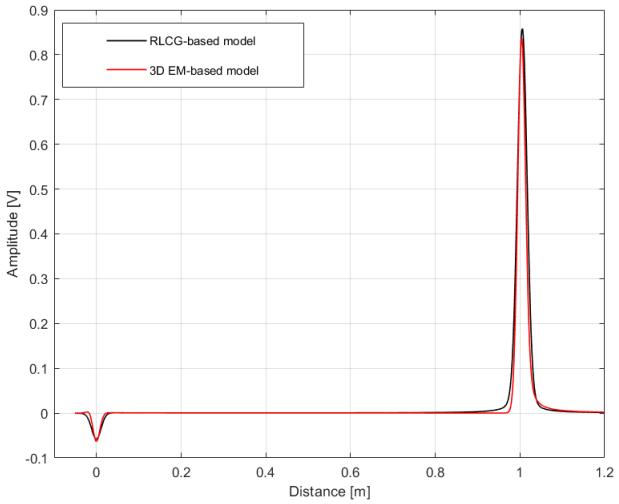


Figure 4: TDR reflectograms of the *RLCG*-based model and 3D EM model

Hereafter, we consider the case for which only the cable's shield is degraded with different widths (θ_f): 45°, 90°,180°. The soft fault of length $L_f = 5mm$ occurs at $x_f = 0.5m$ from the beginning of the cable.

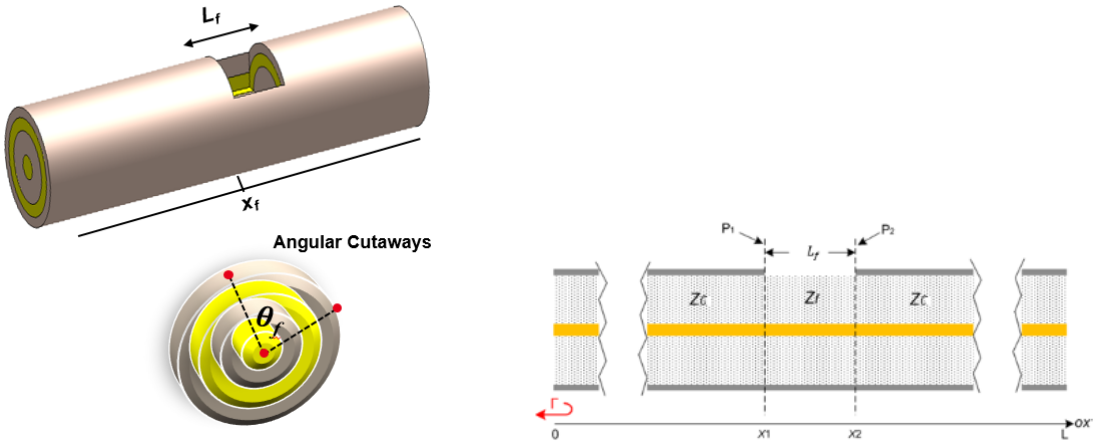


Figure 5: Coaxial cable with a shielding damage model

Figure 6 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, 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 ($0.833V$) due to the reflections in the faulty zone.

Figure 7 illustrates the TDR fault signatures for three angular cutaways. The shape of the reflected signal exhibits positive and negative reflections due to the sign of the reflection

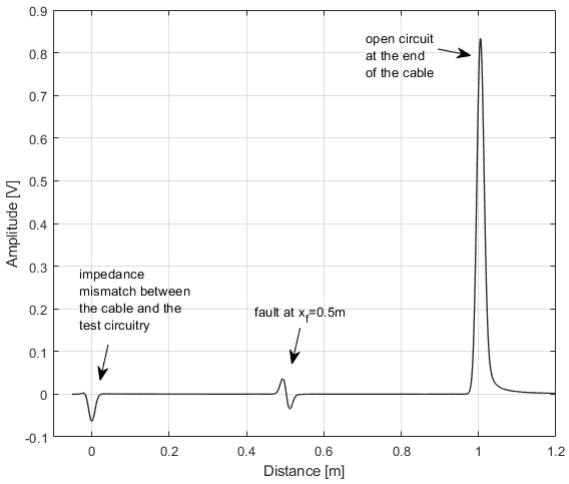


Figure 6: TDR response of a shielding fault with $x_f = 0.5m$, $L_f = 5mm$ and $\theta_f = 180^{\circ}$

coefficient at the discontinuity interfaces. The reflected signal changes significantly with the shielding fault width θ_f. As shown in this figure, the amplitude of the reflected waves is very small (around 5% of the excitation signal amplitude U_i). Thus, the reflected pulse cannot be detected easily with usual measurement devices since according to [29], 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.

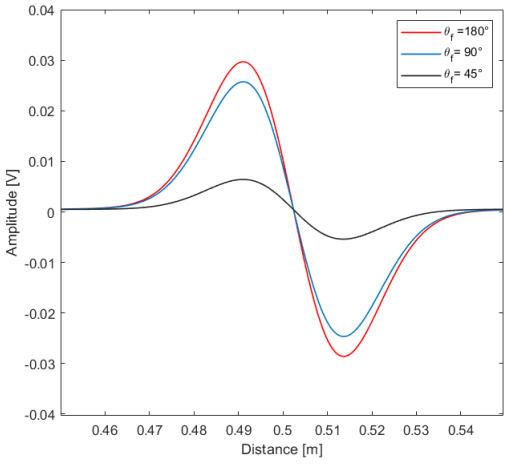


Figure 7: Shielding fault signatures for different fault width's at $x_f = 0.5m$ and $L_f = 5mm$

The detailed process of the characterization method has been represented in [28]. The first step has been done in section 2.2 where the $RLCG$-based model is developed and vali-

dated. The $RLCG$-based model TDR response (Figure 4) is V_{sim} defined as 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 7 is used as V_{mes}, 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. In the case of a shielding damage with $\theta_f = 45°$, $L_f = 5mm$ and $x_f = 0.5m$, the impedance of the fault is $Z_d = 0.9Z_c = 45\Omega$. This fault thus introduces an impedance variation (ΔZ_c) of 5Ω. In the case of a shielding damage with $\theta_f = 90°$, the impedance of the fault is $Z_d = 0.54Z_c = 27\Omega$. This fault thus introduces an impedance variation of 23Ω. Finally, in the case of a shielding damage with $\theta_f = 180°$, the impedance of the fault is $Z_d = 0.46Z_c = 23\Omega$. This fault thus introduces an impedance variation of 27Ω.

From the results in Table 2, we can observe an increase of the variation of the characteristic impedance and the reflection coefficient when the fault width increases.

Table 2: The estimated parameters of the shielding damage for different fault 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 show that the length of the fault has no effect on the change in impedance when the fault width is set constant at 45°. However, the increase of the fault length L_f increases the amplitude of the fault signature.

Table 3: The estimated parameters of the shielding damage for different fault 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 using the Matlab code is compared to the commercial 3D simulator (CST) by simulating a cable with a shielding fault. The TDR results for three different widths are depicted in Figure 8. The correlation coefficients for the different cases are = 0.9841, 0.9839 and 0.9797 respectively. 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 9 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.

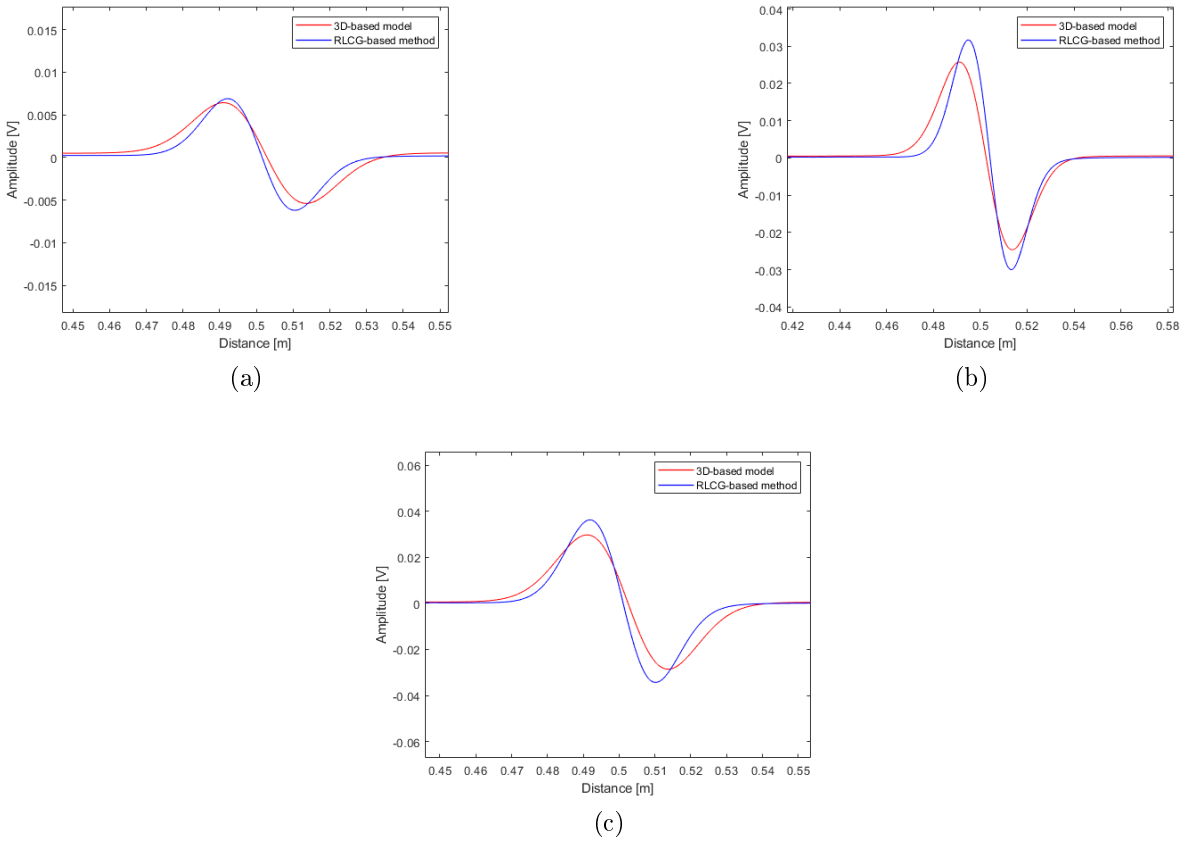


Figure 8: Signatures of the reconstructed and simulated fault with width (a) $\theta_f = 45^\circ$, (b) $\theta_f = 90^\circ$ and (c) $\theta_f = 180^\circ$

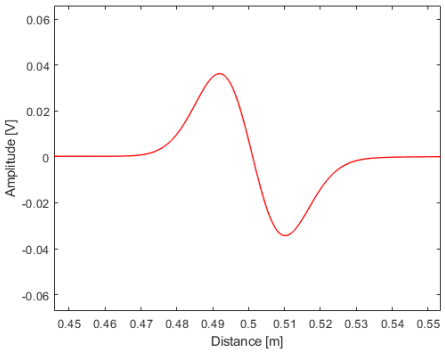
Table 4 shows the amplitude variation of the soft fault signature with the change in the fault position along the cable. It illustrates the attenuation phenomenon and hence, the selection of frequency issue addressed in the following sections.

Table 4: Variation of the fault ($\theta_f = 180^\circ$ and $L_f = 5mm$) amplitude with its position along the cable

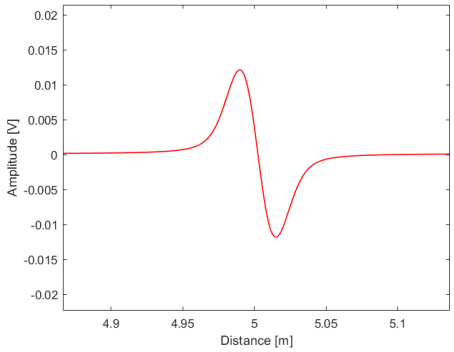
Cable length l_f [m]	1	10	100
Fault position x_f [m]	0.5	5	50
Amplitude [V]	0.036	0.012	0.00012

2.4. Model Validation with Experimental Results

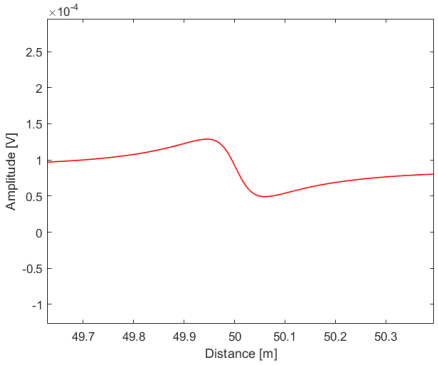
This section aims at validating the developed models with experimental results. The experimental data collection starts with focusing on the coaxial cable type, measuring the TDR response of the healthy cable. Then the cable is damaged using an abrasive apparatus to chafe a small section of the wire and finally measure the TDR response of the faulty cable. In the following, we are interested in how the TDR signal changes depending on the width



(a) $1m$ cable



(b) $10m$ cable



(c) $100m$ cable

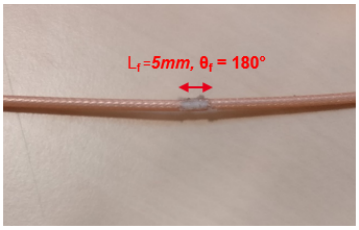
Figure 9: Fault signature for different cable lengths

of the chafing. At $\theta_f = 180°$, the shielding has worn away sufficiently to expose the inner core conductor insulation as it can be observed in Figure 10a.

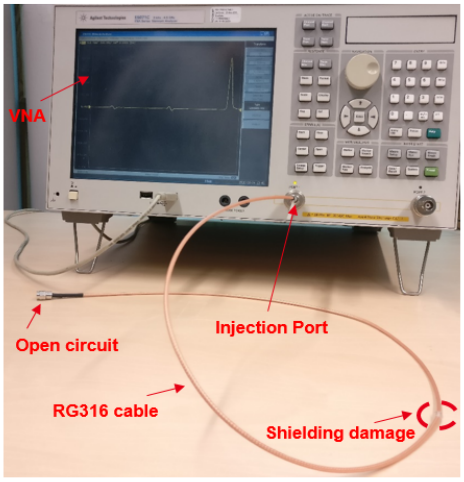
The measurement is carried out using an RG-316 coaxial cable ($Z_c = 50 \mp 2\Omega$ at $4GHz$). The exterior insulation on the cable has a relative permittivity of about 2.1. In this study, we are not interested in measuring scrapes of the exterior insulation, but instead, we are interested when chafing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

A $5mm$ long, 180° wide shielding damage is created at $0.5m$ from the injection point as shown in Figure 10a. The length of the Cable Under Test (CUT) is $1m$. In Figure 10b, the 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 9kHz-4.5GHz Keysight Vector Network Analyzer, calibrated 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 $4GHz$). Note that the VNA is a Keysight E5071C ENA Vector Network Analyzer. This analyzer combines high RF performances and powerful analysis capabilities. It allows many functionalities such as fixture simulator, equation editor for real time data processing, time domain analysis, dielectric and magnetic properties measurement, ... Based on the embedded fixture simula-

tor, this VNA allows mixed mode S-parameters measurements, port impedance conversion, matching circuit simulation, ... The conversion between the S-parameters and the TDR is done automatically by this VNA. For our experiments, the VNA generate a sinusoidal test signal. A variation of the chirp z inverse Fast Fourier Transform (FFT) is used to transform from the frequency domain to the time domain. This permits the user to "zoom in" on a specific time (distance) range of interest for the data display. The reflected signal after post processing (FFT) is converted into time domain (TDR) to obtain the reflectogram. Then we can compare simulation results (TDR) with experimental results that are represented in time domain.



(a) Experimental chafed cable



(b) Experimental bench

Figure 10: Experimental setup

Figure 11 and Figure 12 represent the simulated and experimental reflectograms for healthy and faulty cable, respectively. The difference at the input is due to the presence of the connector and the generator in the experiments, whereas for the simulations, only a generator is used.

Three sections are considered in the experimental and simulation TDR responses: input, faulty part, and the end (here, open circuit). The energy for each section is calculated using (2). the results presented in Table 5 show 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. This difference is probably also due to the larger actual cable loss compared to what is modeled.

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

where $X(f)$ is the Fourier transform of $x(t)$ defined as the reflected signal.

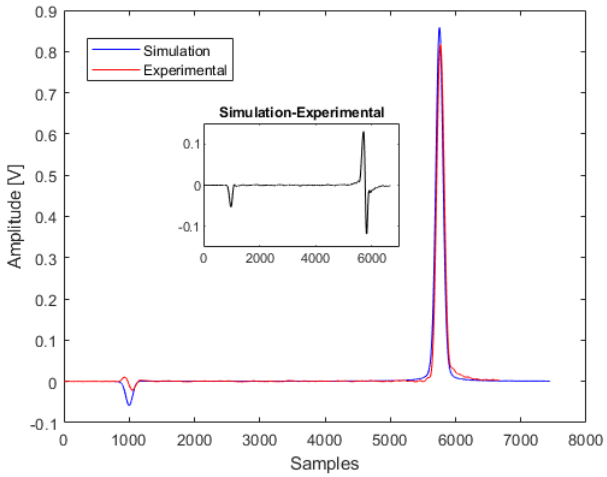


Figure 11: Healthy TDR responses

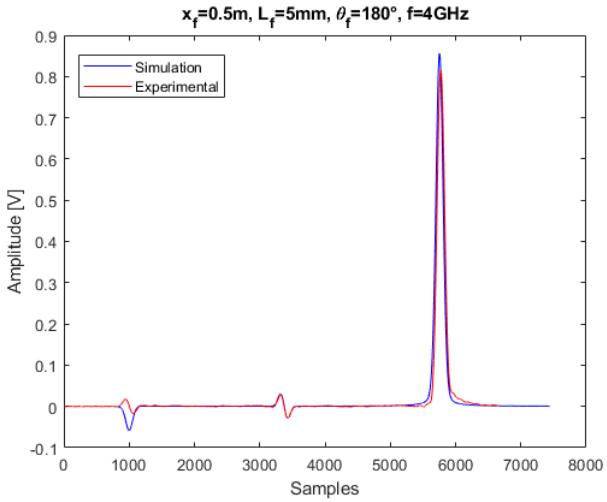


Figure 12: Faulty TDR responses

Table 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

The Root Mean Square Error (RMSE) between the two reflectograms in Figure 12 is equal to 0.024. This value is relatively low since the observations mean value is 0.168. It can be noticed that we can find the mean of observations by dividing the sum of all the observations by the total number of observations (here we use the reflectometry data). The RMSE based only on the signature in the fault section is 0.001. Since it constitutes only 4% of the total RMSE (0.024), the majority of the error is due to the output section. These

results show that the magnitude of the simulation and experimental signals are in the same range.

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

3. Frequency Selection Algorithm for Soft Fault Diagnosis

The first step is to design the PCA model from the fault-free cable. This model is then used to evaluate new measurement data from the cable under test. Two statistical measures can be used; the Q (Squared Prediction Error (SPE)) in the residual subspace, and the Hotelling T^2 in the principal subspace [30]. If a fault is detected, the best frequency corresponding to the one with the highest contribution is selected. Figure 13 shows the flowchart of the proposed method , which 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 explained in [25]. Then, 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 tests. If the Q or T^2 value falls outside their confidence limits for several samples, then there is an abnormality.
3. Best frequency selection phase, where fault diagnosis is 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.

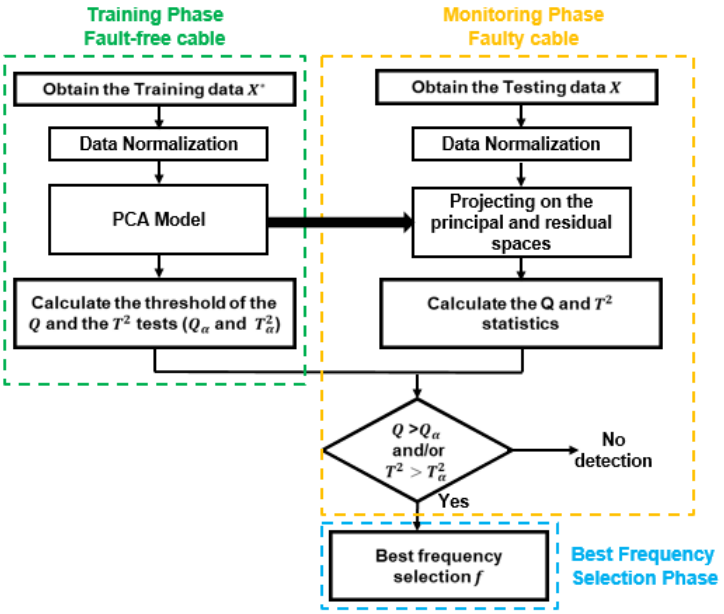


Figure 13: Flowchart of the proposed approach

3.1. Simulation Results

3.1.1. Data-base Building

From the simulation results in [25], we have found that soft faults are better detected in the residual subspace using the Q test. Thus in the following sections, only the Q test will be used. Simulated data are obtained from CST for the $1m$ long cable, and from the $RLCG$-based model for longer cables.

(1) CST simulations: for the cable defined in 2.1 of length $l = 1m$, 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 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 step (1)
- Insert the signatures in the simulation Matlab code
- Choose a cable of length $l = 10m$

- Simulate fault cases characterized by three parameters (36 cases in total):
  (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 $\theta_f = [45°, 90°, 180°]$

Finally, for the $10m$ cable, the TDR responses at the different frequencies ($1GHz$, $2GHz$, $3GHz$, $4GHz$) cab be obtained for the healthy cable and for the 36 faulty cases.

3.1.2. Training

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

$$X^* = [\mathbf{x_1^*} \quad \mathbf{x_2^*} \quad \mathbf{x_3^*} \quad \mathbf{x_4^*}] \tag{3}$$

The obtained data matrix X^* is then used for the construction of the PCA model according to equations (10), (11) and (12) in [25]. Table 6 indicates that the cumulative variance of the first score is 99.9412% that is greater than the lower limit. This implies that the data is well described by one principal component. Thus, k, the number of the retained PC, is equal to one.

Table 6: PCA model variances

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

3.1.3. Monitoring

Our methodology's second step leads to considering a new measurement data set X corresponding to the faulty cases. Each variable (vector) is a concatenated vector of the fault signature data for the 36 fault cases at the same operating frequency. The new data matrix is defined as:

$$X = [\mathbf{x_1} \quad \mathbf{x_2} \quad \mathbf{x_3} \quad \mathbf{x_4}] \tag{4}$$

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

The new measurement matrix is projected into the reference frame obtained from the PCA model, and compared to the reference data using the Q test. The 95% confidence limit $Q_\alpha = 0.0118$ according to [25].

It is observed that several samples have crossed Q_α. This indicates that faults have occurred. The first occurrence of the fault corresponds to the highest impedance variation

in the reflectogram. For each round-trip of the wave, its peak amplitude decreases. Due to the presence of several round-trip peaks for the same fault, several abnormal samples will be produced in the Q chart. Therefore, we consider the samples with the highest Q value as the abnormal ones corresponding to the fault. Twenty eight (28) fault cases out of the thirty six (36) are detected.

3.1.4. Best Frequency Selection

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 are plotted in Figure 14: $\mathbf{x_4}$ had the highest contribution, so the frequency $4GHz$ leads to the best fault detection capability. Table 7 summarizes the results for all the faulty cases: detection result and the selected frequency for the detected faults.

Table 7: Fault detection 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
$F36$		$L_f = 20mm$ $\theta_f = 180°$		

.

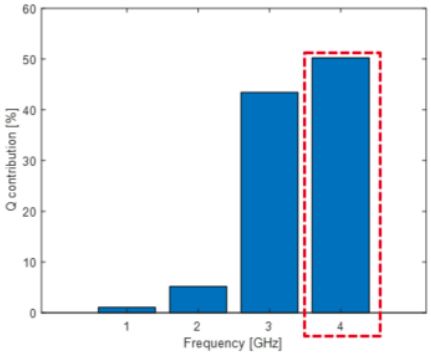
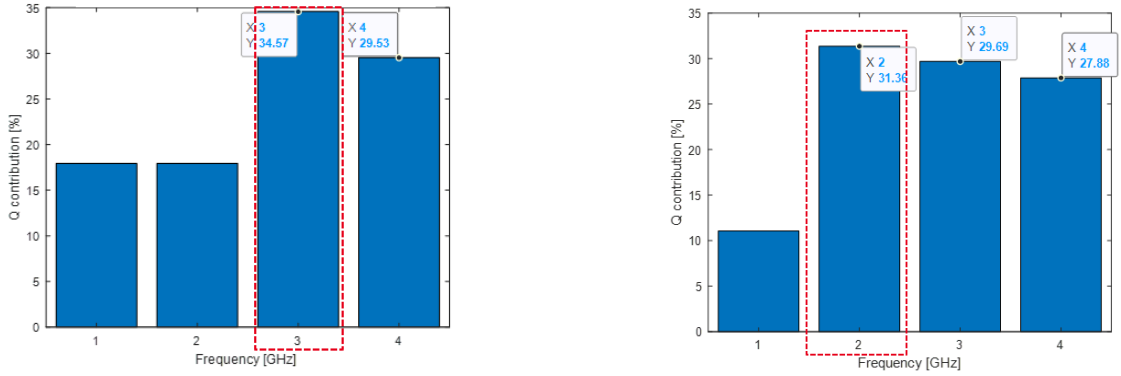


Figure 14: Contribution plot of the fault $F9$

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 15 shows the contribution plots of the abnormal sample corresponding to $F30$ and $F36$.



(a) $F30$ (b) $F36$

Figure 15: Contribution plot of the fault cases at $x_f = 90\%$

In order to explain the obtained results, the energy E_f of the faulty section at each frequency f for each of the indicated fault case is presented in Table 8. It is noticed that the method chooses the frequency with the highest energy.

Table 8: Energy calculation for the detected 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

The combination of reflectometry and PCA coupled to Q statistic, in the case study above, shows that:

- For $x_f = 20\%$, all the faults are detected whatever their severity level, and the selected 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 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 faults with lengths of $5mm$ 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 16, the selected frequency decreases as the fault position approaches the end of the cable.

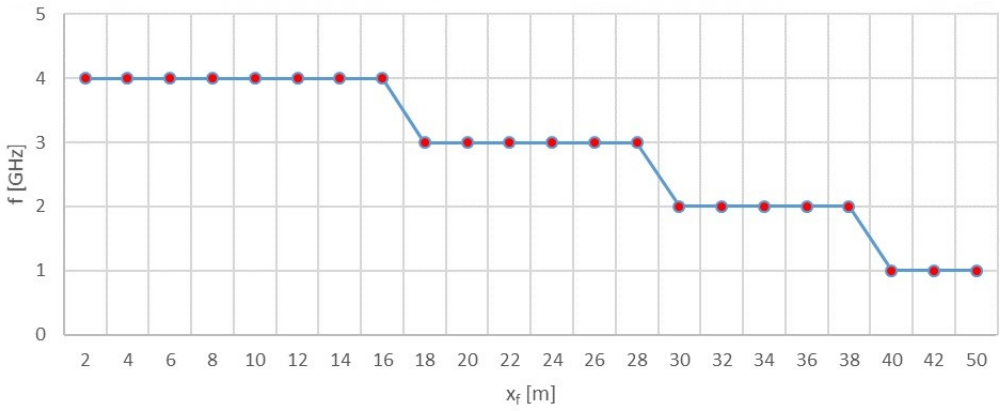


Figure 16: Variation of the selected frequency with the fault position for a $100m$ cable for the fault case ($L_f = 5mm, \theta_f = 180°$)

3.2. Experimental Validation

The methodology for fault detection using PCA, and frequency selection is now evaluated on real cables using the setup displayed in Figure 10b. The reflectometry responses of the

healthy cable and the faulty one (shield damage located at $0.5m$ on a $1m$ cable) are measured and displayed in Figure 11 and Figure 12.

Table 9 indicates that the cumulative variance of the first two scores is 99.99% greater than the lower limit. Using CPV (cumulative percentile variation) method [31], the dimension of the reference frame is set to two. The value of the limit is $Q_\alpha = 0.628$.

Table 9: PCA model variances

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

Figure 17 shows the Q test results of this experiment where several samples are above the threshold. The contribution plot of the abnormal sample is represented in Figure 18. The variable corresponding to frequency $4GHz$ contributes the most to this sample. Hence, $f = 4GHz$ is the selected maximal frequency.

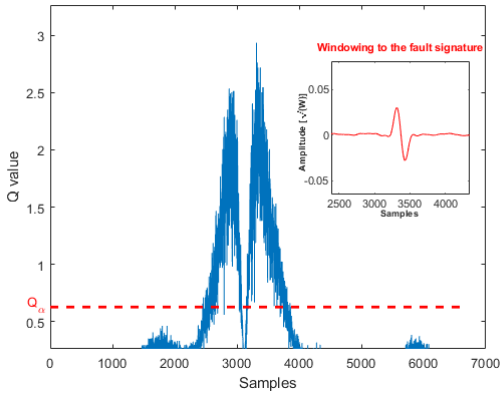


Figure 17: Q values of experimental measurement samples

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

Note that the choice of the frequency depends on several factors from the cable length to the fault characteristics and the environment where the cable is operating. Hence, generalizing the frequency choice is not applicable, it is always depending on the case study and that what is about the proposed method concerns the automatic frequency choice among settled ones according to the case study we have. In our work, even the R2RLCG technology

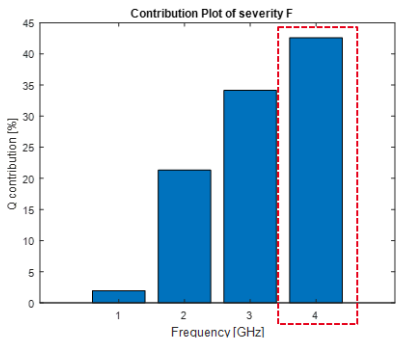


Figure 18: Contribution plot of the abnormal experimental sample corresponding to a $5mm$ long, 180° wide shielding fault

used to obtain the cable RLCG model remains effective on longer cables (around tens of meters). The experimental validation is done for a shorter cable. Nevertheless with no loss of generality, the proposed procedures can be applied and are still valid for longer cable cases.

4. Fault Detection Performance Analysis in the Presence of Noise

The performance of the Q test in the presence of noise is evaluated in this section. Figure 19 represents the flowchart of the methodology organised in three steps; First, the soft fault detection, and the computation of the probability of detection P_d in the presence of noise. Second, the calculation of the probability of false alarm P_{FA} for each SNR. Finally, the best frequency selection in the noisy environment is performed.

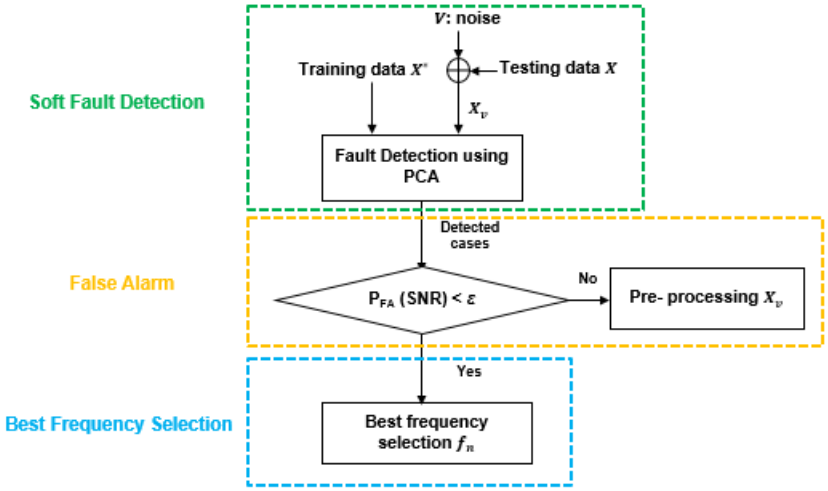


Figure 19: Flowchart of the Q test performance analysis

4.1. Soft Fault Detection

4.1.1. 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 3.1.2 where a $10m$ cable is used and simulated at four different excitation frequencies.

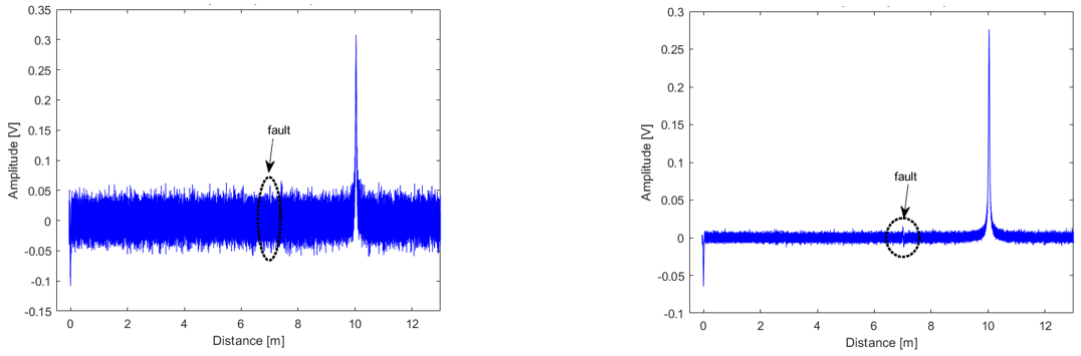
4.1.2. Testing Data Generation: Noise introduction

The testing data used here is X in 3.1.3 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 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 (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 as expressed in 5. Then, 500 realizations are performed at each SNR level.

$$SNR = 10 \log\left(\frac{p_s}{p_\nu}\right) \tag{5}$$

Figure 20 shows the TDR responses in the faulty case ($L_f = 5mm$, $\theta_f = 180°$ at $x_f = 70\%$) for SNR= $-5dB$ and SNR= $10dB$. It is noticed that in both cases, the fault is hidden in the noise making the detection more difficult.



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

Figure 20: TDR response in the presence of F at $x_f = 70\%$ with the effect of the added noise

4.1.3. Fault Detection

For each realization, we apply the Q test to find if the fault 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 SNR. ϵ_d is determined by the application domain. For example in the military, medical and health domains, anomaly detection is a very critical problem and requires a high degree of accuracy [32].

Table 10 gives the P_d of the fault cases $F3$, $F12$, $F21$ and $F30$ for the different SNRs. It is shown that the faults in the presence of noise are detected with a $P_d \geqslant 0.85$. We assume that the detection probabilities are sufficient for the next steps.

Table 10: P_d for different noise levels

SNR [dB]	$F3$	$F12$	$F21$	$F30$
-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

4.2. False Alarm Analysis

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. Then P_{FA} is calculated as the average on the 500 realizations. The results are displayed in Table 11. We can notice that P_{FA} is not null when the noise level is equal or higher than the signal level. Hence, if $P_{FA} < \epsilon$, we proceed to the best frequency selection part. We assume in the

Table 11: P_{FA} for different noise levels

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

following that the obtained false alarm probabilities are lower than the threshold.

4.3. Best Frequency Selection

4.3.1. Robustness to Noise Evaluation

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

For the case under study, Figure 21 represents the variation of the selected frequency with the SNR for different fault positions. It is noted that the selected frequency depends on fault position and the SNR: the variation is similar for $x_f = 20\%$ and $x_f = 50\%$. As the fault reaches the cable end, the selected frequency decreases. This variation was predictable 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.

The trend of the selected frequency is similar for a $100m$ cable as shown in Figure 22.

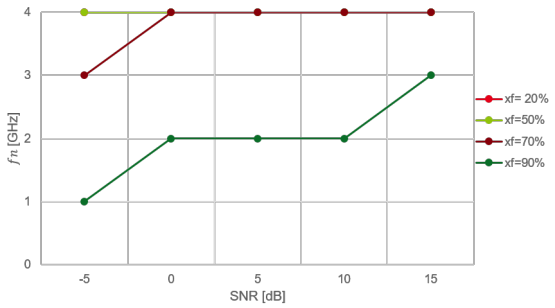


Figure 21: Frequency variation with the SNR for a $10m$ cable with a $5mm$ and 180° fault

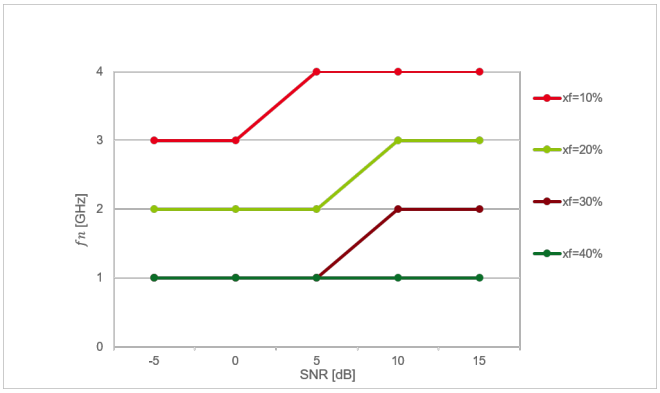


Figure 22: Frequency variation with the SNR for a $100m$ cable with a $5mm$ and 180° fault

4.3.2. Frequency Occurrence

For all the 28 detected fault cases in section 4.1.3, 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%:

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? Table 12 presents the results for 500 realizations which are similar for 2000 realizations. We can observe the the percentages of occurrence for both frequencies are close, even it is a little higher for $3GHz$.

Table 13 and Table 14 display the results for $F30$ adn $F35$.

Table 12: 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%**

Table 13: 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%

Table 14: 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%

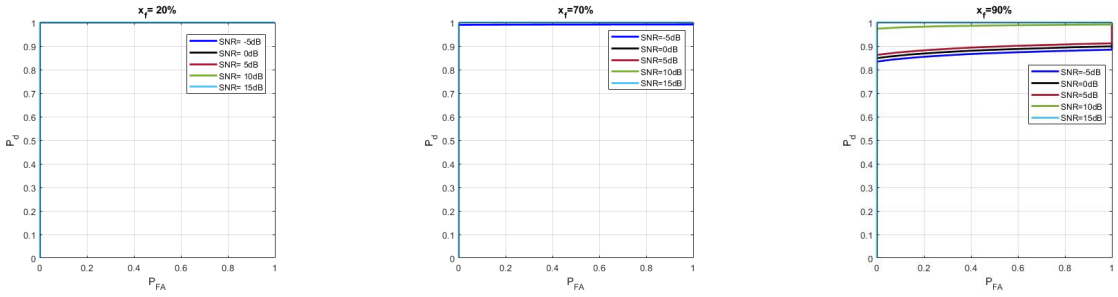
4.3.3. ROC Curves Investigation

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°) for a $10m$ cable case. The detection performance results using the Q test for different SNR levels are displayed in Figure 23, for $x_f = 20\%$, $x_f = 70\%$ and $x_f = 90\%$, respectively.

Thanks to the simulation results for 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 100% detection capability for the noise levels (SNR $\geq 0dB$) with a low false alarm probability for $x_f = 20\%$ and $x_f = 70\%$.



(a) (b) (c)

Figure 23: ROC curves for a $10m$ cable with a $5mm$ and 180° fault

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

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$= -5dB$. To conclude, Q test is efficient for detecting soft faults, but its performance is affected by the SNR when the fault position is close to the cable end.

5. Frequency Selection Robustness to Environmental Noise

The performance of the selected frequency f in the presence of noise in the fault case detection is now studied. Figure 24 represents the flowchart of the proposed methodology. It consists of the following steps:

- From the PCA results in section 3.1.4, 28 fault cases are detected and the selected frequency f for each fault case is indicated in Table 7. Noise is added to the signal at frequency f and 500 realizations are performed.
- A fusion algorithm [33] of several post processing methods (Signature Magnification by Selective Windowing (SMSW) [18], subtractive correlation method, and the method based on the integral of a reflectogram [34]) 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 4.3.2, 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 using the Fusion algorithm.

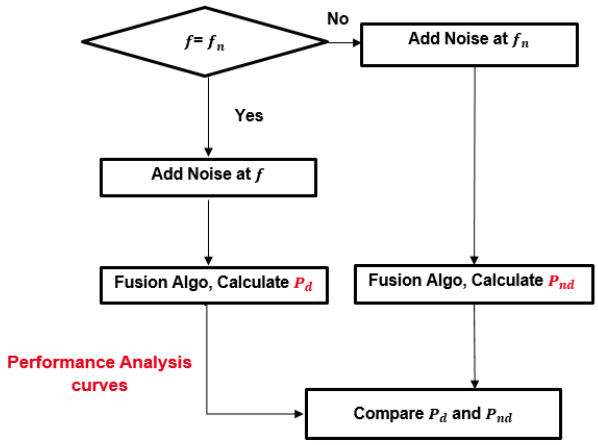


Figure 24: Flowchart of the performance of the selected frequency f in the presence of noise

Now, to obtain the performance analysis curves, each detected fault case in Table 7, at frequency f, for each SNR, P_d is calculated using (6) by averaging the probability of detection of all the realizations $r = \{1, 2, \ldots, 500\}$:

$$P_d = \frac{\sum_{r=1}^{500} p_r}{500} \tag{6}$$

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

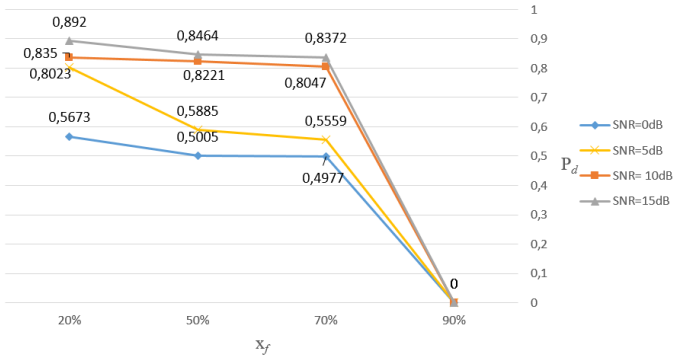


Figure 25: Performance analysis curve in the case of a $20mm$ and 45° fault

The case study is the fault case with $L_f = 20mm$ and three different severities ($\theta_f = 45°$, $\theta_f = 90°$ and $\theta_f = 180°$). For each severity, the fault position varies along the cable. P_d is calculated for different fault positions and for each SNR.

(1) The performance analysis curve for the first severity ($L_f = 20mm$, $\theta_f = 45°$) is represented in Figure 25. It is observed that, for each SNR, the detection capability vanishes as the fault reaches the cable end ($x_f = 90\%$).

(2) Figure 26 represents the performance analysis curve for the second severity ($L_f = 20mm$,$\theta_f = 90°$). The detection capability compared to the previous case is higher for all noise levels. We can also notice that for low noise levels, the probability of detection is higher than 0.8.

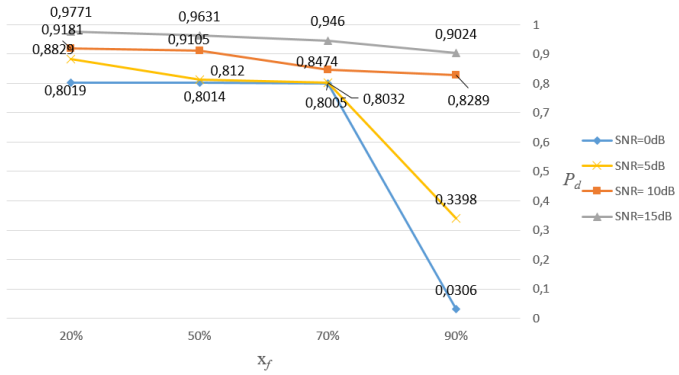


Figure 26: Performance analysis curve in the case of a $20mm$ and 90° fault

(3) Figure 27 represents the performance analysis curve for the third severity level ($L_f = 20mm$,$\theta_f = 180°$). The detection capability for this severity is higher than the two previous cases. We can notice the same decreasing trend when the fault is close to the cable end. However, for low noise levels (10 and $15dB$), the detection capability is almost constant and higher than 0.94.

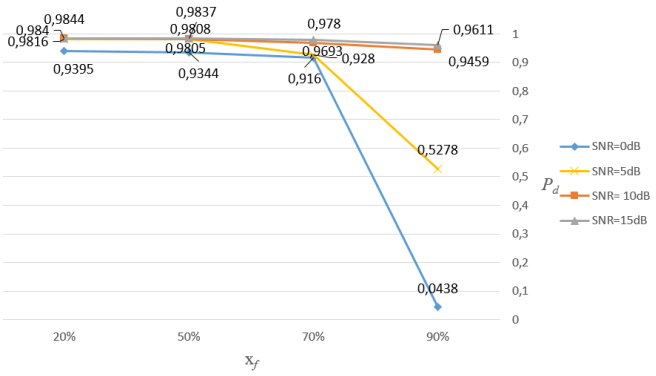


Figure 27: Performance analysis curve in the case of a $20mm$ and 180° fault

From these results, we can conclude that the probability of detection is higher when the fault occurs near the beginning of the cable and as the SNR increases for the same fault severity. The soft faults with the lowest severity level cannot be detected near the cable's end. And the situation is even worse when the noise level is close to the signal level.

6. Conclusion

This study 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 to design the PCA model. The new measurements are projected into the PCA model reference frame and compared to the reference data. If a difference is observed, 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 enabling automatic computerized decision-making.

Our results have shown 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 noise level. Regarding the statistical tests, the Q criteria is more relevant. The simulation results are based on a 3D electromagnetic model using CST for short cables and an $RLCG$-based model for longer cables. The models have been validated with experimental measurements in the case of shielding damage. Once validated, the parametric model has been used for extensive simulations to cope with the different fault configurations.

Experimental validations in the presence of soft faults were carried out. The performance analysis of this method has also been investigated. For the considered use case, when the soft fault is near the injection point, the detection probability equals one even for SNR values as low as $0dB$. As the fault position approaches the end of the cable, the performance is still acceptable, but for lower fault severities, the detection is almost impossible. Future works can focus on other types of noise such as impulsive noise in vehicular networks, mechanical noise in airplanes, etc. Their impact on detection performance should be evaluated. Moreover, the case of different types of cables should be also considered to see how the methodology can be adapted to them based on their specific properties.

References

[1] Schuet S, Timuçin D, Wheeler K. Physics-Based Precursor Wiring Diagnostics for Shielded-Twisted-Pair Cable. IEEE Transactions on Instrumentation and Measurement; 64(2):378-391, Feb 2015.

[2] Osman O, Sallem S, Sommervogel L, Carrion MO, Bonnet P and Paladian F. Distributed Reflectometry for Soft Fault Identification in Wired Networks Using Neural Network and Genetic Algorithm. IEEE Sensors Journal; 20(9):4850-4858, May 2020.

[3] Addad M, Djebbari A. Spread Spectrum Sensing Based on ZCZ Sequences for the Diagnosis of Noisy Wired Networks. IEEE Sensors Journal; 21(2):914-920, Jan 2021.

[4] Furse C, Kafal M, Razzaghi R, Shin Y-J. Fault Diagnosis for Electrical Systems and Power Networks: A Review. IEEE Sensors Journal; 21(2):888-906, Jan 2021.

[5] Schonfeld J, Greulich O, Patterson-Hine A, Lee L, Cockrell J, Hofland L. Wire integrity research (WIRe Pilot Study). Doc. No. A0SP-0001-XB1. Aug, 2000.

[6] Chung YC, Amarnath NN, Furse CM, Mahoney J. Capacitance and inductance sensors for location of open and short circuited wires. IEEE Transaction on Instrumentation and Measurement; 58(8):2495-2502, Aug 2009.

[7] Auzanneau F, Ravot N, Incarbone L. Chaos Time Domain Reflectometry for Online Defect Detection in Noisy Wired Networks. IEEE Sensors Journal; 16(22):8027-8034, Nov 2016.

[8] Cabanillas E, Layer C, Kafal M, Dupret A. Enhancing the Spatial Resolution for Wire Fault Detection Systems Using Multi-Carrier Signals. IEEE Sensors Journal; 18(23):9857-9866, Dec 2018.
[9] Lee C-K, and Chang SJ. A Method of Fault Localization Within the Blind Spot Using the Hybridization Between TDR and Wavelet Transform. IEEE Sensors Journal; 21(4):5102-5110, Feb 2021.
[10] Hashemian HM. State-of-the-art predictive maintenance techniques. IEEE Transactions on Instrumentation and measurement; 60(1):226-36, Dec 2010.
[11] Furse C, Haupt R. Aging, brittle wiring within aircraft poses a hidden hazard that emerging technologies aim to address. IEEE spectrum; 38(2):34–39, February 2001.
[12] Furse C, Chung YC, Dangol R, Nielsen M, Mabey G, Woodward R. Frequency-domain reflectometry for on-board testing of aging aircraft wiring. IEEE Transactions on Electromagnetic Compatibility; 45(2):306–315, May 2003.
[13] Cerri G, De Leo R, Della Nebbia L, Pennesi S, Primiani VM, Russo P. Fault location on shielded cables: Electromagnetic modelling and improved measurement data processing. IEE Proceedings-Science, Measurement and Technology; 152(5):217-26, Sep 2005.
[14] Buccella C, Feliziani M, Manzi G. Accurate detection of low entity cable faults by wavelet transform. In 2004 International Symposium on Electromagnetic Compatibility (IEEE Cat. No. 04CH37559), (Vol. 3, pp. 936-941). IEEE, Aug 2004.
[15] Shin YJ. Theory and application of time-frequency analysis to transient phenomena in electric power and other physical systems. The University of Texas at Austin, 2004.
[16] Franchet M, Ravot N, Picon O. Soft fault detection in cables using the cluster time-frequency domain reflectometry. IEEE Electromagnetic Compatibility Magazine; 2(1):54-69, May 2013.
[17] Sallem S, Ravot N. Self-adaptive correlation method for soft defect detection in cable by reflectometry. In SENSORS, (pp. 2114-2117). IEEE, Nov 2014.
[18] Sallem S, Ravot N. Soft defects localization by signature magnification with selective windowing. In 2015 IEEE SENSORS (pp. 1-4). IEEE, Nov 2015.
[19] Oumri M, Zhang Q, Sorine M. A reduced model of reflectometry for wired electric networks. In 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS), Budapest, Hungary, Jul 2010.
[20] Auzanneau F. Wire troubleshooting and diagnosis: Review and perspectives. Progress In Electromagnetics Research B; 49:253-79, 2013.
[21] Cohen L. Wave propagation with dispersion and damping. In Advanced Signal Processing Algorithms, Architectures, and Implementations XIV (Vol. 5559, pp. 201-220). International Society for Optics and Photonics, Oct 2004.
[22] Chen C, Roemer L. Attenuation and dispersion compensation via cepstral processing. IEEE transactions on acoustics, speech, and signal processing; 29(5):1080-4, Oct 1981.
[23] Penha R, Hines JW. 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, May 2001.
[24] Smaïl MK, Le Bihan Y, Pichon L. Fast diagnosis of transmission lines using neural networks and principal component analysis. International Journal of Applied Electromagnetics and Mechanics; 39(1-4):435-41, Jan, 2012.
[25] Taki N, Ben Hassen W, Ravot N, Delpha C, Diallo D. Frequency selection for reflectometry-based soft fault detection using principal component analysis. In2019 Prognostics and System Health Management Conference (PHM-Paris), (pp. 273-278). IEEE May 2019.
[26] Wheeler KR, Twombly IX, Goebel KF, Wysocki PF. Aging aircraft wiring fault detection survey. Technical report, NASA Ames Research Center, Moffett Field CA 94035, June 2007.
[27] Cohen J, Gregis N, inventors; Commissariat a l'Energie Atomique et aux Energies Alternatives, assignee. Method of determining lineal parameters of a transmission line. United States patent US 10,746,771. Aug 18 2020.
[28] Ben Hassen W, Kafal M. Shielding Damage Characterization in Twisted Pair Cables Using OMTDR-based Reflectometry and Inverse Problems. In 2019 PhotonIcs & Electromagnetics Research

Symposium-Spring (PIERS-Spring), (pp. 3093-3101). IEEE, Jun 2019.
[29] Griffiths LA, Parakh R, Furse C, Baker B. The invisible fray: A critical analysis of the use of reflectometry for fray location. IEEE Sensors Journal;6(3):697-706, Jun, 2006.
[30] Slišković D, Grbić R, Hocenski Ž. Multivariate statistical process monitoring. Tehnicki Vjesnik-Technical Gazette; 19(1):33-41, Mar 2012.
[31] Zumoffen D, Basualdo M. From large chemical plant data to fault diagnosis integrated to decentralized fault-tolerant control: pulp mill process application. Industrial & Engineering Chemistry Research; 47(4):1201-20, Feb 2008.
[32] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM computing surveys (CSUR); 41(3):1-58, Jul 2009.
[33] Ben Hassen W, Roman MG, Charnier B, Ravot N, Dupret A, Zanchetta A, Morel F. Embedded OMTDR sensor for small soft fault location on aging aircraft wiring systems. Procedia Engineering; 168:1698-701, Jan 2016.
[34] Incarbone L, Roman MG, inventors; Commissariat a l'Energie Atomique et aux Energies Alternatives, assignee. Method for detecting soft faults in a cable, which method is based on the integral of a reflectogram. United States patent US 10,690,712. Jun 2020.
[35] Loete F, Zhang Q, Sorine M. Inverse scattering experiments for electrical cable soft fault diagnosis and connector location. PIERS proceedings, Mar 2012.
[36] Papazyan R, Pettersson P, Edin H, Eriksson R, Gafvert U. Extraction of high frequency power cable characteristics from S-parameter measurements. IEEE Transactions on Dielectrics and Electrical Insulation; 11(3):461-70, Oct 2004.
[37] Shi Q, Tröltzsch U, Kanoun O. Analysis of the parameters of a lossy coaxial cable for cable fault location. The 8th International Multi-Conference on Systems, Signals & Devices;(pp. 1-6), Mar 2011. IEEE.
[38] Baker-Jarvis J, Vanzura EJ, Kissick WA. Improved technique for determining complex permittivity with the transmission/reflection method. IEEE Transactions on microwave theory and techniques; 38(8):1096-103, Aug 1990.
[39] Shin YJ , Powers J, Choe TS, Hong CY, Song ES, Yook JG, and Park JB. Application of Time-Frequency Domain Reflectometry for Detection and Localization of a Fault on a Coaxial Cable, IEEE Transactions on Instrumentation and Measurement, Vol. 54, No. 6, Dec. 2005
[40] Cataldo A, De Benedetto E, Masciullo A, Cannazza G. A new measurement algorithm for TDR-based localization of large dielectric permittivity variations in long-distance cable systems, Measurement, Volume 174, 2021
[41] Shi Q, Kanoun A. Analysis of the Parameters of a Lossy Coaxial Cable for Cable Fault Detection, IEEE Transaction on Systems, Signals and Devices, vol.7, no.4, pp. 311-325, 2012
[42] Zhang A, Gao C, Yang W and Li Q. Physical Defect Localizing Methodology for Coaxial Cable Based on Quadratic Propagation Coefficient Model, IEEE Sensors Journal, vol. 21, no. 2, pp. 1017-1025, 15 Jan.15, 2021
[43] Furse CM, Chung YC, Lo C and Pendayala P. A critical comparison of reflectometry methods for location of wiring faults, Smart Structures and Systems, 2006, 2, 25-46
[44] Qi G, Zhu Z, Erqinhu K, Chen Y, Chai Y, Sun J. Fault-diagnosis for reciprocating compressors using big data and machine learning. Simulation Modelling Practice and Theory; 80:104-2, Jan 2018.
[45] Kourti T. Process analysis and abnormal situation detection: from theory to practice. IEEE control systems magazine; 22(5):10-25, Nov 2002. ‘