Transient-based fault identification algorithm for the protection of meshed HVDC grids

Paul Verrax

université
PARIS-SACLAY

Transient-based fault identification algorithm for the protection of meshed HVDC grids

Algorithme d'identification des défauts pour la protection des réseaux maillés HVDC

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
Unité de recherche : Laboratoire des signaux et systèmes, Université Paris-Saclay, CNRS, CentraleSupélec, 91 190, Gif-sur-Yvette, France
Référent : Faculté des sciences d'Orsay

Thèse présentée et soutenue à Villeurbanne, le 16 avril 2021, par

Paul VERRAX

Composition du jury :

Marc PETIT Professeur, Université Paris-Saclay	Président
Roberto BENATO Professeur, Università di Padova	Rapporteur & Examinateur
Thierry POINOT Professeur, Université de Poitiers	Rapporteur & Examinateur
Corinne ALONSO Professeure, Université Toulouse III - Paul Sabatier	Examinatrice
Filipe FARIA DA SILVA Professeur, Aalborg Universitet	Examinateur

Direction de la thèse :

Michel KIEFFER Professeur, Université Paris-Saclay	Directeur de thèse
Bertrand RAISON Professeur, Université Grenoble Alpes	Co-directeur de thèse
Alberto BERTINATO Ingénieur, Supergrid Institute	Co-encadrant de thèse
Sebastian DAMBONE SESSA Docteur, Università di Padova	Invité
Li ZOU Docteur, GE Renewable Energy	Invité

Thèse de doctorat

NNT: 2021UPASG028

Thesis to get the degree of a doctor of philosophy

Transient-based fault identification algorithm for the protection of meshed HVDC grids

Paul Verrax

Université Paris-Saclay

"Failure is not an option"

Apocryphal, attributed to Gene Kranz

« Le vent se lève ! . . . Il faut tenter de vivre !
L'air immense ouvre et referme mon livre »

Paul Valéry

Contents

List of Figures

List of Tables

Remerciements

A la fin de ces trois années de travail, mes remerciements vont tout d'abord à mes encadrants de thèse, Alberto Bertinato, Bertrand Raison, et tout particulièrement à mon directeur de thèse Michel Kieffer. Davantage que le sujet, l'encadrement est peut-être l'élément le plus important d'un projet de thèse. Bien qu'à Paris et Grenoble respectivement, Michel et Bertrand ont su m'accompagner et me guider tout au long de ces trois années, sans jamais ménager leurs efforts et leur temps.

Je souhaite également remercier mes collègues à SuperGrid, tant au programme Architecture & Systems que dans les autres programmes de recherche, et jusqu'au hors programme. Grâce à eux, de séances d'escalades en séances de sports de combats, et de pause cafés en afterworks, j'ai passé trois années aussi agréables que stimulantes. Je voudrais saluer les collègues de la protection, et en particulier Pascal Torwelle, indéfectible partenaire de thèse, de simulations, et de courses...Cheers mate !

Au delà de SuperGrid, j'ai pu compter sur le soutient au long cours de mes amis qui ont su supporter mes interminables tentatives de vulgarisation de mes travaux de thèse. Qu'ils soient à Lyon, comme Alix, Adèle, Charlotte, et Rémi ou plus loin, comme Chloé, Sixtine, Sophie, Stéphane, et Walfroy, ils sont pour beaucoup dans ma bonne humeur durant ces trois années.

Enfin, je suis infiniment reconnaissant envers ma famille, et tout particulièrement ma mère et ma sœur. Aillant déjà parcouru ce chemin sinueux, elles m'ont soutenu, encouragé, (et même parfois hébergé) tout au long de cette thèse.

Résumé

La protection des réseaux maillés HVDC se heurte à différents verrous parmi lesquels l'identification fiable et rapide des défauts affectant les lignes du réseau. Cette thèse propose un concept d'algorithme d'identification basé sur l'estimation des paramètres du défaut. Un modèle analytique décrivant l'évolution de la tension et du courant dans la ligne en défaut est développé. Le modèle obtenu dépend explicitement des paramètres du défaut (*e.g.* la distance et la résistance du défaut). Lorsqu'un défaut est détecté, le relais protégeant la ligne considérée estime les paramètres du défaut, a priori inconnus, à partir des mesures reçues des capteurs les plus proches. Le résultat de l'estimation des paramètres permet de confirmer ou non que la ligne considérée est en défaut. En particulier, un indicateur de confiance évalue la précision des paramètres estimés. La méthode proposée est testée sur un réseau 4 terminaux basé sur des lignes aériennes, implémenté dans un logiciel de simulation des transitoires électromagnétiques. La présence d'inductances de lignes n'est pas requise pour identifier fiablement les défauts. L'algorithme est capable d'identifier la plupart des défauts en utilisant une fenêtre de mesure réduite (moins de 0.5 ms). L'algorithme d'identification peut être intégré dans une stratégie de protection sélective, où chaque élément du réseau est protégé individuellement. La compatibilité de l'approche avec les aspects systèmes tels que le dimensionnement des disjoncteurs ou les spécifications des capteurs est aussi vérifiée.

Abstract

The protection of meshed HVDC grids comprises many locks, among which the reliable and fast identification of faults affecting the lines of the grid is seen as particularly challenging. The present thesis investigates the concept of a single-ended fault identification algorithm based on the estimation of the fault parameters. A closed-form model of the transient evolution of the voltage and current in a faulty line is first developed. The obtained model depends explicitly on the fault parameters (*e.g.* the fault distance and resistance) and is thus adaptable to various fault cases. When a fault is suspected at a relay protecting one of the transmission lines, the proposed algorithm estimates the unknown fault parameters based on the received measurements and the developed model. The result of the parameter estimation process allows the relay to confirm or reject that the protected faulty line is faulty or not. In particular, a confidence indicator is employed to evaluate the accuracy of the estimated fault parameters. The proposed method is tested on a 4 station meshed grid based on overhead lines, implemented in Electro-Magnetic Transient software. Contrary to most existing approaches, the presence of DC reactors at the extremity of each lines is not required for the algorithm to distinguish between faults occurring on different transmission lines. The proposed approach is able to identify most of the faults while using a short observation window (less than 0.5

ms long). The fault identification algorithm is thus a good candidate to be used in a selective fault clearing strategy, where each elements of the grid is protected individually. The compliance of the approach with system requirements such as the DC circuit breakers and the sensor specifications is also investigated.

Résumé étendu

L'intégration massive des énergies renouvelables intermittentes entraîne une évolution du système de transmission d'électricité. La transmission par courant continu est appelée à jouer un rôle de plus en plus important, tant pour le raccordement des sources d'énergies distantes comme les parcs éoliens en mer que pour la transmission de puissance sur de longues distances. Si de nombreuses liaisons à Courant Continu Haute Tension (CCHT) existent aujourd'hui en point-à-point, la future mise en place de réseaux maillés, voir Figure 0.1, permettrait d'accroître les capacités de transmission à un coût réduit. L'établissement de tels réseaux se heurte toutefois à des verrous tant techniques que réglementaires ou économiques. Parmi les verrous techniques, la protection du réseau est vue comme particulièrement difficile. La protection des réseaux électriques vise à diminuer l'impact des défauts, tels que les courts-circuits, sur l'approvisionnement en électricité et la stabilité du réseau. Les dommages éventuels aux équipements comme les stations de conversions doivent aussi être minimisés.

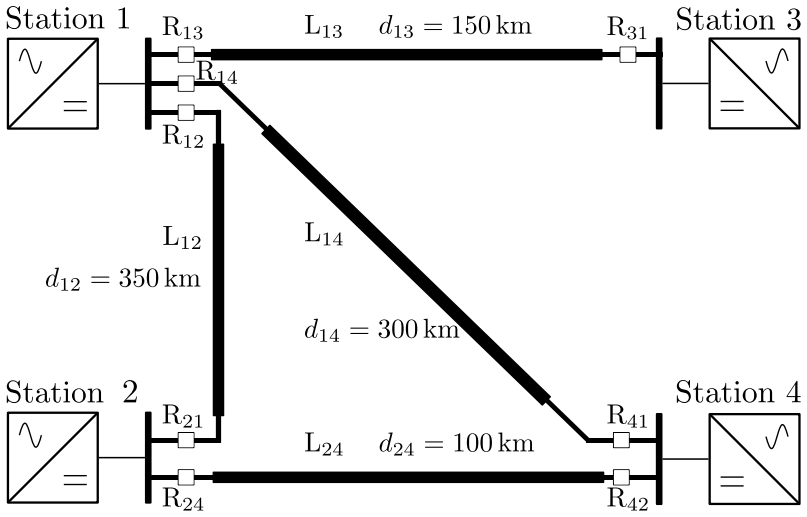


Figure 0.1. : Schéma simplifié d'un réseau maillé CCHT à 4 terminaux

Lorsqu'un défaut affecte une ligne du réseau, différentes stratégies pour l'éliminer existent. Les stratégies dites non sélectives entraînent la déconnexion temporaire de tout le réseau avant l'isolation de la ligne en défaut. A l'inverse, dans une stratégie sélective, seule la ligne en défaut est déconnectée et le reste du réseau reste opérationnel. Les stratégies sélectives sont donc préférables en ce qu'elles diminuent

l'impact des défauts sur le réseau CCHT ainsi que sur le réseau alternatif environnant. Dans une stratégie sélective, la déconnexion des lignes de transmission repose sur les disjoncteurs placés aux extrémités de chaque ligne. L'ouverture des disjoncteurs, commandée par les relais de lignes, permet la déconnexion de la ligne en cas de défaut. Chaque relais doit donc être capable de détecter la survenue d'un défaut dans le réseau puis d'identifier si le défaut affecte la ligne qu'il protège. Cette tâche d'identification du défaut doit garantir que tous les défauts survenant sur la ligne protégée par le relais sont bien identifiés mais que les défauts affectant une autre partie du réseau ne le sont pas. De plus, l'identification doit être suffisamment rapide, de l'ordre de la milliseconde, pour permettre l'ouverture des disjoncteurs avant que le courant de défaut ne soit trop important.

Le problème de l'identification rapide et fiable, par les relais de ligne, des défauts, est un sujet de recherche actif. La plupart des solutions existantes tirent parti de la présence d'inductances à l'extrémité de chaque ligne. Ces inductances ralentissent montée en courant suite au défaut mais facilitent aussi la discrimination entre les défauts affectant différentes lignes. Toutefois, ces inductances ont un coût important et peuvent également diminuer la stabilité du réseau. Pour les réseaux basés uniquement sur des lignes aériennes, le comportement inductif de ces dernières rend la montée en courant moins rapide que dans le cas des câbles souterrains. Combiné à l'utilisation de disjoncteurs hybrides, une identification rapide du défaut permet de déconnecter la ligne en quelques millisecondes, sans pour autant avoir recours à des inductances de lignes. Une telle solution requiert alors des algorithmes identifiant les défauts en moins d'une milliseconde, et se basant donc uniquement sur les mesures des capteurs à proximité immédiate du relais.

Cette thèse propose un algorithme d'identification des défauts se basant sur l'estimation des paramètres du défaut. Le principe général de l'algorithme est présenté sur la Figure 0.2. Chaque relais dispose d'un modèle décrivant l'évolution transitoire de la tension et du courant en cas de défaut affectant la ligne protégée par le relais. Un défaut affectant une ligne peut être caractérisé par deux paramètres principaux : sa position le long de la ligne et son impédance. Le modèle embarqué à chaque relais dépend explicitement de ces paramètres, qui sont a priori inconnus lorsque un défaut est suspecté. Après la détection du défaut, l'algorithme estime les paramètres du défaut à partir du modèle et des mesures reçues en présupposant que la ligne protégée est en défaut. Si l'estimation des paramètres permet de faire correspondre le modèle avec les mesures reçues, l'algorithme identifie le défaut sur la ligne protégée. A l'inverse, si l'algorithme n'aboutit pas à une estimation des paramètres de défaut jugée suffisamment précise, le défaut n'est pas identifié et est donc considéré comme extérieur à la ligne protégé par le relais. Cette discrimination se fait par une logique de décision ayant notamment recours à une évaluation de la région de confiance des paramètres estimés.

L'estimation des paramètres requiert un modèle décrivant les phénomènes transitoires affectant la ligne en défaut. La survenue d'un défaut engendre la propagation d'ondes de tension et de courant à travers la ligne puis le réseau. L'évolution de ces

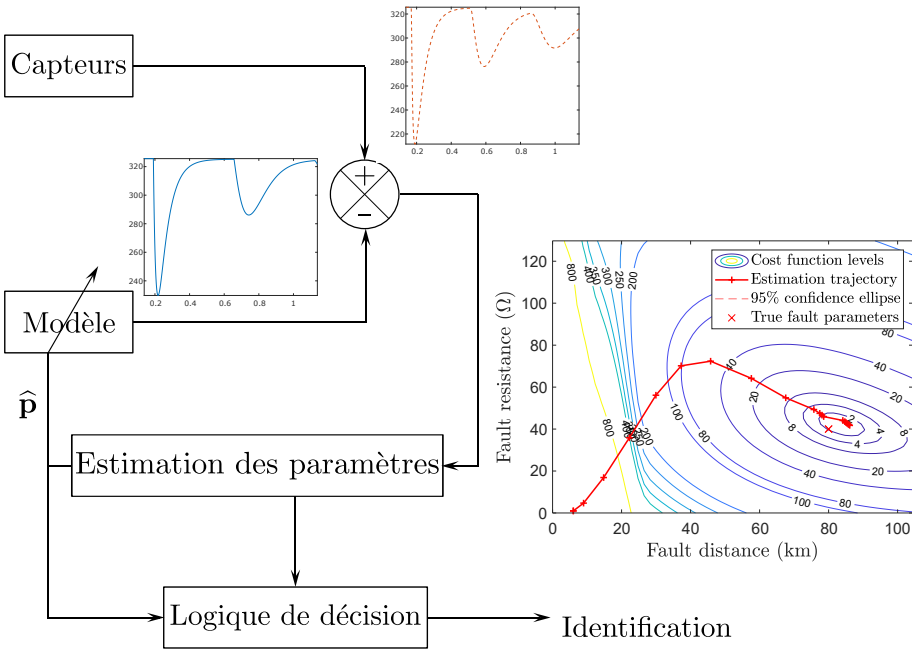


Figure 0.2. : Principe de l'algorithme d'identification basé sur l'estimation des paramètres du défaut.

ondes peut être modélisée avec des outils tels que l'équation des télégraphistes et, dans le cas de lignes multi-conducteurs, l'analyse modale. Négliger la variation en fréquence des paramètres de lignes ainsi que l'emploi de modèles simplifiés pour les stations de conversion permet d'établir un modèle analytique des premières ondes progressives. Ce modèle ne rend toutefois pas compte de la distorsion due à l'effet de la résistivité du sol. La prise en compte de ces effets requiert généralement des calculs importants mal adaptés à l'évaluation rapide du modèle. Nous proposons une approche comportementale qui représente l'essentiel de la distorsion liée à la résistivité du sol en utilisant des filtres passe-bas. L'effet de la résistivité dépendant de la distance parcourue par l'onde, des filtres différents doivent être réglés pour différentes distances de défauts. Nous introduisons une description explicite de ces filtres en fonction de la distance du défaut, permettant la modélisation de l'effet de terre pour n'importe quelle distance parcourue, à partir d'un nombre réduits de filtres pré-déterminés. Dans le cas d'une ligne multi-conducteurs, le filtrage n'est appliqué qu'au mode terre, les modes aériens n'étant que peu impactés par la distorsion. Le modèle obtenu exploite donc une approche à la fois physique, notamment pour les interactions des ondes avec les stations, et comportementale, pour la prise en compte de l'effet de terre. Malgré cette approche, le modèle dépend explicitement des paramètres du défaut, ce qui est un atout pour l'estimation de ces paramètres.

Lorsqu'un défaut est suspecté à un relais, l'algorithme d'identification estime les paramètres de défaut à partir des mesures reçues et du modèle paramétrique. Une estimation au sens du maximum de vraisemblance est utilisée. Sous certaines hypo-

thèses, cela conduit à la résolution d'un problème des moindres carrés pondérés. Des outils d'optimisation itératifs tels que l'algorithme de Levenberg Marquardt offrent une résolution assez efficace de ce problème. La logique de décision confirmant ou non que le défaut soit sur la ligne protégée par le relais se base sur deux tests. Un test de validité compare la valeur des paramètres estimés avec des valeurs minimales et maximales attendues pour la résistance et la distance du défaut. Un test de précision compare ensuite la surface de la région de confiance à 95% des paramètres estimés avec un seuil préétabli. Le défaut est identifié sur la ligne lorsque les deux tests sont satisfaits. Si ce n'est pas le cas, de nouvelles mesures sont ajoutées pour améliorer l'estimation des paramètres. L'estimation est donc récursive en ce qu'elle considère une fenêtre de mesure de plus en plus grande. Si le défaut n'est pas identifié après l'utilisation d'une fenêtre de mesure d'une taille maximale, le défaut est considéré comme externe à la ligne protégée.

L'algorithme d'identification proposé est testé en utilisant comme source de données le logiciel de simulation des transitoires électromagnétiques EMTP-RV. Le fonctionnement de l'algorithme est d'abord détaillé sur des exemples particuliers. Des simulations extensives permettent ensuite d'évaluer les performances de l'algorithme sur un grand nombre de défauts affectant le réseau test représenté sur la Figure 0.1. L'algorithme est capable d'identifier correctement la grande majorité des défauts internes tout en rejetant les défauts se produisant affectant les autres lignes du réseau. La fenêtre de mesure requise pour l'identification des défauts internes est particulièrement courte, de l'ordre de 0.3 ms pour la plupart des défauts, et toujours en dessous de 1 ms. Ces performances sont comparées avec une méthode de la littérature, montrant en particulier que l'algorithme que nous proposons est davantage robuste vis-à-vis des défauts externes.

Les cas pour lesquels l'algorithme échoue à identifier correctement le défaut concernent les défauts se produisant à l'extrémité opposée de la ligne protégée, voire sur la ligne voisine. De tels défauts n'engendrent pas une montée en courant importante au relais considéré, ce qui rend possible l'utilisation de la communication. A l'inverse, les défauts situés à quelques kilomètres du relais nécessitent d'ouvrir les disjoncteurs immédiatement. La distinction entre ces deux types de défauts, appelés respectivement non sévères et sévères, peut être faite par l'algorithme à partir de l'estimation des paramètres du défaut. Dans le cas de défauts non sévères, la collaboration des deux relais protégeant une même ligne permet de diminuer le risque d'échec de l'identification.

L'algorithme d'identification de défauts peut être intégré dans une stratégie de protection sélective. Dans une telle approche, chaque ligne de transmission est protégée individuellement par les deux disjoncteurs situés aux extrémités de la ligne. Il faut donc s'assurer que le courant dans la ligne en défaut ne dépasse pas le pouvoir de coupure de ces disjoncteurs à l'instant d'ouverture. L'emploi de disjoncteurs hybrides concilie un pouvoir de coupure assez important (16 kA) et un temps d'ouverture rapide (2 ms). Associés à une identification rapide du défaut (en moins d'une milliseconde comme suggéré par les simulations extensives), de tels disjoncteurs hy-

brides permettent l'ouverture de la ligne en défaut sans pour autant recourir à des inductances de ligne pour ralentir la montée en courant. L'identification et donc la déconnexion rapide de la ligne en défaut limite également le nombre de défauts menant au blocage d'une ou plusieurs stations de conversion, ce qui diminue l'impact du défaut sur le réseau.

L'approche proposée présente des résultats prometteurs, en particulier concernant la fiabilité de la méthode et le nombre très limité de mesures nécessaires à l'identification du défaut. Toutefois, les temps de calculs obtenus avec une implémentation en Matlab ($\sim 50\,\mathrm{ms}$) sont environ 500 fois trop lents, comparés à un objectif d'exécution en temps réel ($\sim 0.5\,\mathrm{ms}$). Si des pistes existent pour diminuer le nombre de calculs effectués par l'algorithme, l'implémentation temps réel de la méthode reste un défi important.

1. Introduction

1.1. Motivation

Variable renewable energy (VRE) sources such as wind and solar power may represent respectively 42% and 15% of the European electric mix by 2040 [ENTSO-E, 2020]. The shift from the existing grid dominated by conventional thermal power plants to a grid where a large part of the demand may be covered by intermittent sources rises a number of challenges. Contrary to thermal power plants, VRE location is largely restricted by geographical considerations such as the wind speed or the solar irradiance, and may be placed far from the main consumption areas. This is significant in the case of offshore wind power plants, located at dozens or a hundred kilometers from shore. The connection of VRE sources and the power dispatch towards the consumption areas may thus require the development of the transmission grid, in particular to avoid congestion issues [Pesch et al., 2014].

An increase in the bulk power transmission capacities is also seen as enabler for the integration of large amount of renewables [Henry et al., 2014]. The strengthening of the existing transmission grid allows to diminish the overall power system cost [Schlachtberger et al., 2017] and limits the recourse to overcapacity and storage [Schaber et al., 2012]. In Europe, the ENTSO-E predicts an important development of the interconnections between countries, see Figure 1.1. Various scenarios have been proposed to develop the European power network, including point to point upgrades as well as interconnected grids, see Figure 1.2.

While the existing transmission grid is almost entirely based on AC transmission, the growing use of HVDC technology is seen as a promising option to tackle the challenges raised by the transmission grid evolution [Pierri et al., 2017]. HVDC transmission offers, for long transmissions, lower costs than HVAC transmission due to lower losses, see Figure 1.3. The use of DC is almost mandatory for underground or underwater cables of dozens of kilometers due to the compensation of reactive power losses. The interconnection of asynchronous areas such as the UK and continental European grids also requires DC technology. Currently, most of the existing DC lines consist of point-to-point links. The emergence of Multi-Terminal HVDC (MTDC) grids is seen as a promising option to connect a large amount of renewable energy sources as well as distant consumption areas. Two MTDC grids have been commissioned in China, namely the Zhoushan [Pipelzadeh et al., 2015] and Zhangbei [Pang and Wei, 2018] projects. In Europe, the north sea is the subject of much

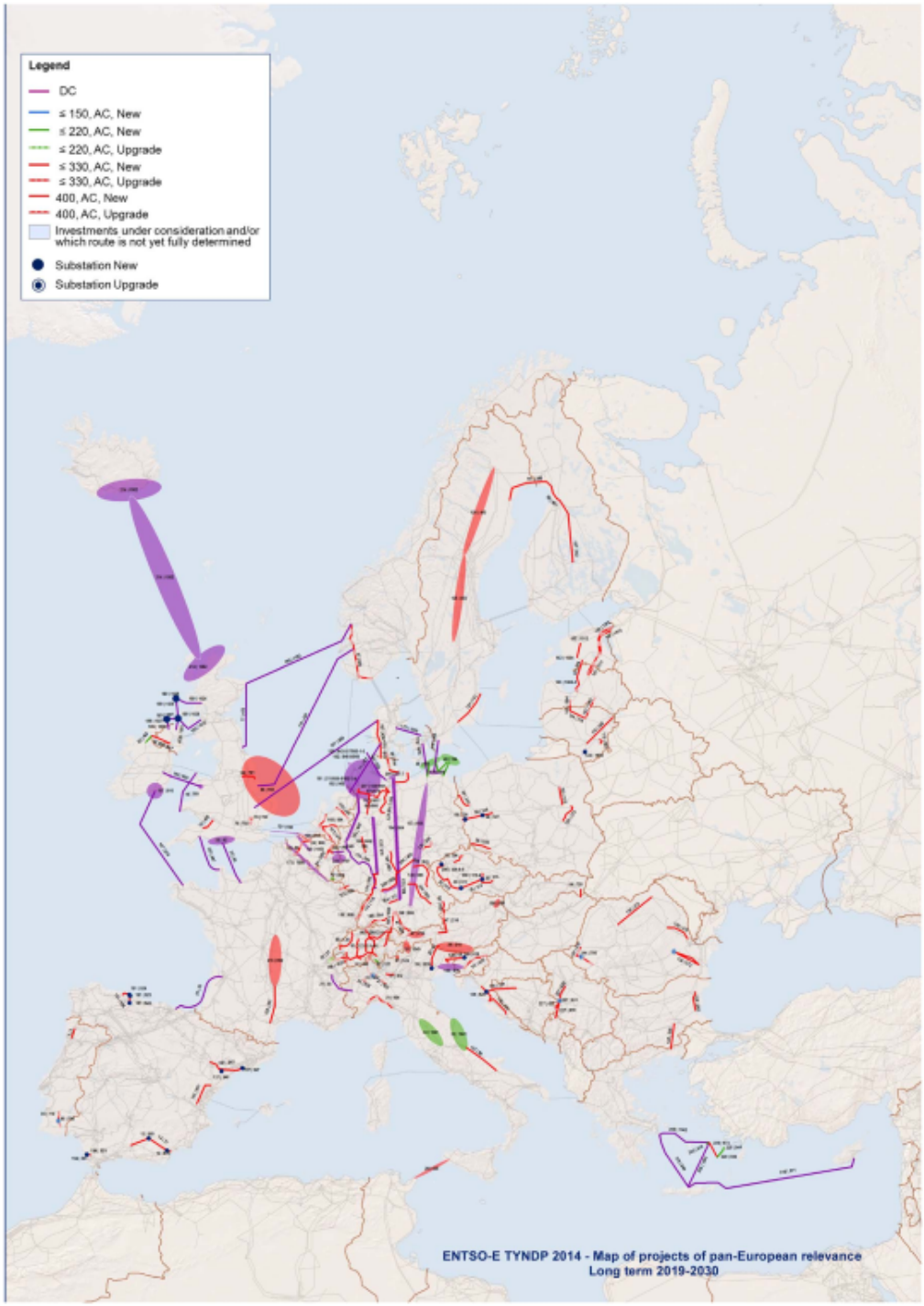
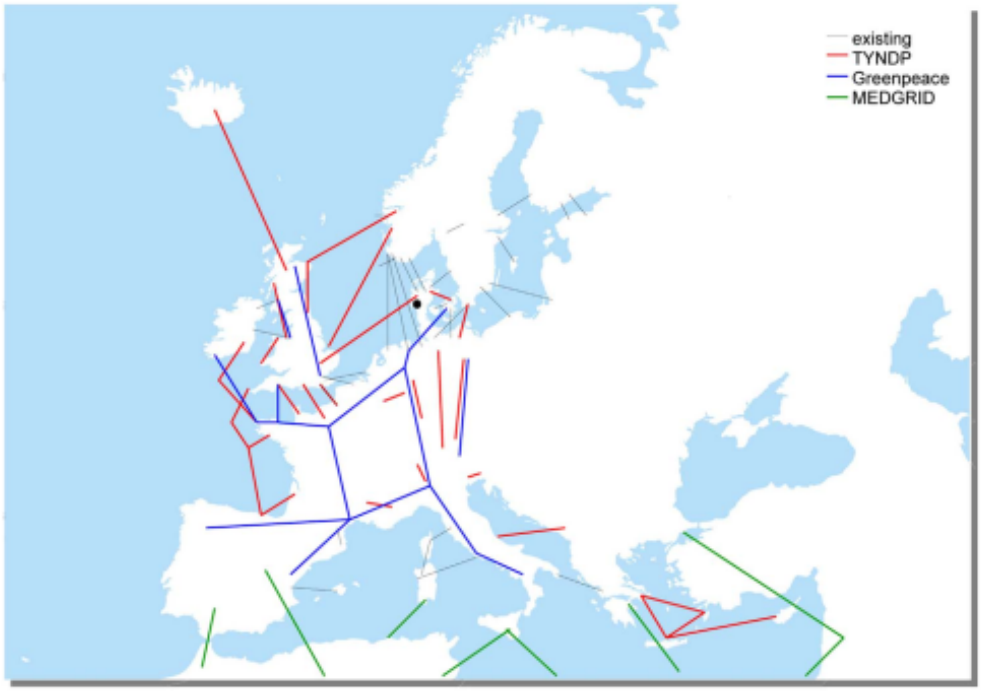


Figure 1.1.: Map of the projects of pan European relevance (2019-2030), from [ENTSO-E, 2020].



(a) Scenario 1: single upgrades

(b) Scenario 2: meshed grids

Figure 1.2.: Evaluation of different models for the development of the European HVDC Supergrid, from [Pierri et al., 2017].

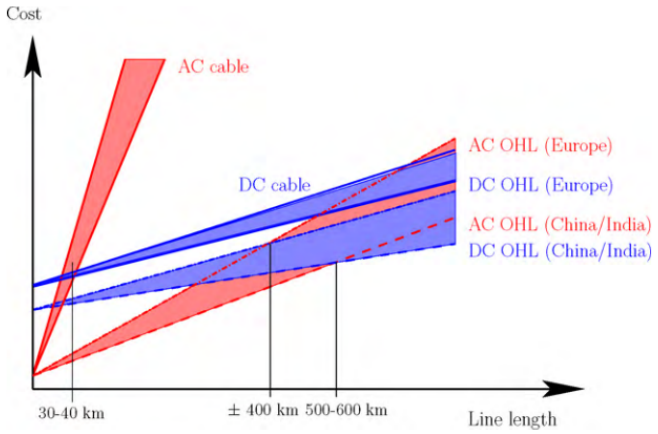


Figure 1.3.: Cost break-down for AC and DC systems for overhead lines (OHL) and cables, from [Van Hertem and Ghandhari, 2010].

attention due to its large wind power capacity and strategic location between countries. HVDC transmission is also considered to improve onshore transmission, such as in the German corridor project [Fairley, 2013]. Despite the difficulty to obtain new right-of-ways for overhead lines (OHL) in Europe, the upgrade of existing AC lines to DC is a promising alternative to increase the power ratings of transmission lines [Reed et al., 2019]. The full integration of HVDC grids into the existing HVAC system, also known as the Supergrid, is thus seen as the future of power transmission [Macilwain, 2010].

1.2. Problem statement and literature overview

1.2.1. The fault identification problem

The development of supergrids still faces a number of non-technical and technical barriers [Zarazua de Rubens and Noel, 2019]. Among those, the control and protection of the grid is seen as particularly challenging [Van Hertem and Ghandhari, 2010]. The protection of power networks refers to the ability to minimize the impact of faults on power systems. It includes the detection and clearing of faults that affect the different components of the grid. In this thesis, we focus on the protection of the transmission lines. The concepts developed for the protection of HVAC grids can not be directly applied to HVDC systems as, for instance, the faulty line current does not cross zero. This makes the disconnection of the faulty line by the line DC circuit breakers (DCCB) a potentially difficult task.
The fault clearing strategy refers to the course of actions that lead to the elimination of the fault and the disconnection of the faulty line. In non-selective strategies, a fault leads to the temporary disconnection of the entire grid, before the faulty line is isolated, see for instance [Torwelle et al., 2020]. Conversely, selective fault

clearing strategies allow the disconnection of the faulty elements and the healthy parts of the grid can pursue to operate. In partially selective approaches, the grid is decomposed into several protection zones that are protected individually. Fully selective strategies are generally preferred as they diminish the impact of fault on both the DC and AC grids [Abedrabbo et al., 2017].
An example of a four station meshed HVDC grid is represented in Figure 1.4 (top). Each line is individually protected by the line breaking modules (BM) located at its extremities. The converter breaker modules, located between the bus-bar and the converter stations, may be involved in the fault clearing strategy as well as in back-up strategies. Each breaking module, see Figure 1.4 (bottom) comprises a DC circuit breaker controlled by a relay which receives measurements from the voltage and current sensors. The relay is responsible for the control of the DCCB and in particular for the sending of the tripping orders when required. Each relay is able to detect a fault or abnormal behavior which occurred in the grid. The assessment by a relay that a fault occurred on the line it protects is referred to as the *fault identification*. Each relay should thus be able to distinguish between the faults that affect the line it protects (or internal faults) and the faults that occur elsewhere in the grid (or external faults). This thesis focuses thus on the fault identification carried out at the line relays. The presence of the BM is thus generally omitted in the grid diagrams and only the location of the line relays is indicated.

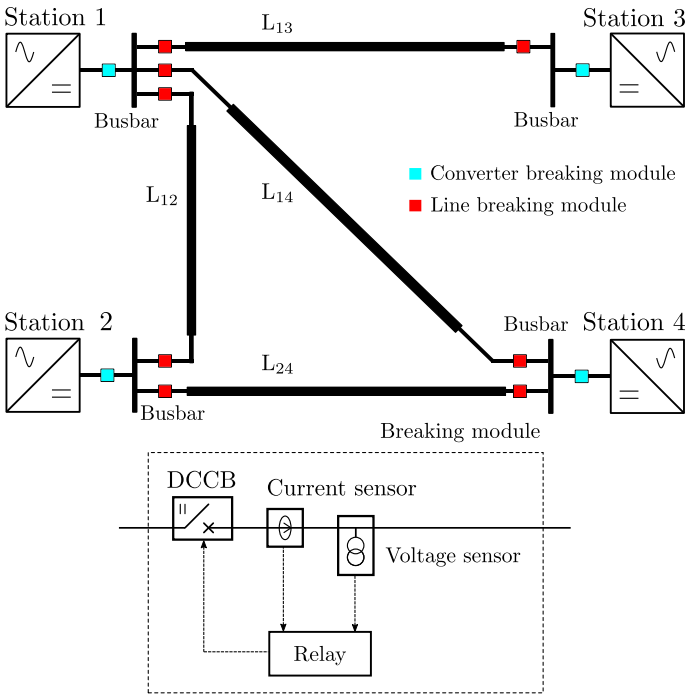


Figure 1.4.: Example of a four station meshed HVDC grid (top). Each line is protected by the two breaking modules (bottom) located at its extremities.

The reliable identification of faults is an ongoing research topic. The important rise of current in the faulty line and the existing DCCB capabilities require to trip the

breakers in few milliseconds. The current rise can be mitigated by placing DC reactors (DCR) at the extremity of each line [Kontos et al., 2014]. Nevertheless, large DCR may have a negative impact on dynamics of the grid [Shinoda et al., 2019], as well as an important cost. In this thesis, the possibility of totally discarding DCR in grid involving only overhead lines (OHL) is investigated. The absence of DCR also makes the fault identification task more difficult as the discontinuity between neighboring lines become less significant. Inversely, the protection of only cable-based grids is not investigated in the manuscript as DCR can not be omitted in this case. The case of grids comprising hybrid lines, *i.e.*, lines with portions of underground cables and overhead lines is nevertheless studied as an extension of the OHL based approach. In hybrid lines, the identification of the faulty segment (underground or aerial) allows to attempt the reclosing of the line as faults in OHL are generally temporary, see for instance [Lewis et al., 2016].

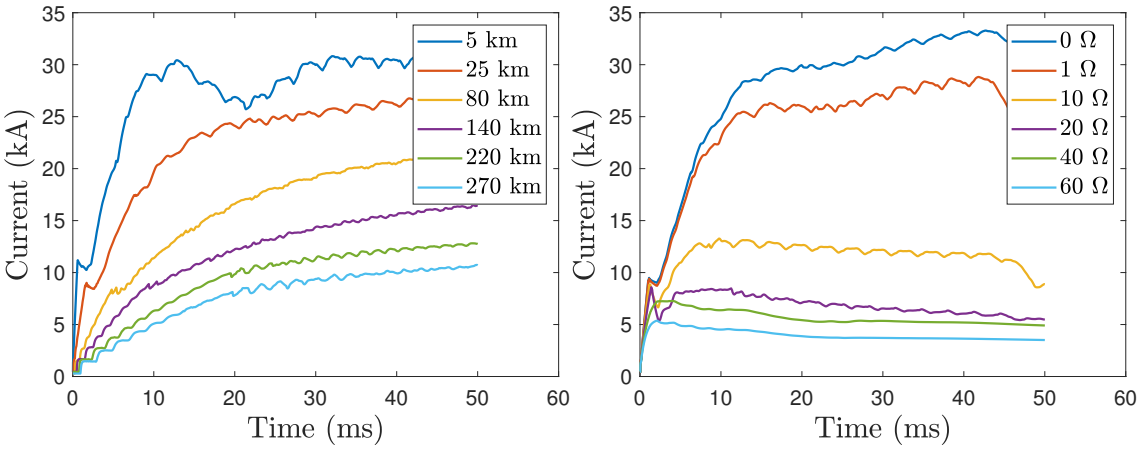


Figure 1.5.: Rise of the current in the absence of DCR for faults occurring at different distances from the breaker with a fault resistance $R_f = 1\,\Omega$ (left) and with different fault resistances for a distance $d_f = 15\,\text{km}$ (right).

An example of the current rise in the absence of DCR due to faults occurring at different distances from the breaker and with different fault resistances is provided in Figure 1.5. Hybrid DCCB with an operating time of 2 ms and breaking capability of 16 kA may be considered [PROMOTIoN WP6, 2016]. In such a case, the breakers must be tripped in few milliseconds, before the line current exceeds the breaker capabilities as observed for various fault cases in Figure 1.5. Double-ended methods, which rely on the communication between distant relays, suffer from communication delays of several milliseconds [Johannesson and Norrga, 2018] and are thus not suited for the fast identification of faults. This thesis focus thus on a single-ended fault identification algorithm able to identify faults using a very short observation window, typically less than 1 ms.

1.2.2. Brief review of fault identification algorithms

This section overviews existing detection, identification, and localization methods, focusing on single-ended algorithms. More details can be found in [Le Blond et al., 2016], [Chang et al., 2017], and [Ashouri et al., 2018].

Fault identification techniques can be classified according to multiple characteristics. Following [Isermann, 1984], we propose to differentiate measurement-based and model-based techniques. As we focus on fault identification, methods developed outside the context of HVDC transmission such as HVAC or micro-grids can also be of interest.

A fault is generally associated with high a variation rate in both current and voltage. Various fault detection and identification methods using thresholds on the voltage and the voltage derivative [Leterme et al., 2016], [Li et al., 2017] or current [Pirooz Azad and Van Hertem, 2017] or a combination of both [Marten et al., 2015], [Auran et al., 2017] have been proposed.

These algorithms share the assumption of the presence of rather large inductances at the end of each line, see Figure 1.4, generally required to limit the rise of current and ensure the reliability of the algorithm. As previously mentioned, such inductances are costly and may degrade the dynamic performance of the system and cause instability [Shinoda et al., 2019].

By contrast, model based approaches try to benefit from previous knowledge on the system or to exploit a model of the grid after the fault occurrence. A fault affecting a transmission line generates traveling waves (TW) that propagate throughout the grid. The TW can be detected using signal processing tools such as the wavelet transform [De Kerf et al., 2011].

Numerous methods are based on the arrival times of the different traveling waves to locate the fault along a line. Among those, but in the context of AC transmission, [Guzman-Casillas et al., 2018] uses multiple traveling wave arrival times to spot the pattern of the reflections between the observation point and the fault, as represented in Figure 1.6. The method is then able to locate a fault on a line inserted in a meshed grid. Several traveling wave arrival times (about a dozen) need to be acquired, which limits the time to react.

In [Xi et al., 2016], the S-transform of the voltage is used to detect the arrival times of the traveling waves. Along with the polarities of the waves, this allows one to locate the fault on a given line.

In those two methods, the key challenges consist in being able to spot accurately the wave fronts, and in being able to distinguish waves due to reflection at the fault from waves due to other reflections in the grid.

For multi-conductor lines, the TW propagate at various speeds which correspond to the speeds of different modes (*e.g.,* ground and aerial modes). The surge arrival time difference between the different modes is analyzed in [Tong et al., 2019]. This time

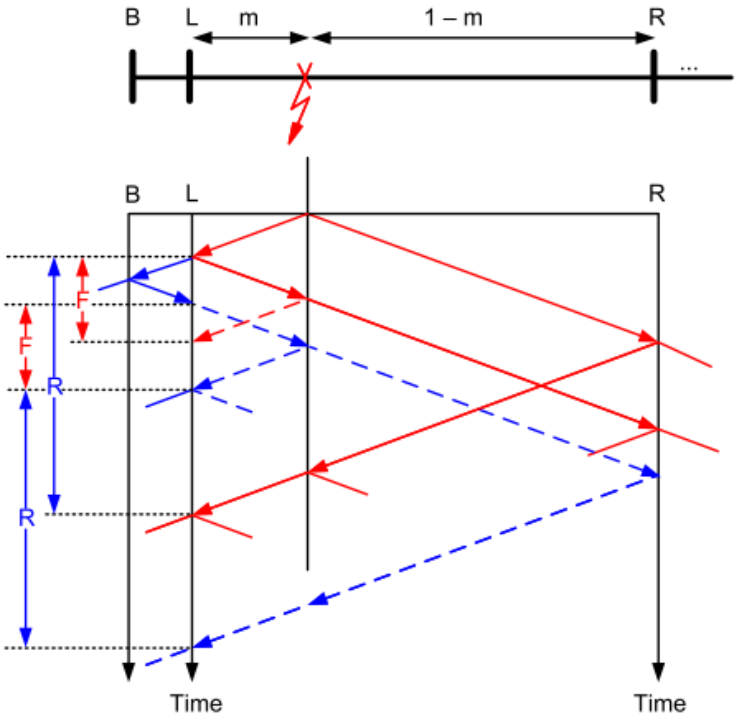


Figure 1.6.: Fault location based on the reflection of multiple traveling waves, from [Guzman-Casillas et al., 2018].

difference depends on the fault distance and is used to identify faults by comparison with the expected maximum time difference for internal faults. Nevertheless, the detection of the arrival times of the different modes may prove difficult, in particular for faults occurring close to a station. The method also requires to distinguish between the arrival of the ground and aerial modes, which may be difficult depending on the station configuration.

The application of electromagnetic time reversal (EMTR) to the fault localization and identification problem has been extensively studied in [Razzaghi et al., 2013] with a focus on the HVDC application in [Razzaghi et al., 2014]. The EMTR approach takes advantage of the time reversal invariance of electromagnetic equations such as the telegraph equations. Based in the recorded voltage at an observation point within the grid, the current waveform at the (supposed) fault location can be reconstructed, using a simplified model of the transmission line. The fault location can then be estimated considering the true fault location maximizes the current energy (in the signal processing sense). As the method is meant as a location algorithm, the behavior towards external faults is yet to be explored. In addition, it requires more than 5 ms long measurement windows.

In the context of DC micro-grid, [Meghwani et al., 2018] proposes to estimate the fault distance and resistance using a simple RL model of the distribution lines. The confidence level of the estimated parameters as well as the value of the estimated

fault distance are used to identify internal faults. The method shows good results on a cable-based micro-grid but the extension to an HVDC grid requires a different approach for modeling long distance transmission lines as the propagation time along such lines can not be neglected. The Bergeron's model is employed along with rational filters in [Suonan et al., 2013] to describe the voltage and current at a fixed set-point along the line. The fault distance with respect to this set-point is computed based on a simplified RL model. As the RL approximation is valid for short lines the estimated fault distance is accurate for a fault close to the set-point but erroneous for faults located elsewhere, see Figure 1.7. This is acceptable as the goal here is to identify whether the fault occurred before or after the set-point. The distance protection principle is then applied: the obtained fault distance is compared with the pre-determined set-point. The approach is able to differentiate faults occurring before and after the set-point but is only tested on point-to-point links. In addition, a measurement window of 15 ms is required which is not compliant with the fast fault clearing of the line.

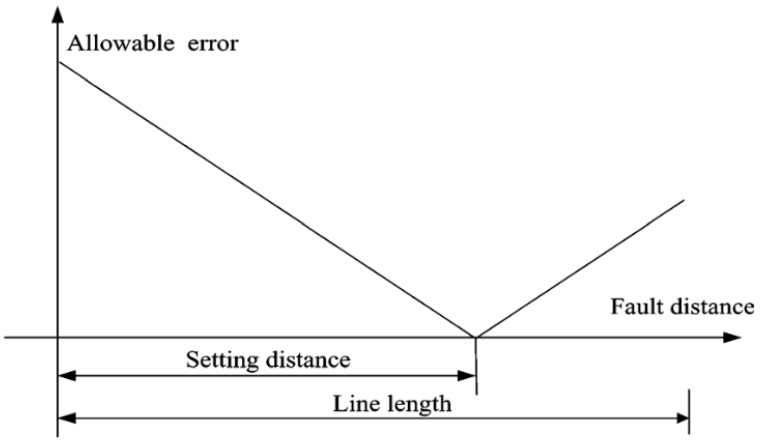


Figure 1.7.: Distance protection principle: the fault is located with respect to a fixed set-point, from [Suonan et al., 2013].

In [Ali Al Hage et al., 2016], a multiple behavioral model-based approach has been developed. Accurate line models are derived for a finite set of possible fault cases. These models are combined in a bank of Kalman filters used to perform the fault identification, see Figure 1.8. Measurements from the relay are then compared to the predictions obtained from the filters. The best predicting filter provides an estimate of the fault characteristics. This technique is able to estimate precisely the fault location but requires considering many filters to be able to identify faults with a large variety of characteristics, see [Ali Al Hage et al., 2016].

The concept of parameter estimation for fault identification has been applied in [Zhang et al., 2020], where a behavioral model of the ground mode of the first TW is used. The fault is identified if one of the estimated parameter exceeds some

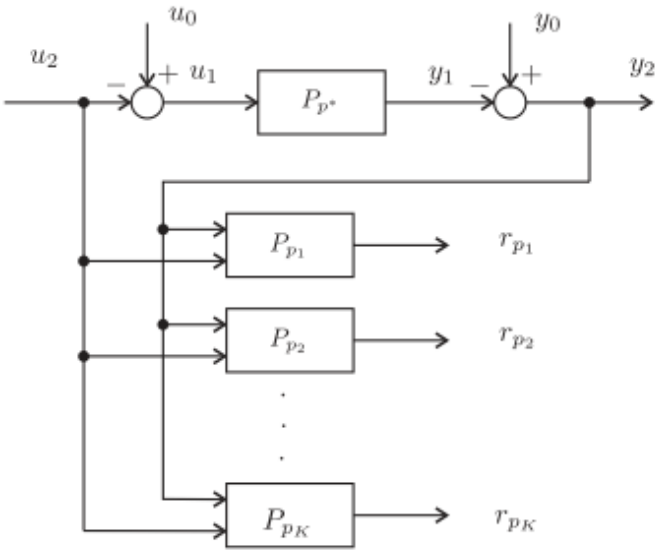


Figure 1.8.: Bank of Kalman filters, each filter describes the evolution of the voltage and current after a specific fault case, from [Ali Al Hage et al., 2016].

predetermined threshold. As the identified parameters by this algorithm do not directly relate to physical quantities, setting the threshold may be difficult. In addition, the modeling of only the first TW makes it difficult to handle faults close to the station, which are particularly important as they lead to a fast rise of the current. This method has been used as a reference comparison for the proposed fault identification algorithm in Section 4.2.2 where more details on the method are provided.

In [Johannesson et al., 2016], the first incident wave of the current is extracted from local measurements in time domain using a rational approximation of the characteristic admittance of the line. The presence of an important incident wave is characteristic of an internal fault. The fault identification is then achieved using a predetermined static threshold on the incident wave. The presence of DC reactors at the end of the lines is still assumed to help distinguishing external disturbances from internal faults.

Model-based algorithms are hence considered for fault location as well as identification. Fault location algorithms, that must determine the precise position of the fault along the line are generally unable to distinguish between internal and external faults. Transient based methods compliant with a fast fault clearing must rely on the transient behavior that prevails in the first milliseconds after the fault occurrence. Many techniques employ the propagation of the traveling waves to get information on the fault location by analyzing the arrival times of the TW at one or several points within the grid. The observation of several waves is then required to locate accurately the fault as well as distinguish external from internal faults.

If one aims at using shorter observation windows for fault identification, for instance less than 1 ms long, all the information contained in the first or second voltage and current waves should be exploited. The accurate description of the current and voltage evolution in this time frame generally involves complex models such as the universal line model [Morched et al., 1999]. Such models, used for instance in [Suonan et al., 2013], [Ali Al Hage et al., 2016], and [Meghwani et al., 2018], involve a fitting of the propagation function and admittance matrix using rational transfer functions and must thus be computed off-line. The obtained model only describes the behavior of a specific fault affecting a line and may not be easily adapted to faults occurring at a different locations or with different fault resistances. By contrast, less accurate models such as RL equivalent are easier to compute but have a poor accuracy. Pure behavioral approaches, as employed in [Zhang et al., 2020], may be a good compromise but they do not relate directly to the physical characteristics of the grid and in particular to the fault parameters. This limits their application to protection purposes as it may be difficult to decide if a fault occurred on the protected line based only on behavioral quantities instead of the fault distance for instance.

The possibility to omit the DC reactors at the end of the transmission lines is today limited as it renders the distinction between internal and external faults more difficult. The reliable identification of faults by the relays in the absence of line reactors can also benefit from a more accurate representation of the TW propagation and interactions at the stations within the grid.

Finding a good trade-off between the accuracy of the model and its simplicity is thus particularly challenging. In this thesis, a combined physical and behavioral model is introduced to represent the first waves after the fault while keeping the computations at a manageable level. The proposed model depends explicitly on the characteristics of the fault and is thus adaptable to various fault cases.

1.3. Thesis contribution and outline

This thesis explores a novel concept for the fast identification of faults based on the estimation of the fault parameters, described in Figure 1.9. A fault affecting a transmission line can be characterized by its fault parameters, namely the fault distance and the fault resistance. A model of the voltage and current transient evolution after a fault is established. This model depends explicitly on the fault parameters. The model is obtained in an off-line stage and is supposed to be embedded in the relay. When a fault is suspected at a relay protecting a transmission line, the fault identification algorithm (FIA) estimates of the unknown fault parameters based on the parametric model and on the received measurements, see the parameter tuning block in Figure 1.9. Based on the results of the parameter estimation, the algorithm is then able to confirm or refute that a fault occurred on the protected line, see the decision logic block.

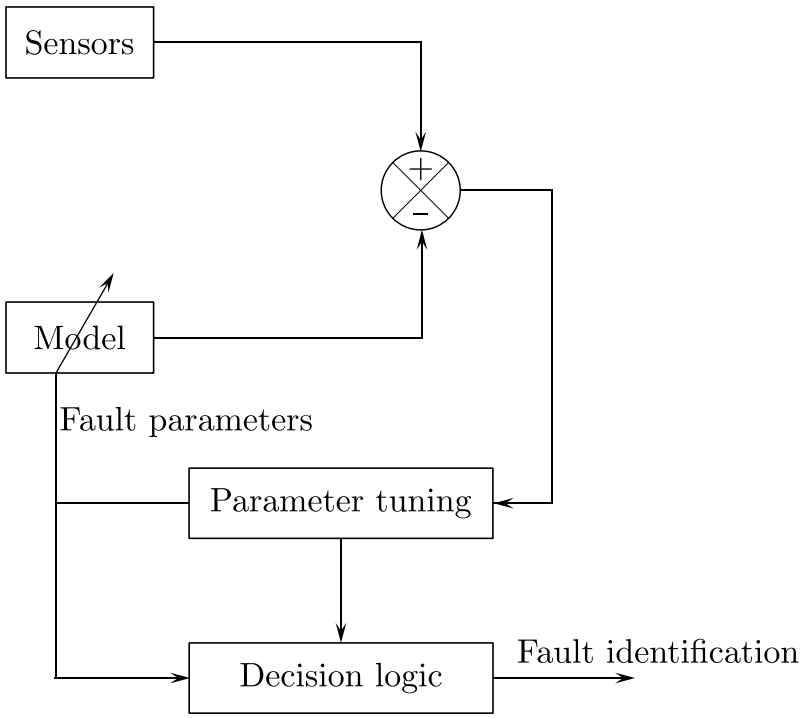


Figure 1.9.: Principle of the fault identification algorithm (FIA) based on the estimation of the fault parameters.

The development of a DC fault model accurate in the first instant (*e.g.* during the first millisecond) after the fault is described in Chapter 2. This model depends on the characteristics of the grid, which are supposed known, and on the unknown vector of the fault parameters $\mathbf{p}$. The model is based on well-known tools such as the TW theory as well as modal analysis for multi-conductor lines. Due to the soil resistivity effects, the accurate representation of the TW propagation involves complex computations ill-suited for fast protection purposes. A combined knowledge-based and behavioral approach is proposed to represent accurately the TW propagation while keeping the computations at a manageable level. In particular, the distortion due to the ground is taken into account by ground filters whose parameters depend explicitly on the fault distance is made explicit. This feature induces that the obtained model, despite comprising a behavioral part, depends explicitly on the fault parameters and is thus adjustable to any fault occurring on a given line. The DC fault model was specifically applied to the asymetric monopole and bipole configurations. For the case of hybrid lines comprising portions of underground cables and overhead lines, a graph representation of the grid is introduced so that a large number of TW are taken into account.

When a fault is suspected at a relay protecting some line e, the algorithm evaluates a maximum likelihood estimate $\widehat{\mathbf{p}}$ of the fault parameter vector $\mathbf{p}$. After each iteration of the estimation algorithm, an approximate confidence region of the estimated parameter vector $\widehat{\mathbf{p}}$ is evaluated. An hypothesis test is then considered involving the

estimate $\hat{\mathbf{p}}$ and its confidence region to determine whether a fault actually occurred in the line e. The hypothesis test may conclude that the fault affects line e, in which case a selective clearing strategy is triggered. When it is unable to conclude, a new batch of measurements is added to update $\hat{\mathbf{p}}$ and the associated confidence region. Once enough measurements have been made available without allowing the FIA to conclude, the fault is deemed to be located elsewhere in the grid. The hypothesis test and the decision logic are detailed in Section 3.

The proposed fault identification algorithm is tested using simulation data from EMT software EMTP-RV [Mahseredjian et al., 2007] to accurately represent the behavior of the grid after a fault. The behavior of the algorithm is studied in three application cases: asymmetric monopole and bipole with OHL as well as asymmetric monopole with hybrid lines. For grids involving only OHL, extensive simulations allowed to test the proposed approach on a broad range of fault cases. The algorithm is able to identify most internal faults using a very short observation window, typically of less than 0.5 ms. The FIA presents good performances compared to a reference method [Zhang et al., 2020], in particular regarding the security against external faults. Dependability or security failures may occur for the faults occurring close to the remote station connected to the protected line.

The system integration of the FIA into a fault clearing strategy is studied in Chapter 5. The estimated fault parameters are employed to assess the severity of the fault, leading to the distinction between severe faults that require the fast tripping of the line breakers and non-severe faults. For the latter, the communication with the other relay protecting the line can improve the robustness of the fault identification. The hardware requirements are also studied, showing the compliance of the proposed strategy with an OHL based grid without line reactors. The impact of the sensor accuracy on the identification algorithm is also investigated.

A general conclusion and some future research perspectives are introduced in Chapter 6. A multiple-hypothesis approach is proposed to increase the robustness of the algorithm with respect to faults close to a station. The problem of identifying faults for which the input voltage is unknown, such as a fault due to a lighting strike, is finally discussed.

The work performed during the thesis resulted in various scientific publications. The application of the algorithm to the mono-conductor case was first published in EPSR 2020[1] and presented at the conferences DPSP 2020[2] and IFAC WC 2020[3]. The more general case of a bipolar configuration has been accepted for publication

[1] Verrax, P., et al. "Transient-based fault identification algorithm using parametric models for meshed HVDC grids." *Electric Power Systems Research* 185 (2020): 106387. DOI: 106387

[2] Verrax, P., et al. "A novel single-ended fault identification algorithm for full selective protection strategy." *15th International Conference on Developments in Power System Protection* (DPSP 2020), Liverpool. DOI: 10.1049/cp.2020.0059.

[3] Verrax, P., et al. " Fault identification in HVDC grids using a transient parametric model.", *21st IFAC World Congress*, Berlin (online) DOI: 10.1049

in IEEE Transactions on Power Delivery[4]. The general method is the subject of a regular patent application[5].

[4]Verrax, P., et al. "Fast fault identification in bipolar HVDC grids: a fault parameter estimation approach." *IEEE Transactions on Power Delivery* (2021). DOI: 10.1109/TPWRD.2021.3056876.

[5]EP3726681A1, "Transient based method for identifying faults in a high / medium voltage electric power transmission system, fault identification module and power transmission system".

2. DC Fault transient modeling

In this chapter, we establish a model that describes the transient evolution of the voltage and the current after a fault occurrence in a transmission line. This model is to be employed to estimate the fault parameters and should thus be easily adjustable to different unknown fault parameters. The voltage and current evolution can be described using the theory of Traveling Waves (TW) which is presented in Section 2.1 in the general case of a multi-conductor transmission line. This approach, after some simplifying assumptions, leads to an analytic physical model of the TW that depends explicitly on the grid characteristics as well as on the fault parameters. Nevertheless, this model neglects the distortion along the line, related to the frequency dependent effect of the soil resistivity. A behavioral approach is proposed in Section 2.2 to represent this effect while keeping an explicit and parametric model of the TWs. The proposed modeling approach for multi-conductor overhead lines is illustrated on a bipolar configuration throughout the Sections 2.1 and 2.2. The particular application to a mono-conductor line is considered in Section 2.3. Finally, the case of hybrid lines comprising portions of overhead lines and underground cables is investigated in Section 2.4. The parameters of the grid in implemented EMT software are detailed in Appendix A for the three considered configurations: mono and multi-conductor OHL as well as underground cables.

2.1. Physical modeling of the TW phenomenon

This section presents a physical approach to model the transient phenomena occurring in a transmission line just after a fault occurrence. The general theory of traveling wave propagation and modal analysis for multi-conductor transmission lines is reminded in Section 2.1.1. The fault and station modeling are presented in Section 2.1.2. Explicit formulas for any wave traveling through the grid are then given in Section 2.1.3. More detailed presentations of the theory of propagation on multi-conductor overhead lines can be found in many references such as [Wedepohl, 1963], or more recently [Franklin, 2009].

2.1.1. The multi-conductor transmission line

In this section, we consider a line e composed of n_c conductors connecting the substations q and q'.

2.1.1.1. The distributed parameters model

The distributed parameters model allow to represent transient phenomena such as the propagation of waves along a transmission line. The voltage and current of all the conductors are gathered in the n_c-dimensional vectors $\mathbf{v}$ and $\mathbf{i}$, whose evolution described by the coupled partial derivative equations

$$\begin{aligned}\frac{\partial \mathbf{v}(t,x)}{\partial x} &= -[R]\mathbf{i}(t,x) - [L]\frac{\partial \mathbf{i}(t,x)}{\partial t} \\ \frac{\partial \mathbf{i}(t,x)}{\partial x} &= -[G]\mathbf{v}(t,x) - [C]\frac{\partial \mathbf{v}(t,x)}{\partial t}\end{aligned} \tag{2.1}$$

The matrices $[R]$, $[L]$, $[G]$ and $[C]$ are the distributed line parameters, respectively the series resistance and inductance and the shunt conductance and capacitance, see Figure 2.1.

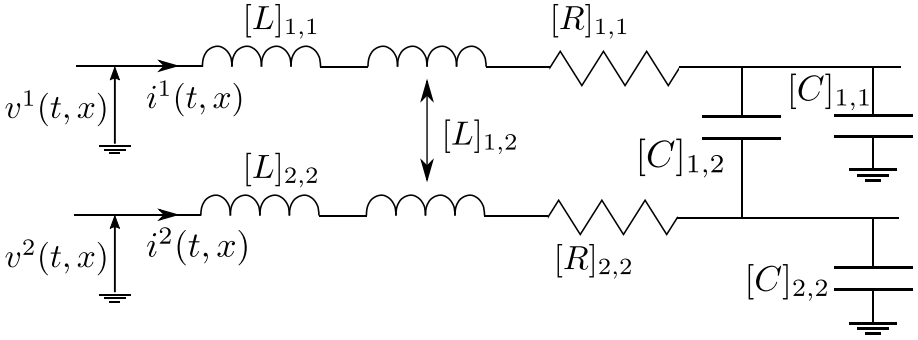


Figure 2.1.: The distributed parameters model on a two conductor transmission line.

Equations (2.1) lead to the Telegraphers equations in time domain

$$\begin{aligned}\frac{\partial \mathbf{v}(t,x)^2}{\partial x^2} &= [R][G]\mathbf{v}(t,x) + ([R][C] + [L][G])\frac{\partial \mathbf{v}(t,x)}{\partial t} + [L][C]\frac{\partial \mathbf{v}(t,x)^2}{\partial t^2} \\ \frac{\partial \mathbf{i}(t,x)^2}{\partial x^2} &= [G][R]\mathbf{i}(t,x) + ([C][R] + [G][L])\frac{\partial \mathbf{i}(t,x)}{\partial t} + [C][L]\frac{\partial \mathbf{i}(t,x)^2}{\partial t^2}\end{aligned} \tag{2.2}$$

which can be expressed in the Laplace domain as

$$\begin{aligned}\frac{\partial^2 \mathbf{V}(s,x)}{\partial x^2} &= [Z(s)][Y(s)]\mathbf{V}(s,x) \\ \frac{\partial^2 \mathbf{I}(s,x)}{\partial x^2} &= [Y(s)][Z(s)]\mathbf{I}(s,x).\end{aligned} \tag{2.3}$$

$[Z(s)] = [R] + s[L]$ and $[Y(s)] = [G] + s[C]$ are the distributed series impedance and shunt admittance matrices, respectively. The line parameters depend on the characteristics of the conductors, the topology of the towers as well as the surrounding environment and in particular the soil resistivity.

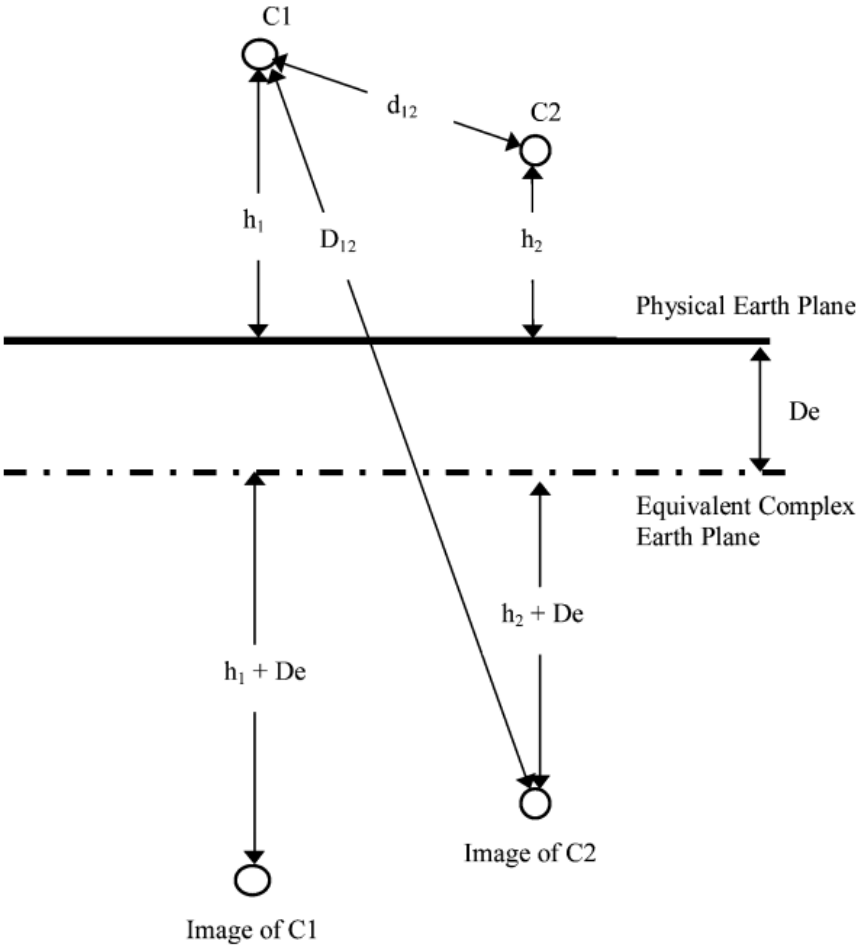


Figure 2.2.: Illustration of the geometry of a multi-conductor transmission line with two conductors [Franklin, 2009].

The general geometry of a multi-conductor tower is represented in Figure 2.2 where the soil resistivity effect is modeled using the complex return plan method. The different parameters introduced are

h_i	the average height of conductor i
r_i	the radius of conductor i
d_{ij}	the distance between conductor i and j
D_{ij}	the distance between conductor i and the image of conductor j with respect to earth

Analytic formulas can be found in [Franklin, 2009] and [Wedepohl and Nguyen, 1996] to compute the value of the distributed parameters. As suggested in [Franklin, 2009],

the sky-wires can be first neglected in the analysis of the distributed line parameters. The capacitance can then be computed through the potential coefficient matrix $[P]$ whose elements are

$$\begin{aligned}[P]_{i,i} &= \frac{1}{2\pi\varepsilon_0}\ln\left(\frac{2h_i}{r_i}\right)\\ [P]_{i,j} &= \frac{1}{2\pi\varepsilon_0}\ln\left(\frac{D_{ij}}{d_{ij}}\right)\end{aligned} \tag{2.4}$$

where ε_0 is the vacuum permittivity.

The capacitance matrix

$$[C] = [P]^{-1}\,. \tag{2.5}$$

depends only on the transmission line geometry. In particular, it does not depend on the soil resistivity and its behavior is constant with frequency.

The series impedance can be divided into an internal term $[Z_{\text{int}}]$ only related to the conductor characteristics and a external term $[Z_{\text{ext}}]$ that takes into account the mutual inductance with the ground return path. The internal impedance for conductor i is a diagonal matrix whose entries are approximated by [Wedepohl and Nguyen, 1996]

$$[Z_{\text{int}}]_{i,i} = \frac{\rho_{\text{c}_i}}{\pi r_i^2}\left(\frac{r_i}{2\delta_{\text{c}_i}}\coth\left(\frac{0.777 r_i}{\delta_{\text{c}_i}}\right) + 0.356\right) \tag{2.6}$$

where $\delta_{\text{c}_i} = \sqrt{\frac{\rho_i}{j\omega\mu_{\text{c}_i}}}$ is the penetration depth of conductor i which depends on the frequency through the skin effect; μ_{c_i} and ρ_{c_i} are the permeability and resistivity of conductor i, respectively.

The external impedance terms are given by

$$\begin{aligned}[Z_{\text{ext}}]_{i,j} &= j\omega\frac{\mu_0}{2\pi}\ln\left(\frac{D_{ij}}{d_{ij}}\right)\\ [Z_{\text{ext}}]_{i,i} &= j\omega\frac{\mu_0}{2\pi}\ln\left(\frac{2\left(h_i + D_{\text{e}}\right)}{r_i}\right)\end{aligned} \tag{2.7}$$

and D_{e} is the depth of the complex ground return plan, see Figure 2.2, of soil resistivity ρ_{g},

$$D_{\text{e}} = \sqrt{\frac{\rho_{\text{g}}}{j\omega\mu_0}} \tag{2.8}$$

The distributed resistance and inductance can then be grouped such that $[Z] = [Z_{\text{int}}]+[Z_{\text{ext}}] = [R]+j\omega\,[L]$. The distributed resistance and inductance matrices have a strong frequency dependent behavior, in particular through the soil resistivity.

2.1.1.2. Solving the Telegraphers equation

In order to solve analytically the Telegraphers equation (2.3), the distributed parameter matrices $[R], [L], [G]$, and $[C]$ are considered at a fixed frequency f_0. The influence of conductor i on conductor j is identical to the influence of conductor j on i, hence the matrices $[Z(s)]$ and $[Y(s)]$ are symmetric, which is evidenced in (2.4) and (2.7). Furthermore we assume the conductors are regularly transposed so that all off-diagonal terms of $[Z]$ and $[Y]$ are equal. The more general case of untransposed transmission lines is discussed in [Wedepohl and Nguyen, 1996]. Thus there exists a unitary matrix (in practice orthogonal) $[T_V]$ such that

$$[Z(s)][Y(s)] = [T_V]^T [D_{ZY}(s)] [T_V] \tag{2.9}$$

where $[D_{ZY}(s)]$ is a diagonal matrix gathering the (real) eigenvalues of $[Z(s)][Y(s)]$. The same transformation matrix diagonalizes $[Y(s)][Z(s)]$ since

$$\begin{aligned}[Y(s)][Z(s)] &= \left([Z(s)]^T [Y(s)]^T\right)^T \\ &= [T_V]^T [D_{ZY}(s)] [T_V].\end{aligned}$$

The modal (or component) voltages and currents are defined as

$$\mathbf{V}^c = [T_V]\mathbf{V} \tag{2.10}$$
$$\mathbf{I}^c = [T_V]\mathbf{I} \tag{2.11}$$

The solution of (2.3) consists of the superposition of two waves traveling in opposite directions. The wave traveling towards the positive x is

$$\mathbf{V}^c(s,x) = \underbrace{\exp(-[\Gamma(s)]x)}_{=[H(s,x)]} \mathbf{V}^c_{\text{init}}(s) \tag{2.12}$$

where

$$\begin{aligned}[\Gamma(s)] &= \sqrt{[D_{ZY}(s)]} \\ &= \sqrt{[T_V][Z(s)][Y(s)][T_V]^T}\end{aligned} \tag{2.13}$$

is the propagation matrix. The initial surge at the fault location $\mathbf{V}^c_{\text{init}}$ computed in modal domain is described in Section 2.1.2.

The elements of $[\Gamma]$ can be divided into real and imaginary parts which account respectively for the attenuation and the propagation of the different modes along the line. The slowest mode, or *ground mode*, is placed in the first entry in $\mathbf{V}^c$ while the *aerial modes*, which travel faster, correspond to the remaining entries. The phase current waves are related to the phase voltage waves through the surge (or characteristic) admittance matrix $[Y_s]$

$$\mathbf{I}(s,x) = [Y_s]\mathbf{V}(s,x) \tag{2.14}$$

where

$$[Y_{\mathrm{s}}] = [Z]^{-1} [T_{\mathrm{V}}]^{T} [\Gamma] [T_{\mathrm{V}}] . \tag{2.15}$$

The surge (or characteristic) impedance $[Z_{\mathrm{s}}]$ is

$$[Z_{\mathrm{s}}] = [T_V]^{T} [\Gamma]^{-1} [T_{\mathrm{V}}] [Z] . \tag{2.16}$$

In modal domain, the surge impedance and admittance matrices are

$$[Z_{\mathrm{s}}^{\mathrm{c}}] = [T_{\mathrm{V}}] [Z_{\mathrm{s}}] [T_{\mathrm{V}}]^{T} \tag{2.17}$$

and

$$[Y_{\mathrm{s}}^{\mathrm{c}}] = [T_{\mathrm{V}}] [Y_{\mathrm{s}}] [T_{\mathrm{V}}]^{T} \tag{2.18}$$

so that modal currents and voltages are related as

$$\mathbf{I}^{\mathrm{c}} (s, x) = [Y_{\mathrm{s}}^{\mathrm{c}}] \mathbf{V}^{\mathrm{c}} (s, x) . \tag{2.19}$$

Furthermore, line parameters in modal domain can be defined as

$$\begin{aligned} [R^{\mathrm{c}}] &= [T_V] [R] [T_V]^{T} \\ [L^{\mathrm{c}}] &= [T_V] [L] [T_V]^{T} \\ [C^{\mathrm{c}}] &= [T_V] [C] [T_V]^{T} . \end{aligned} \tag{2.20}$$

The accuracy of the analytic formulas introduced in Section 2.1.1.1 for the series impedance and admittance as well as their evolution with the frequency are evaluated against the values obtained using an EMT software in the case of a 2-conductor line in Figure 2.3. The distributed parameters are provided over the frequency range $[1 : 10^6]$ Hz in modal domain to evidence the different behavior of the ground and aerial modes. The frequency dependent behavior of the ground mode inductance, related to the soil resistivity ρ through (2.8), is more significant at low frequencies. As expected from the analytic formulas (2.5), the shunt capacitance does not depend on the frequency.

The precision of the analytic formulas for the distributed parameters appears satisfying for both modes and all parameters, except for the series conductance $[G]$ which is generally neglected. The series resistance at high frequency is slightly more difficult to represent due to the skin effect.

Different approximations can be considered to compute the surge impedance and propagation function explicitely. First, the distributed conductance $[G]$ has practically no impact, thus it is assumed $[G] = 0$. As evidenced in Figure 2.3, the losses due to the distributed resistance $[R]$ are low compared to the effect of $[L]$ and $[C]$, in particular at relatively high frequencies. Thus, in a low-loss approach one has

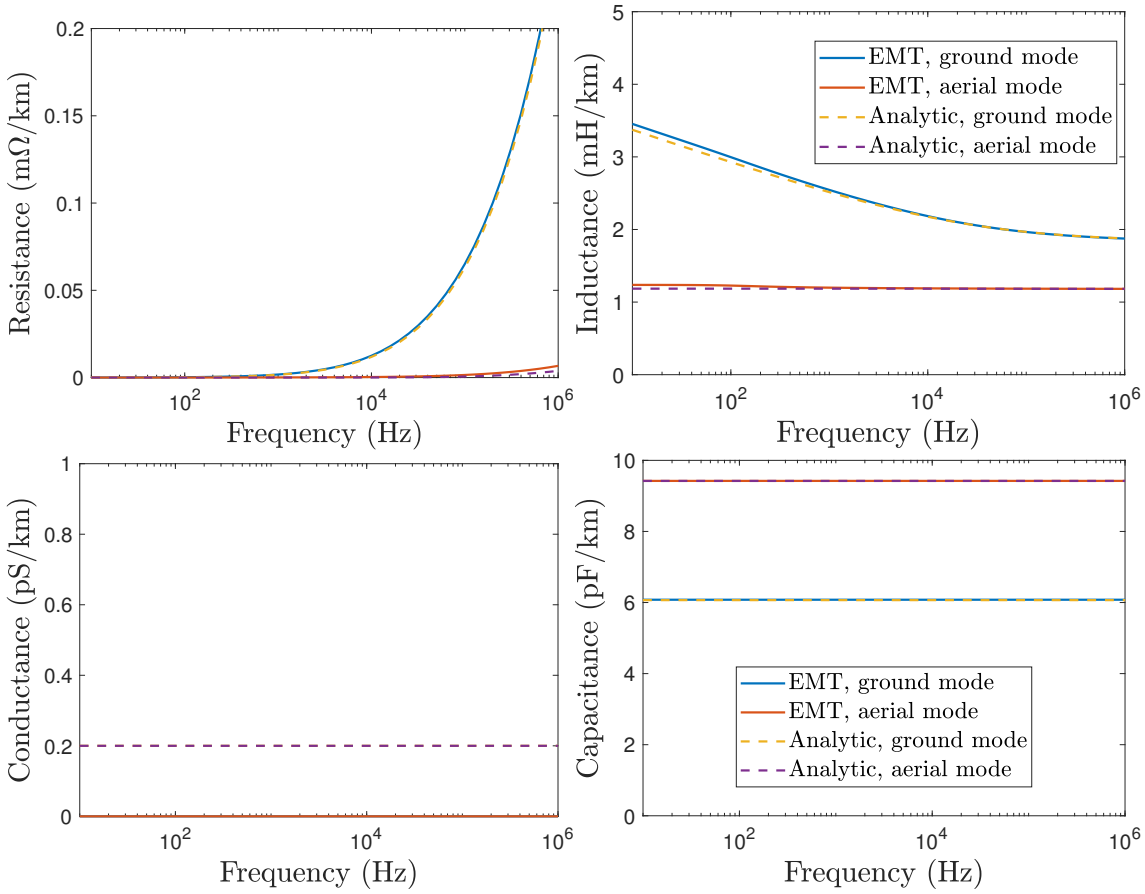


Figure 2.3.: Evolution of the distributed modal parameters analytic expressions and EMT data with the frequency: The series resistance (top-left), inductance (top-right), shunt capacitance (bottom-right), and conductance (bottom-left).

$$\begin{aligned}[\Gamma(s)]^2 &= [T_{\mathrm{V}}]\,[Z(s)]\,[Y(s)]\,[T_{\mathrm{V}}]^T \\ &= [T_{\mathrm{V}}]\,([R]+s\,[L])\,s\,[C]\,[T_{\mathrm{V}}]^T \\ &= [T_{\mathrm{V}}]\left(\frac{1}{s}\,[R]\,[L]^{-1}+I_{n_{\mathrm{c}}}\right)[T_{\mathrm{V}}]^T\,[T_{\mathrm{V}}]\,s^2\,[L]\,[C]\,[T_{\mathrm{V}}]^T\end{aligned} \tag{2.21}$$

Thus,

$$\begin{aligned}[\Gamma(s)]^2 &= [T_{\mathrm{V}}]\left([R]\,[L]^{-1}/s+I_{n_{\mathrm{c}}}\right)[T_{\mathrm{V}}]^T\,[T_{\mathrm{V}}]\,s^2\,[L]\,[C]\,[T_{\mathrm{V}}]^T \\ &= \left(\frac{1}{s}\,[R^{\mathrm{c}}]\,[L^{\mathrm{c}}]^{-1}+I_{n_{\mathrm{c}}}\right)s^2\,[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]\end{aligned}$$

which is composed of two diagonal terms. The square root of $[\Gamma(s)]^2$ can thus be taken as the square root of the diagonal entries. In a low-loss approach, the square root of the first term can be approximated considering that for the i^{th} diagonal entry one has $\frac{1}{s}\frac{[R^{\mathrm{c}}]_{i,i}}{[L^{\mathrm{c}}]_{i,i}} \ll 1$,

$$\begin{aligned}\sqrt{\left(\frac{1}{s}\,[R^{\mathrm{c}}]\,[L^{\mathrm{c}}]^{-1}+I_{n_{\mathrm{c}}}\right)_{i,i}} &= \sqrt{\frac{1}{s}\frac{[R^{\mathrm{c}}]_{i,i}}{[L^{\mathrm{c}}]_{i,i}}+1} \\ &\simeq 1+\frac{1}{2}\frac{[R^{\mathrm{c}}]_{i,i}}{s\,[L^{\mathrm{c}}]_{i,i}}.\end{aligned}$$

Hence,

$$[\Gamma(s)] \simeq \left(I_{n_{\mathrm{c}}}+\frac{1}{2}\frac{[R^{\mathrm{c}}]}{s\,[L^{\mathrm{c}}]}\right)s\sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]}. \tag{2.22}$$

If the series resistance is completely neglected, the loss-less propagation matrix is

$$[\Gamma(s)] \simeq s\sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]} \tag{2.23}$$

which amounts to a pure delay when considering (2.23) in the propagation function $H(s,x)=\exp(-\Gamma x)$ in time domain.

Regarding the surge impedance, one has

$$[Z_{\mathrm{s}}] = [T_V]^T\,[\Gamma]\,[T_{\mathrm{V}}]\,[Y]^{-1}$$

thus

$$\begin{aligned}[Z_{\mathrm{s}}] &\simeq [T_V]^T\left(\left(I_{n_{\mathrm{c}}}+\frac{1}{2}\frac{[R^{\mathrm{c}}]}{s\,[L^{\mathrm{c}}]}\right)s\sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]}\right)[T_{\mathrm{V}}]\,[T_{\mathrm{V}}]^T\,[T_{\mathrm{V}}]\,[sC]^{-1}\,[T_{\mathrm{V}}]^T\,[T_{\mathrm{V}}] \\ &\simeq [T_V]^T\left(\left(I_{n_{\mathrm{c}}}+\frac{1}{2}\frac{[R^{\mathrm{c}}]}{s\,[L^{\mathrm{c}}]}\right)s\sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]}\right)[sC^{\mathrm{c}}]^{-1}\,[T_{\mathrm{V}}]\,,\end{aligned}$$

as the diagonal matrices can commute, low-loss approximation leads to

$$[Z_{\mathrm{s}}] \simeq [T_V]^T \sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]^{-1}} \left(I_{n_{\mathrm{c}}} + \frac{1}{2}\frac{[R^{\mathrm{c}}]}{s\,[L^{\mathrm{c}}]} \right) [T_{\mathrm{V}}]\,. \tag{2.24}$$

In a lossless approach the surge impedance is a real matrix,

$$[Z_{\mathrm{s}}] \simeq [T_V]^T \sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]^{-1}}\,[T_{\mathrm{V}}]\,. \tag{2.25}$$

Similar approximations for the surge impedance in modal domain are directly obtained. In the low-loss case

$$[Z_{\mathrm{s}}^{\mathrm{c}}] \simeq \sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]^{-1}} \left(I_{n_{\mathrm{c}}} + \frac{1}{2}\frac{[R^{\mathrm{c}}]}{s\,[L^{\mathrm{c}}]} \right) \tag{2.26}$$

and in the loss-less case

$$[Z_{\mathrm{s}}^{\mathrm{c}}] \simeq \sqrt{[L^{\mathrm{c}}]\,[C^{\mathrm{c}}]^{-1}}. \tag{2.27}$$

The accuracy of the approximations is extensively studied in the bipolar case in Appendix B.2. The evolution with frequency of the distributed parameters $[R]$, $[L]$ and $[C]$ and of the surge impedance and propagation function is also detailed.. The impact of considering the distributed parameters at fixed frequency instead of accounting for their evolution with the frequency is also evidenced.

2.1.2. Modeling of the interactions at the station and fault location

Consider an incident voltage wave $\mathbf{V}_{\mathrm{f}}$ defined as (2.12) traveling through a line e. When a change of propagation medium occurs (*e.g.*, at the junction between a line and a station), the forward wave $\mathbf{V}_{\mathrm{f}}$ gives rise to a transmitted wave $\mathbf{V}_{\mathrm{t}}$ and a reflected wave $\mathbf{V}_{\mathrm{r}}$,

$$\begin{aligned} \mathbf{V}_{\mathrm{t}} &= \mathbf{V}_{\mathrm{f}} + \mathbf{V}_{\mathrm{r}} \\ &= \left([I_{n_{\mathrm{c}}}] + [K] \right) \mathbf{V}_{\mathrm{f}} \\ &= [T]\,\mathbf{V}_{\mathrm{f}}, \end{aligned} \tag{2.28}$$

where the transmission and reflection matrices $[T]$ and $[K]$ depend on the surge admittances of the media. Consider a wave from medium 1 of surge admittance $[Y_{\mathrm{s},1}]$ propagating towards $n-1$ media of surge admittances $[Y_{\mathrm{s},2}], \ldots, [Y_{\mathrm{s},n}]$. The reflection and transmission matrices are respectively [Wedepohl, 1963]

$$\begin{aligned} [K_{1\to 2\ldots n}] &= \left(\sum_{\ell=1}^{n} [Y_{\mathrm{s},\ell}] \right)^{-1} \left([Y_{\mathrm{s},1}] - \sum_{\ell=2}^{n} [Y_{\mathrm{s},\ell}] \right) \\ [T_{1\to 2\ldots n}] &= \left(\sum_{\ell=1}^{n} [Y_{\mathrm{s},\ell}] \right)^{-1} \left(2\,[Y_{\mathrm{s},1}] \right). \end{aligned} \tag{2.29}$$

Any reflection (resp. transmission) matrix $[K]$ (resp. $[T]$) can be defined in modal domain using the transformation matrix

$$[K^{\mathrm{c}}] = [T_{\mathrm{V}}]\,[K]\,[T_{\mathrm{V}}]^{T}\,. \tag{2.30}$$

For reflections occurring at the junction with a converter station, approximate linear models are considered. The MMC stations can be represented using an RLC equivalent[Leterme and Van Hertem, 2014]

$$\begin{aligned} R_{\mathrm{MMC}} &= n_{\mathrm{SM}} \times r_{\mathrm{SM}} \\ L_{\mathrm{MMC}} &= \frac{2}{3} L_{\mathrm{arm}} \\ C_{\mathrm{MMC}} &= 3C_{\mathrm{arm}} \end{aligned} \tag{2.31}$$

where L_{arm} and C_{arm} are the arm inductance and capacitance, in per-unit. The n_{SM} submodules (SM) per arm are characterized by their individual resistance r_{SM}. For the considered fast transients, the capacitance has a negligible impact and can be neglected. In the situations where the MMC station autoblock for self protection, the model introduced above is no longer valid. In such cases however, the blocking occur after that several waves have been reflected at the station. The representation of the first two or three traveling waves thus does not require to model the MMC stations in blocking mode. More details regarding the derivation of (2.31) are provided in Appendix A.1.

The initial surge at the fault location in equation (2.12) depends on the considered fault. The superposition principle is applied so that the fault condition is analyzed independently from the normal operating behavior, as in [Tang et al., 2018]. A pole-to-ground fault is modeled as a switch closed at $t = t_{\mathrm{f}}$ in series with the fault resistance R_{f} and the opposite pre-fault voltage, $-\mathbf{v}_{\mathrm{bf}}$, see Figure 2.4. The fault resistance is considered constant during the time interval of interest in the order of a millisecond [De Andrade Suárez and Sorrentino, 2010].

Thus, the initial surge can be expressed as

$$\mathbf{V}_{\text{ init}}\left(s, R_{\mathrm{f}}, t_{\mathrm{f}}\right) = -\left(2\left[Y_{\mathrm{s},e}\right] + \left[Y_{\mathrm{s,f}}\right]\right)^{-1}\left(\left[Y_{\mathrm{s,f}}\right]\right)\frac{\exp\left(-t_{\mathrm{f}} s\right)}{s}\mathbf{v}_{\mathrm{bf}}, \tag{2.32}$$

where the $[Y_{\mathrm{s},e}]$ is the surge admittance matrix (2.15) of the faulty line e. The fault surge admittance matrix $[Y_{\mathrm{s,f}}]$ is an $(n_{\mathrm{c}}, n_{\mathrm{c}})$ matrix with non-zero elements in the entries corresponding to the conductor(s) affected by the fault. For a fault between the i^{th} conductor and the ground with fault resistance R_{f}, one has

$$\left[Y_{\mathrm{s,f}}\right]_{i,i} = \frac{1}{R_{\mathrm{f}}}.$$

The initial surge expression (2.32) can be re-written using the reflection coefficient from the line e to the fault $[K_{\mathrm{e}\to\mathrm{f}}(R_{\mathrm{f}})]$,

$$\mathbf{V}_{\text{ init}}\left(s, R_{\mathrm{f}}, t_{\mathrm{f}}\right) = \left[K_{\mathrm{e}\to\mathrm{f}}\left(R_{\mathrm{f}}\right)\right]\frac{\exp\left(-t_{\mathrm{f}} s\right)}{s}\mathbf{v}_{\mathrm{bf}}. \tag{2.33}$$

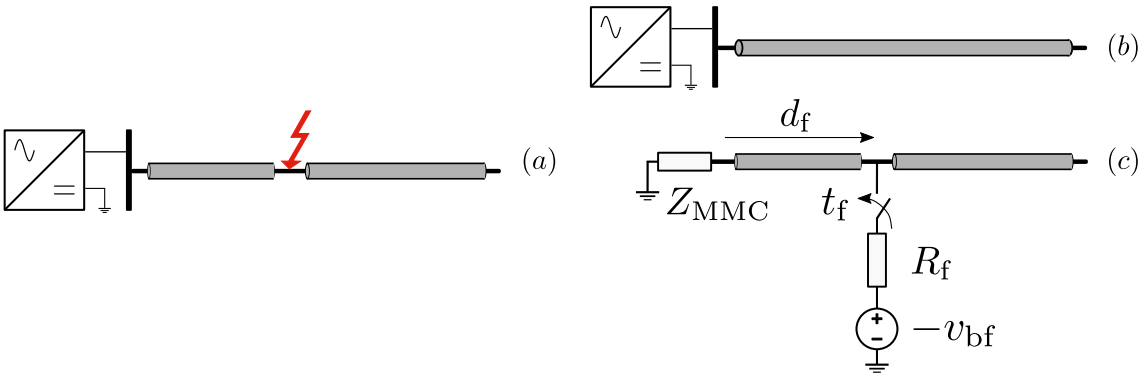


Figure 2.4.: Fault modeling using the superposition principle. The faulty network (a) is represented as the superposition of the healthy network (b) and the faulty line (c), disconnected from any other sources. The faulty line is modeled as an ideal switch closing at $t = t_\mathrm{f}$ on the fault resistance R_f and a voltage source $-v_\mathrm{bf}$ that accounts for the voltage collapse at the fault instant.

The initial surge in modal domain is then

$$\mathbf{V}^\mathrm{c}_\mathrm{init}\left(s, R_\mathrm{f}, t_\mathrm{f}\right) = \left[T_V\right]\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]\frac{\exp\left(-t_\mathrm{f}s\right)}{s}\mathbf{v}_\mathrm{bf} \tag{2.34}$$

which depends on the fault resistance R_f and instant t_f. The voltage at the fault location just before the fault occurrence $\mathbf{v}_\mathrm{bf}$ cannot be measured. In steady state condition however, it can be approximated by the measured pre-fault voltage by the sensor located at one extremity of the line e.

2.1.3. Parametric model of any traveling wave

In Section 2.1.1, we considered the distributed line parameters at a fixed given frequency f_0, allowing us to explicitly solve the telegraphers equation (2.3). We combine this approach to the fault and station modeling presented in Section 2.1.2 to obtain a *physical model* of the different TWs.

The physical model of the first TW starting from the fault location that arrives and is reflected at the substation q after traveling a distance d_f is thus, in modal domain,

$$\begin{aligned}\mathbf{V}^{\mathrm{m,c}}_{\mathrm{f},1}\left(s, \mathbf{p}\right) &= \left[H\left(s, d_\mathrm{f}\right)\right]\mathbf{V}^\mathrm{c}_\mathrm{init}\left(s, R_\mathrm{f}, t_\mathrm{f}\right)\\ \mathbf{V}^{\mathrm{m,c}}_{\mathrm{r},1}\left(s, \mathbf{p}\right) &= \left[K^\mathrm{c}_\mathrm{q}\right]\mathbf{V}^{\mathrm{m,c}}_{\mathrm{f},1}\left(s, \mathbf{p}\right)\end{aligned} \tag{2.35}$$

where $\left[K^\mathrm{c}_\mathrm{q}\right]$ is the reflection matrix from line e to station q. The vector of the fault parameters $\mathbf{p} = \left(R_\mathrm{f}, d_\mathrm{f}, t_\mathrm{f}\right)^T$ gathers the fault resistance, distance and instant. The wave due to the reflection of $\mathbf{V}^{\mathrm{m,c}}_{\mathrm{r},1}$ at the fault location and back towards the station

q is

$$\begin{aligned}\mathbf{V}_{\mathrm{f},2}^{\mathrm{m,c}}(s,\mathbf{p}) &= \left[H(s,d_{\mathrm{f}})\right]\left[K_{\mathrm{e}\to\mathrm{f}}^{\mathrm{c}}(s,R_{\mathrm{f}})\right]\left[H(s,d_{\mathrm{f}})\right]\mathbf{V}_{\mathrm{r},1}^{\mathrm{m,c}}(s,\mathbf{p})\\ \mathbf{V}_{\mathrm{r},2}^{\mathrm{m,c}}(s,\mathbf{p}) &= \left[K_{\mathrm{q}}^{\mathrm{c}}\right]\mathbf{V}_{\mathrm{f},2}^{\mathrm{m,c}}(s,\mathbf{p})\,.\end{aligned} \tag{2.36}$$

The different waves traveling through the grid can be represented using a Bewley lattice diagram, see for instance [Kimbark, 1970]. The first waves are plotted in the case of a point-to-point link between stations q and q' in Figure 2.5. The aerial (in blue) and ground (in brown) modes travel at different speeds along the transmission line but are mixed when reaching a station. This illustrates that the different TW consist of modes with different arrival times. In Figure 2.5, the voltage at station q associated to the first and second TW is, respectively, $\mathbf{V}_1^{\mathrm{m,c}} = \mathbf{V}_{\mathrm{f},1}^{\mathrm{m,c}} + \mathbf{V}_{\mathrm{r},1}^{\mathrm{m,c}}$ and $\mathbf{V}_2^{\mathrm{m,c}} = \mathbf{V}_{\mathrm{f},2}^{\mathrm{m,c}} + \mathbf{V}_{\mathrm{r},2}^{\mathrm{m,c}}$, from (2.35) and (2.36).

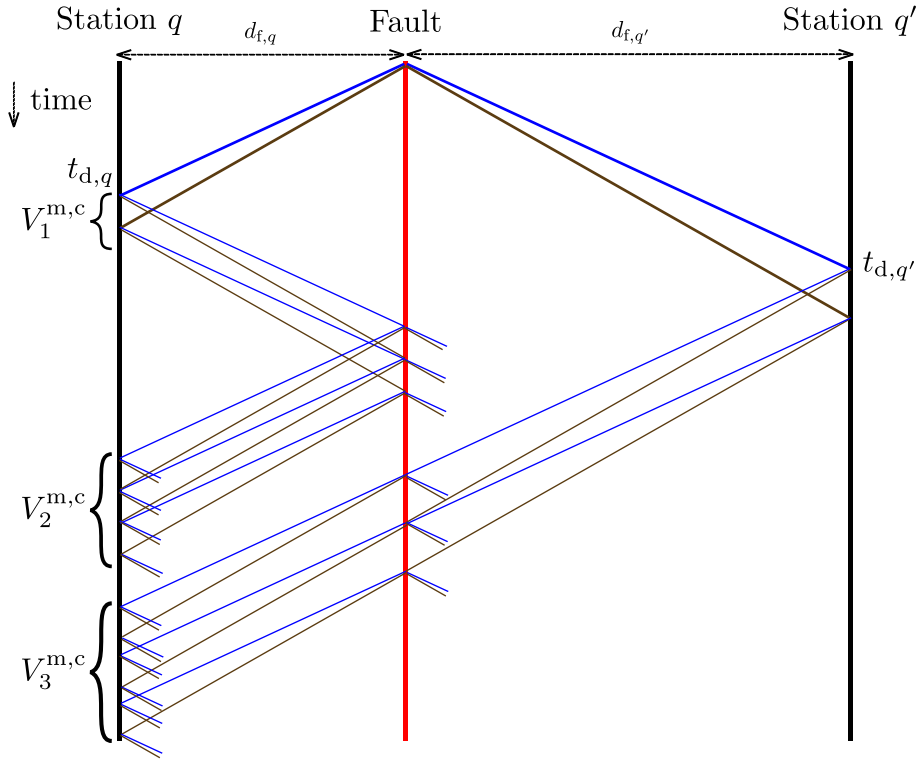


Figure 2.5.: Example of a Bewley lattice diagram for a point-to-point link; The aerial (blue) modes and the ground (brown) mode travel at different speeds along the transmission line.

2.2. Accounting for soil resistivity effects

To obtain the physical model in Section 2.1 the matrices $[R]$, $[L]$, $[G]$, and $[C]$ have been assumed independent of the frequency. In practice, at least $[L]$ is frequency-

dependent as evidenced in Section 2.1.1.1. Taking this dependency into account significantly increases the complexity of the evaluation of the time-domain expressions of the TWs. This section presents a combined physical and behavioral approach to represent the frequency-dependent effects while keeping the evaluations of the expression of the TWs at a manageable level of complexity. The soil resistivity is first assumed to be constant along the whole path of the line in Section 2.2.1. An extension to the case of non-uniform resistivities is introduced in Section 2.2.2.

2.2.1. Uniform soil resistivity

First, one considers a lossless model, in which $[Z(s)] \simeq s[L]$ and $[Y(s)] \simeq s[C]$. This is consistent with the high-frequency components of the transients, leading to negligible series resistance and shunt conductance.

The dependence of $[L]$ with frequency impacts unequally the different modes. While the ground mode endures significant distortion and time delay due to the impact of the soil resistivity on $[L]$, the aerial modes travel with negligible attenuation and at a speed which does not depend on the frequency, as evidenced in [Kimbark, 1970]. This will be illustrated in the case of a 2–conductors OHL. Hence a fixed frequency model is sufficient to represent the propagation of the aerial modes but does not account for the distortion of the ground mode. We thus expresses the propagation matrix $[H(s,x)]$ in (2.12) as the product of a term $[H_{\mathrm{g}}(s,x)]$ accounting for the ground mode distortion only and a term $[H_0(s,x)]$ representing the propagation delays

$$[H(s,x)] = [H_{\mathrm{g}}(s,x)]\,[H_0(s,x)]\,. \tag{2.37}$$

In (2.37), $[H_0(s,x)] = \exp(-[\Gamma_0(s)]\,x)$ and the matrix $[\Gamma_0]$ is computed using the lossless line distributed parameters at a fixed frequency f_0, *i.e.*,

$$[\Gamma_0(s)] = s\sqrt{[T_V]\,[L_0]\,[C_0]\,[T_V]^T}. \tag{2.38}$$

We propose to account for the distortion of the ground mode using a rational transfer function with coefficients $(a_1, \ldots, a_{n_a}, b_0, \ldots, b_{n_b})$

$$[H_{\mathrm{g}}(s,x)] = \begin{pmatrix} \frac{b_0+b_1 s+\cdots+b_{n_b}s^{n_b}}{1+a_1 s+\ldots a_{n_a}s^{n_a}} & 0 & \ldots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & 0 \\ 0 & & 0 & 1 \end{pmatrix}. \tag{2.39}$$

As the distortion due to the soil resistivity depends on the distance traveled by the wave, the coefficients of the ground filter H_{g} are expected to depend on the fault distance d_{f}. The speed of the ground mode

$$c_{\mathrm{g}}(s, d_{\mathrm{f}}) = \| [\Gamma_0(s)]_{1,1} \|^{-1}$$

is also frequency dependent and thus also varies with the traveled distance [Wasserrab and Balzer, 2015].

Consider the vector $\boldsymbol{\theta} = (a_1, \ldots, a_{n_a}, b_0, \ldots, b_{n_b}, c_\mathrm{g})$ gathering the parameters of the filter H_g and the ground mode speed c_g. For a given fault distance d_f, the vector of parameters θ can be identified so that the simulated voltage for the ground mode of the first forward traveling wave $\mathbf{v}_{\mathrm{f},1}^{\mathrm{g}}$ fits the modeled wave $\mathbf{v}_{\mathrm{f},1}^{\mathrm{m,g}}$ by least-squares estimation, *i.e.*,

$$\boldsymbol{\theta} = \arg \min_{\boldsymbol{\theta}} \sum_k \left(v_{\mathrm{f},1}^{\mathrm{g}}(t_k) - v_{\mathrm{f},1}^{\mathrm{m,g}}(t_k, \boldsymbol{\theta}) \right)^2. \tag{2.40}$$

Four different model orders are compared, $(n_\mathrm{a}, n_\mathrm{b}) = (1,0)$, $(n_\mathrm{a}, n_\mathrm{b}) = (1,1)$, $(n_\mathrm{a}, n_\mathrm{b}) = (2,0)$, and $(n_\mathrm{a}, n_\mathrm{b}) = (2,1)$. The output of the four different models are compared with the EMT simulation in Figure 2.6 for two faults located at $d_\mathrm{f} = 70\,\mathrm{km}$ (left) and $d_\mathrm{f} = 220\,\mathrm{km}$ (right) from the point of measure. Both faults have a resistance $R_\mathrm{f} = 0\,\Omega$. The model with $n_\mathrm{a} = 2$, $n_\mathrm{b} = 1$ seems to fit significantly better the EMT data than the three other models, which conversely present a similar accuracy.

The similar waveform of the three models ($n_a = 1, n_b = 0$, $n_a = 1, n_b = 1$, and $n_a = 2, n_b = 0$) may be related to the value of the estimated parameters of the ground filters (2.40). The evolution of the parameters a_1, b_0, a_2, and b_1 with the fault distance is plotted in Figure 2.7 for the four different models. The impact of the parameter b_1 is negligible for the model $n_a = 1, n_b = 1$ as the cut-off frequency of the numerator $\left|\frac{b_0}{b_1}\right| \simeq 10^6\,\mathrm{rad/s}$ is well above the frequency content of the considered signals. Similarly, the highest cut-off frequency of the denominator considering $n_a = 2, n_b = 0$ is about $10^6\,\mathrm{rad/s}$ and the impact of the parameter a_2 is thus negligible compare to a_1. Two metrics are employed to assess the accuracy of the models for different fault distances. First, the value of the cost function at its minimum

$$\min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{k=1}^{N} \left(v_{\mathrm{f},1}^{\mathrm{g}}(t_k) - v_{\mathrm{f},1}^{\mathrm{m,g}}(t_k, \boldsymbol{\theta}) \right)^2 \tag{2.41}$$

is compared for multiple fault parameters. The cost function (2.41) is corrected by the number of sampling points N. For close faults, subsequent waves are quickly generated due to reflections at the fault location. The number of sampling points N is tuned depending on the fault distance such that for all faults, only the first wave is taken into account.

The second employed metrics is the coefficient of determination [Ljung, 1987]

$$R^2 = 1 - \frac{\sum_{k=1}^{N} \left(v_{\mathrm{f},1}^{\mathrm{g}}(t_k) - v_{\mathrm{f},1}^{\mathrm{m,g}}(t_k, \boldsymbol{\theta}) \right)^2}{\sum_{k=1}^{N} \left(v_{\mathrm{f},1}^{\mathrm{g}}(t_k) - \bar{v}_{\mathrm{f},1}^{\mathrm{g}}(t_k) \right)^2}, \tag{2.42}$$

where $\bar{v}_{\mathrm{f},1}^{\mathrm{g}}(t_k)$ is the mean of the measured voltage. The R^2 coefficient can be interpreted as the ratio of the data variation that is explained by the model. R^2

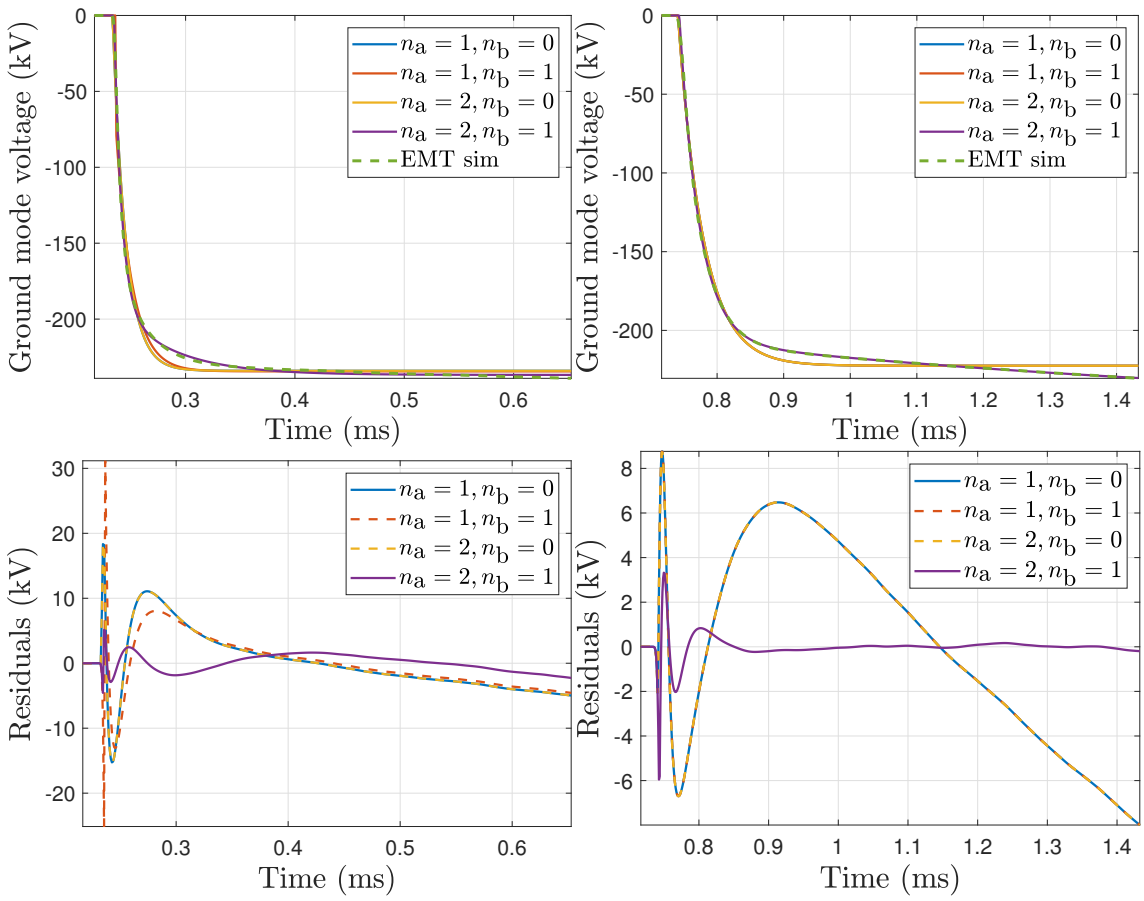


Figure 2.6.: Accuracy of the four different model orders compared to EMT data for two faults located at $d_f = 70\,\text{km}$ (left) and $d_f = 220\,\text{km}$ (right) from the point of measure.

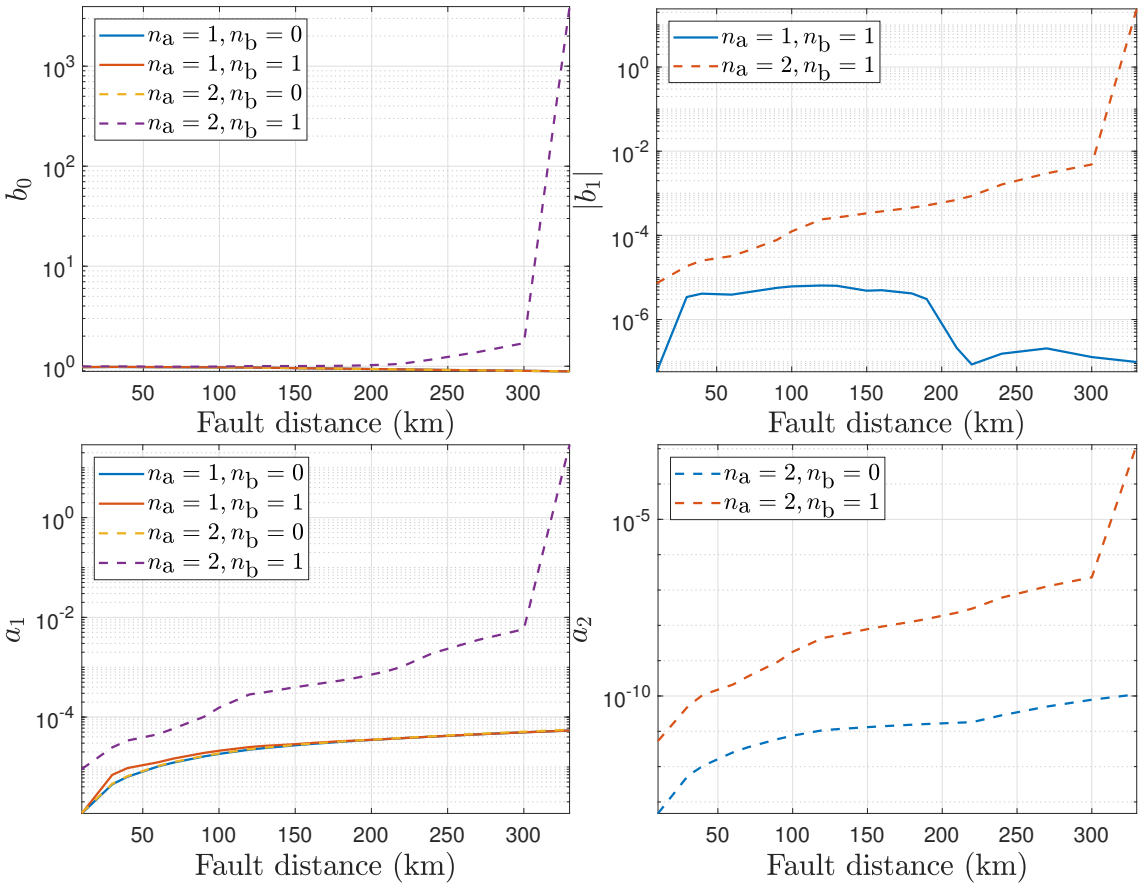


Figure 2.7.: Evolution of the parameters b_0, b_1, a_1, and a_2 with the fault distance for the four different models

gets closer to one as the model is more and more accurate. The values of (2.41) and (2.42) for different fault distances are presented in Figure 2.8. The model with $n_\mathrm{a} = 2, n_\mathrm{b} = 1$ is significantly more accurate, in terms of cost (2.41) and coefficient of determination (2.42). By contrast, the three remaining models present similar performances. This confirms the observation performed in Figure 2.6 on two particular fault cases.

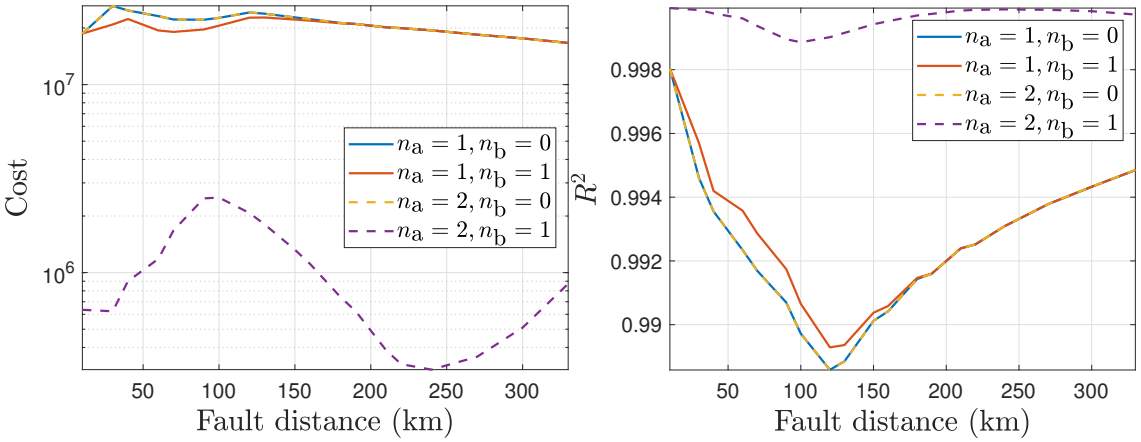


Figure 2.8.: Accuracy of the identified model for different model orders.

Despite its good accuracy, the model $(n_a = 2, n_b = 1)$ is discarded as the complexity of the obtained formulas makes the evaluation of the model too slow for our purposes. We propose instead to use a first-order filter for the ground mode, *i.e.*, $n_a = 1, n_b = 0$ as it allows to capture most of the soil resistivity effects while keeping the complexity at a reasonable level. Multiple fault cases are then simulated with fault distances $\{d_{\mathrm{f},1}, d_{\mathrm{f},2}, \ldots, d_{\mathrm{f},n}\}$. For each fault distance $d_{\mathrm{f},i}$, an estimate $\widehat{\boldsymbol{\theta}}_i$ of $\boldsymbol{\theta}$ is then obtained. The evolution of the parameters a_1, b_0 and c_g with the fault distance is plotted for the bipolar example in Figures 2.9 and 2.10.

The dependence in d_f of the components of $\boldsymbol{\theta}$ is then modeled. The following polynomial model in d_f is proposed for the parameters of the ground mode filter

$$b_0^\mathrm{m}(d_\mathrm{f}) = \beta_0 + \beta_1 d_\mathrm{f} + \beta_2 d_\mathrm{f}^2 \tag{2.43}$$

$$a_1^\mathrm{m}(d_\mathrm{f}) = \alpha_0 + \alpha_1 d_\mathrm{f}. \tag{2.44}$$

The evolution with d_f of the ground mode speed is modeled as

$$c_\mathrm{g}^\mathrm{m}(d_\mathrm{f}) = c_\mathrm{a} \exp\left(\gamma_0 + \gamma_1 d_\mathrm{f} + \gamma_2 d_\mathrm{f}^2\right), \tag{2.45}$$

where c_a is the speed of the aerial mode, in practice equal to the speed of light. It is furthermore imposed that the ground mode speed is always less than the aerial mode speed, *i.e.*,

$$c_\mathrm{g}^\mathrm{m}(d_\mathrm{f}) = \min\left\{c_\mathrm{a}, c_\mathrm{a} \exp\left(\gamma_0 + \gamma_1 d_\mathrm{f} + \gamma_2 d_\mathrm{f}^2\right)\right\}.$$

The parameters $\alpha_i, \beta_i, \gamma_i$ are tuned by least-squares estimation from $d_{\mathrm{f},i}$ and $\widehat{\boldsymbol{\theta}}_i$, $i = 1, \ldots, n$. The results of this approach on a 2-conductor rigid bipole configuration are shown in Figures 2.9 and 2.10. The oscillations of the identified ground mode speed noticeable in Figure 2.10 (left) can be better interpreted considering the ground mode delay with respect to the arrival time of the aerial mode, see Figure 2.10 (right). For small distances, small variations in c_{g} do not affect the arrival time of the ground mode (with the considered sampling frequency $f_{\mathrm{s}} = 1\,\mathrm{MHz}$).

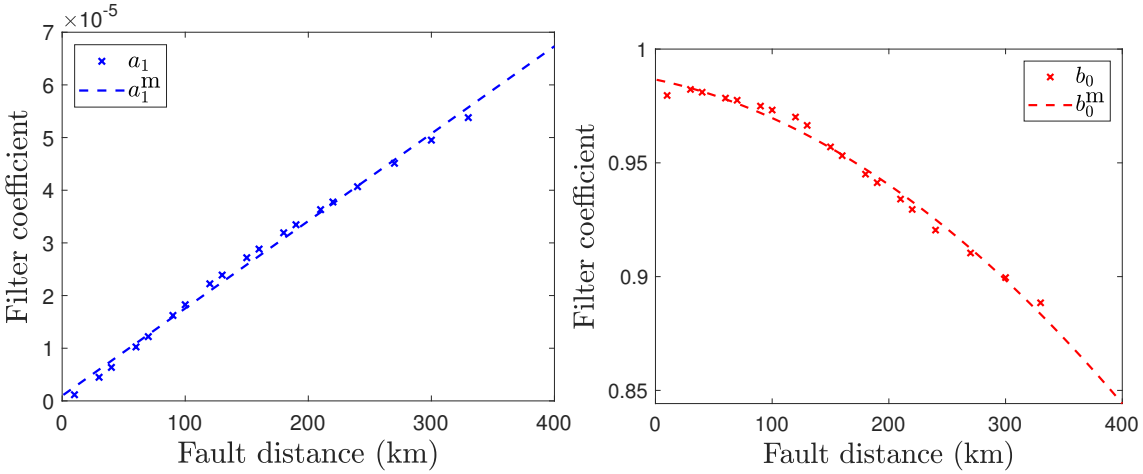


Figure 2.9.: Evolution of the estimated ground filter coefficients $b_0(d_{\mathrm{f},i})$ and $a_1(d_{\mathrm{f},i})$, $i = 1, \ldots, n_{\mathrm{d}}$ and of their models $b_0^{\mathrm{m}}(d_{\mathrm{f}})$ and $a_1^{\mathrm{m}}(d_{\mathrm{f}})$ as a function of the fault distance d_{f}

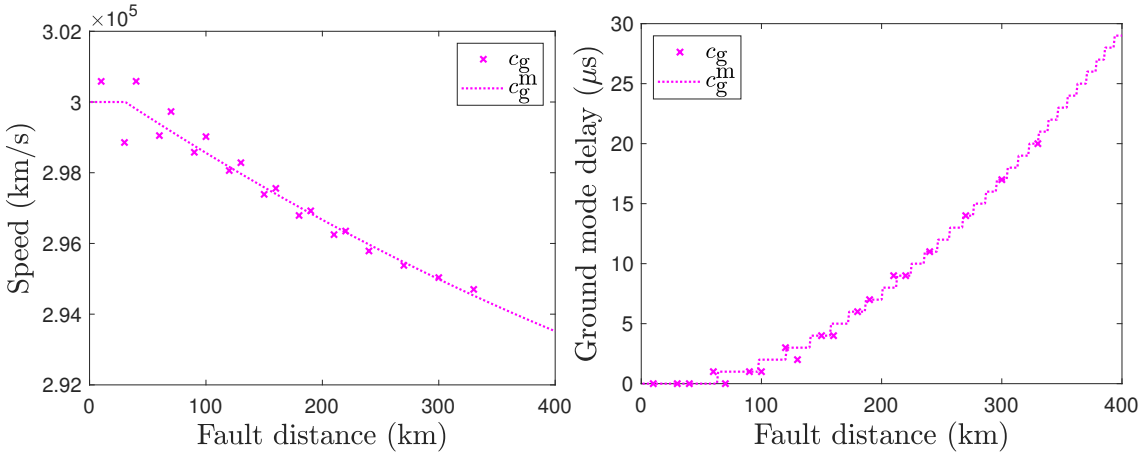


Figure 2.10.: Left: Evolution of the estimates of the ground mode speed $c_{\mathrm{g}}(d_{\mathrm{f},i})$, $i = 1, \ldots, n_{\mathrm{d}}$ and of its model $c_{\mathrm{g}}^{\mathrm{m}}(d_{\mathrm{f}})$ as a function of the fault distance. Right: Delay of the ground mode with respect to the arrival time of the aerial mode.

It can be observed in Figure 2.9 that for low distances, the effect of the ground filter

can be neglected,

$$H_{\mathrm{g}} = \frac{b_0}{1 + a_1 s} \underset{d \to 0}{\longrightarrow} b_0 \simeq 1. \tag{2.46}$$

In particular, if one is interested in modeling only close faults, one may only consider the physical model. A specific close-fault model is employed in the multi-hypothesizes, see Section 6.1.

The explicit expressions of traveling waves expressions obtained in Section 2.1 can be adapted to take into account the ground effects. As the ground filter H_{g} depends explicitly on the fault distance d_{f}, the analytic expressions are adjustable to any fault parameters $\mathbf{p} = (d_{\mathrm{f}}, R_{\mathrm{f}})^{\mathrm{T}}$. As seen in (2.35), the first forward and reflected TWs reaching the station q are expressed in modal domain as

$$\begin{aligned} \mathbf{V}^{\mathrm{m,c}}_{\mathrm{f},1}(s, \mathbf{p}) &= \left[H_{\mathrm{g}}(s, d_{\mathrm{f}})\right] \left[H_0(s, d_{\mathrm{f}})\right] \mathbf{V}^{\mathrm{c}}_{\mathrm{init}}(s, R_{\mathrm{f}}, t_{\mathrm{f}}) \\ \mathbf{V}^{\mathrm{m,c}}_{\mathrm{r},1}(s, \mathbf{p}) &= \left[K^{\mathrm{c}}_{\mathrm{q}}(s)\right] \mathbf{V}^{\mathrm{m,c}}_{\mathrm{f},1}(s, \mathbf{p}). \end{aligned} \tag{2.47}$$

The wave due to the reflection of $\mathbf{V}^{\mathrm{c}}_{\mathrm{r},1}$ at the fault location and back towards the station q is

$$\begin{aligned} \mathbf{V}^{\mathrm{m,c}}_{\mathrm{f},2}(s, \mathbf{p}) &= \left[H_g(s, d_{\mathrm{f}})\right] \left[H_0(s, d_{\mathrm{f}})\right] \left[K^{\mathrm{c}}_{\mathrm{e}\to\mathrm{f}}(R_{\mathrm{f}})\right] \left[H_g(s, d_{\mathrm{f}})\right] \left[H_0(s, d_{\mathrm{f}})\right] \mathbf{V}^{\mathrm{m,c}}_{\mathrm{r},1}(s, \mathbf{p}) \\ \mathbf{V}^{\mathrm{m,c}}_{\mathrm{r},2}(s, \mathbf{p}) &= \left[K^{\mathrm{c}}_{\mathrm{q}}(s)\right] \mathbf{V}^{\mathrm{m,c}}_{\mathrm{f},2}(s, \mathbf{p}). \end{aligned} \tag{2.48}$$

Thus, we obtain a combined physical and behavioral model for the different traveling waves. This model is still parametric and depends explicitly on the fault parameters $\mathbf{p}$. One can get back to temporal domain expressions by applying the inverse Laplace transform,

$$\mathbf{v}^{\mathrm{m,c}}_k(t, \mathbf{p}) = \mathcal{L}^{-1}\left\{\mathbf{V}^{\mathrm{m,c}}_k(s, \mathbf{p})\right\}, k \in \mathbb{N}^* \tag{2.49}$$

Formulas for the current $\mathbf{i}^{\mathrm{m,c}}_k(t, \mathbf{p})$ can be derived from the voltage using the surge admittance matrix (2.14). In the lossless approximation, the surge admittance is a real matrix, and the voltage-current transformation can be applied in the Laplace or time domain indifferently. Phase quantities are obtained from the modal expressions using the inverse transformation matrices, see (2.10) and (2.11). The total voltage and current model is the combination of the different waves that arrive at the considered sub-station

$$\begin{aligned} \mathbf{v}^{\mathrm{m,c}}(t, \mathbf{p}) &= \sum_k \mathbf{v}^{\mathrm{m,c}}_k(t, \mathbf{p}) \\ \mathbf{i}^{\mathrm{m,c}}(t, \mathbf{p}) &= \sum_k \mathbf{i}^{\mathrm{m,c}}_k(t, \mathbf{p}). \end{aligned} \tag{2.50}$$

More or less waves can be included in the model (2.50) depending on the required accuracy of the model. A model comprising more waves would correctly describe

the transient behavior over a larger time interval, at the expense of heavier computations. For most applications, the three waves depicted in Figure 2.5 offer a sufficient accuracy. More waves may be included to describe faults occurring close to a station as many waves get reflected in a short time interval.

The obtained model is now illustrated on a specific fault example and compared with simulation data obtained with EMT software. Consider a positive pole-to-ground fault occurring at $t_f = 0$ on line L_{14} of the network represented in Figure 4.1, located at $d_f = 60\,\text{km}$ from the station 4 with an impedance $R_f = 10\,\Omega$. The model of the two first TW described by (2.47) and (2.48) is compared in Figure 2.11 with the TW obtained from a plain EMT simulation. In Figure 2.11, the voltages are on the left and the currents are on the right sub-figures, while the positive and negative poles are respectively plotted on the top and bottom sub-figures. The model of the two first TW represents quite well the behavior obtained by an EMT simulation during the first 0.8 ms following the fault occurrence. The time interval on which the model is valid thus depends on the number of TW that are described.

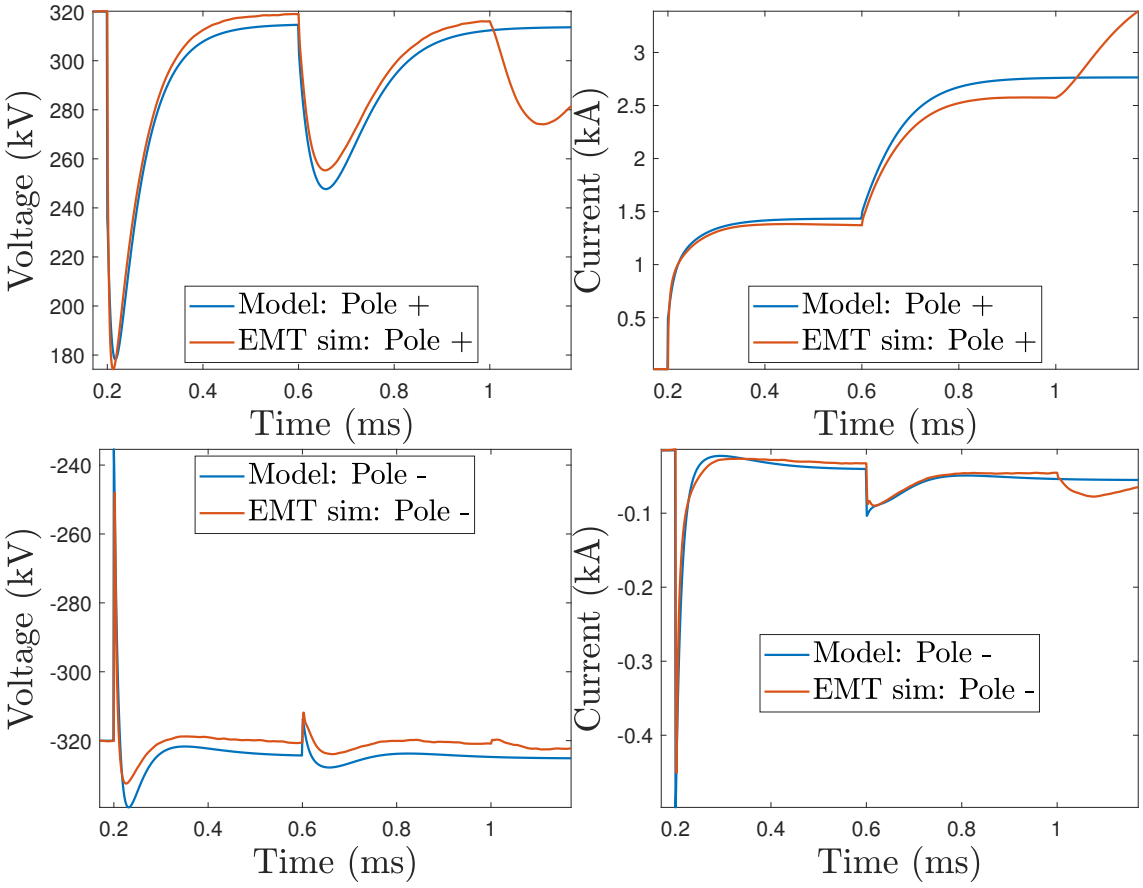


Figure 2.11.: Modeled and simulated phase voltage (left) and current (right) for positive (top) and negative (bottom) poles; The pole-to-ground fault is located at $d_f = 60\,\text{km}$ from the station with an impedance of $R_f = 10\,\Omega$ are displayed.

2.2.2. Non-uniform soil resistivity

The case of a transmission line with a non-uniform soil resistivity is investigated in this section. The influence of the soil resistivity on the parameters of the ground filters is first studied in Section 2.2.2.1. An adaptive formulation of the ground filter is then proposed in the case of a two-portion lines in Section 2.2.2.2.

2.2.2.1. Dependency of the parameters of the ground filters with ρ

In this section we investigate the dependency of the parameters of the soil resistivity filters introduced in Section 2.2.1 with the value of soil resistivity. The same transmission line is considered with different soil resistivities ρ_i and different line lengths $d_{\mathrm{f},j}$. The ground filters

$$H_{\mathrm{g}}\left(s, d_{\mathrm{f}}, \rho\right) = \frac{b_0\left(d_{\mathrm{f}}, \rho\right)}{1 + a_1\left(d_{\mathrm{f}}, \rho\right) s}$$

can be tuned with the procedure described in Section 2.2.1 for each $\rho = \rho_i,\ d_{\mathrm{f}} = d_{\mathrm{f},j}$. The coefficients $b_0\left(d_{\mathrm{f}}, \rho\right)$ and $a_1\left(d_{\mathrm{f}}, \rho\right)$ as well as the speed of the ground mode $c_{\mathrm{g}}\left(d_{\mathrm{f}}, \rho\right)$ depend on the fault distance and the soil resistivity beneath the line. The evolution of $b_0\left(d_{\mathrm{f}}, \rho\right), a_1\left(d_{\mathrm{f}}, \rho\right)$, and $c_{\mathrm{g}}\left(d_{\mathrm{f}}, \rho\right)$ for different fault distances and soil resistivities are plotted in Figures 2.12, 2.13, and 2.14.

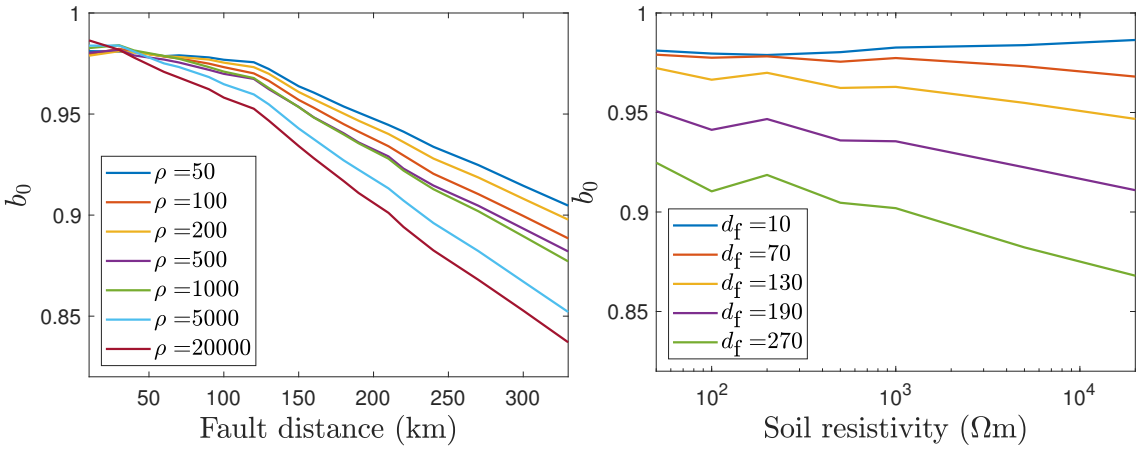


Figure 2.12.: Evolution of the parameter b_0 for several values of the soil resistivity ρ (left) and fault distance d_{f} (right). The parameter seems to depend more on the fault distance than the soil resistivity of the line.

It can be observed that

- The evolution of the three parameters with the fault distance presents a similar shape for various soil resistivities. This motivates the use of the same

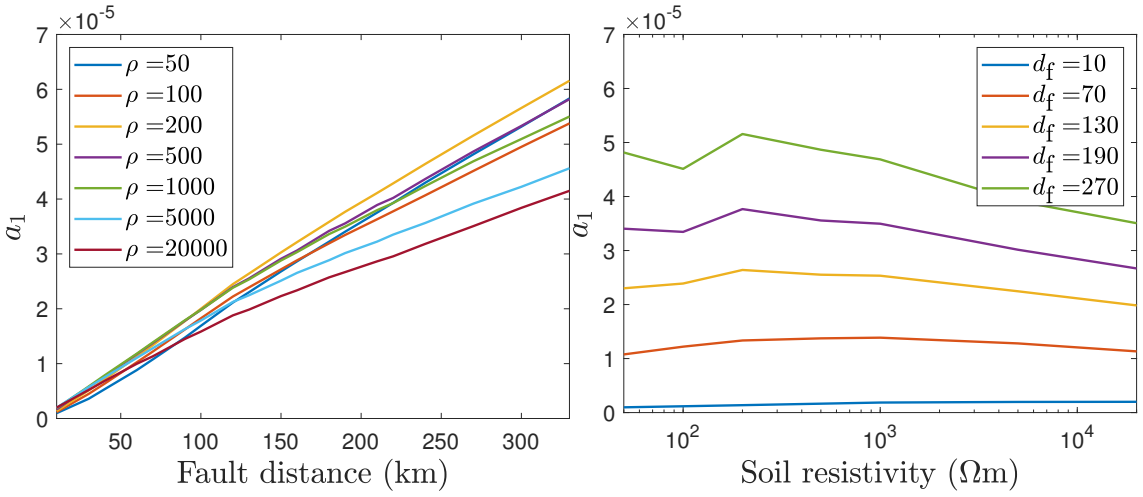


Figure 2.13.: Evolution of the parameter a_1 for several values of the soil resistivity ρ (left) and fault distance d_f (right). The parameter seems to depend more on the fault distance than the soil resistivity of the line.

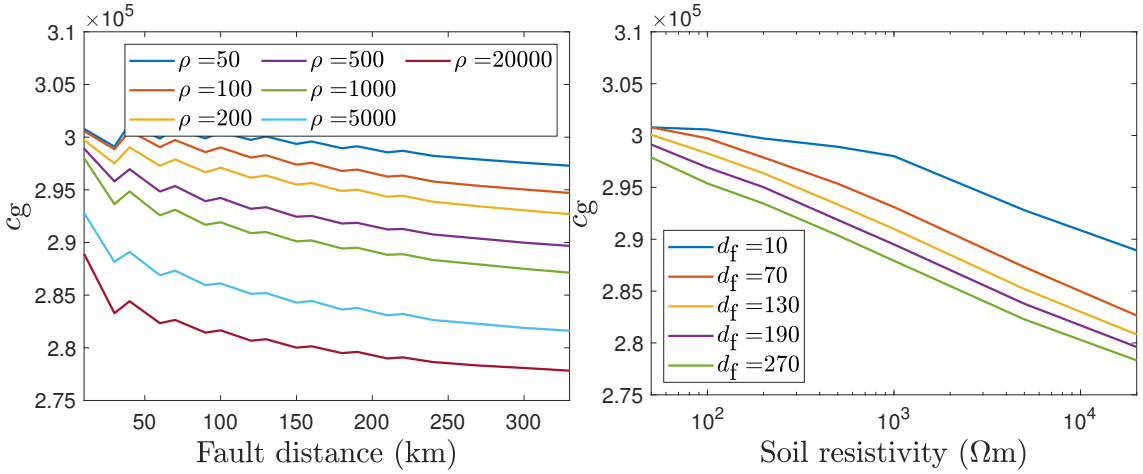


Figure 2.14.: Evolution of the ground mode speed c_g for several values of the soil resistivity ρ (left) and fault distance d_f (right). Except for small values of ρ, the ground speed shows an affine evolution with the logarithm of the soil resistivity.

interpolation to model this evolution, *i.e.*,

$$\begin{aligned} b_0^{\mathrm{m}}(d_{\mathrm{f}},\rho) &= \beta_0(\rho) + \beta_1 d_{\mathrm{f}}(\rho) + \beta_2(\rho)\, d_{\mathrm{f}}^2 \\ a_1^{\mathrm{m}}(d_{\mathrm{f}},\rho) &= \alpha_0(\rho) + \alpha_1(\rho)\, d_{\mathrm{f}}. \\ c_{\mathrm{g}}^{\mathrm{m}}(d_{\mathrm{f}},\rho) &= c_{\mathrm{a}} \exp\left(\gamma_0(\rho) + \gamma_1(\rho)\, d_{\mathrm{f}} + \gamma_2(\rho)\, d_{\mathrm{f}}^2\right), \end{aligned}$$

where the modeling coefficients $\alpha_{0\ldots1}$, $\beta_{0,\ldots,2}$ and $\gamma_{0,\ldots,2}$ are tuned for each specific value $\rho = \rho_i$.

- The parameters $b_0(d_{\mathrm{f}},\rho)$ and $a_1(d_{\mathrm{f}},\rho)$ are more sensitive to the variations of d_{f} than ρ. This is especially true at small fault distances (*e.g.* $d_{\mathrm{f}} < 100\,\mathrm{km}$) where the soil resistivity has a smaller impact on the propagation.
- The speed of the ground mode depends significantly on the soil resistivity, and the evolution seems linear with $\log(\rho)$.

As a consequence, we propose not to model the evolutions of $b_0(d_{\mathrm{f}},\rho)$, $a_1(d_{\mathrm{f}},\rho)$ and $c_{\mathrm{g}}(d_{\mathrm{f}},\rho)$ for small changes of ρ around a given value ρ_0. Such small variations may occur due to environmental conditions such as the humidity or the temperature [Coelho et al., 2015, Afa and Anaele, 2010]. The behavior of the algorithm in such cases where the actual soil resistivity ρ is slightly different from the reference value ρ_0 used in the model is addressed in Section 4.2.3. Large variations of the soil resistivity may nevertheless occur considering the transmission line extends over very different types of grounds with significantly different resistivities. This scenario is investigated in Section 2.2.2.2 where two portions of significant different resistivities are considered.

2.2.2.2. Adaptive ground filter

In this section we consider the case of a transmission line spanning over two portions of lengths d_1 and d_2 with significantly different soil resistivities ρ_1 and ρ_2. The remaining characteristics of the transmission line such as the conductor properties and the pylon geometry are assumed to remain identical along the entire line. The more general case where this property is not satisfied has for instance be studied in [Wedepohl and Nguyen, 1996]. The physical part of the model from Section 2.1 remains identical to the uniform soil case. In particular the transformation matrices to obtain modal quantities are identical along the line. The behavioral approach proposed in Section 2.2.1 to represent the distortion along the line must however be adapted. In the uniform case the propagation function for the ground mode H has been split into

$$[H(s,d_{\mathrm{f}},\rho)] = [H_0(s,d_{\mathrm{f}},\rho)]\,[H_g(s,d_{\mathrm{f}},\rho)] \tag{2.51}$$

where $[H_0]$ represents the propagation delays, and $[H_g]$ the ground filter. As the ground filter only applies to the ground mode, (2.51) can be reduced to a scalar

equation for the ground mode. We model the succession of two portions of different soil resistivity by cascading the propagation functions involved in each section,

$$\begin{aligned} H\left(s, d_{\mathrm{f}}, \rho\right) &= H_0\left(s, d_{\mathrm{f},1}, \rho_1\right) H_g\left(s, d_{\mathrm{f}}, \rho_1\right) H_0\left(s, d_{\mathrm{f},2}, \rho_2\right) H_g\left(s, d_{\mathrm{f}}, \rho_2\right) \\ &= \underbrace{H_0\left(s, d_{\mathrm{f},1}, \rho_1\right) H_0\left(s, d_{\mathrm{f},2}, \rho_2\right)}_{\text{total propagation delay}} \underbrace{H_g\left(s, d_{\mathrm{f}}, \rho_1\right) H_g\left(s, d_{\mathrm{f}}, \rho_2\right)}_{\text{total distortion}} \end{aligned} \tag{2.52}$$

where the two fault distances $d_{\mathrm{f},1}$ and $d_{\mathrm{f},2}$ are defined from the total fault distance d_{f} and the lengths d_1 and d_2 as

$$\begin{cases} d_{\mathrm{f},1} &= d_{\mathrm{f}},\ d_{\mathrm{f},2} = 0 \text{ if fault is in the first portion, } d_{\mathrm{f}} < d_1 \\ d_{\mathrm{f},1} &= d_1,\ d_{\mathrm{f},2} = d_{\mathrm{f}} - d_1 \text{ else.} \end{cases} \tag{2.53}$$

The total propagation delay can be expressed using the ground mode propagation speeds along the two portions, $c^{\mathrm{m}}_{\mathrm{g},1}$ and $c^{\mathrm{m}}_{\mathrm{g},2}$ as follows

$$\begin{aligned} H_0\left(s, d_{\mathrm{f},1}, \rho_1\right) H_0\left(s, d_{\mathrm{f},2}, \rho_2\right) &= \exp\left(-s\frac{d_{\mathrm{f},1}}{c^{\mathrm{m}}_{\mathrm{g},1}}\right) \exp\left(-s\frac{d_{\mathrm{f},2}}{c^{\mathrm{m}}_{\mathrm{g},2}}\right) \\ &= \exp\left(-s\frac{d_{\mathrm{f},1} + d_{\mathrm{f},2}}{c^{\mathrm{m}}_{\mathrm{g,eq}}}\right) \end{aligned} \tag{2.54}$$

where the equivalent propagation speed $c_{\mathrm{g,eq}}$ is

$$c^{\mathrm{m}}_{\mathrm{g,eq}} = \left(d_{\mathrm{f},1} + d_{\mathrm{f},2}\right) \frac{c^{\mathrm{m}}_{\mathrm{g},1} c^{\mathrm{m}}_{\mathrm{g},2}}{d_{\mathrm{f},2} c^{\mathrm{m}}_{\mathrm{g},1} + d_{\mathrm{f},1} c^{\mathrm{m}}_{\mathrm{g},2}}. \tag{2.55}$$

The total distortion (2.52) is

$$H_g\left(s, d_{\mathrm{f},1}, \rho_1\right) H_g\left(s, d_{\mathrm{f},2}, \rho_2\right)$$

$$= \frac{b^{\mathrm{m}}_{0,\rho_1}\left(d_{\mathrm{f},1}\right)}{1 + s a_{1,\rho_1}\left(d_{\mathrm{f},1}\right)} \frac{b^{\mathrm{m}}_{0,\rho_2}\left(d_{\mathrm{f},2}\right)}{1 + s a_{2,\rho_2}\left(d_{\mathrm{f},2}\right)} \tag{2.56}$$

$$= \frac{b^{\mathrm{m}}_{0,\rho_1}\left(d_{\mathrm{f}}\right) b^{\mathrm{m}}_{0,\rho_2}\left(d_{\mathrm{f}}\right)}{1 + s\left(a^{\mathrm{m}}_{1,\rho_1}\left(d_{\mathrm{f},1}\right) + a^{\mathrm{m}}_{2,\rho_2}\left(d_{\mathrm{f},2}\right)\right) + s^2 a^{\mathrm{m}}_{1,\rho_1}\left(d_{\mathrm{f},1}\right) a^{\mathrm{m}}_{2,\rho_2}\left(d_{\mathrm{f},2}\right)} \tag{2.57}$$

which corresponds to the Cascades of two first-order low-pass filters. We approximate the obtained second-order filter by a first-order filter with appropriate cut-off frequency, *i.e.*,

$$H_{g,\rho_1}\left(s, d_{\mathrm{f},1}\right) H_{g,\rho_2}\left(s, d_{\mathrm{f},2}\right) \simeq \underbrace{\frac{b^{\mathrm{m}}_{0,\rho_1}\left(d_{\mathrm{f},1}\right) b^{\mathrm{m}}_{0,\rho_2}\left(d_{\mathrm{f},2}\right)}{1 + s\left(a^{\mathrm{m}}_{1,\rho_1}\left(d_{\mathrm{f},2}\right) + a^{\mathrm{m}}_{2,\rho_2}\left(d_{\mathrm{f},1}\right)\right)}}_{= H_{\mathrm{g,eq}}\left(s, d_{\mathrm{f},1}, d_{\mathrm{f},2}\right)}. \tag{2.58}$$

The main advantage of the proposed approximate approach is that the applied filter remains a first order, which allows to employ the same time domain expressions as

in the homogeneous case. O an adaptation of the value of the parameters $a_1^{\mathrm{m}}, b_0^{\mathrm{m}}$, and the ground mode speed $c_{\mathrm{g}}^{\mathrm{m}}$ is necessary. In addition, the ground mode speed, which depends significantly on the soil resistivity, see Figure 2.14, is adapted to accurately represent the total propagation time along the line using the equivalent speed (2.55).

To facilitate the parameter estimation process described in Chapter 3, the obtained model should be continuous at the junction between the two portions of the line. Consider a fault distance $d_{\mathrm{f}} = d_{\mathrm{f},1} + \varepsilon$ where ε is an arbitrary small positive number. In this case,

$$\begin{aligned} H_{\mathrm{g,eq}}\left(s, d_{\mathrm{f},1}, \varepsilon\right) &= H_{g,\rho_1}\left(s, d_{\mathrm{f}1}\right) \frac{b_{0,\rho_2}^{\mathrm{m}}\left(\varepsilon\right)}{1 + s a_{2,\rho_2}^{\mathrm{m}}\left(\varepsilon\right)} \\ &\underset{\varepsilon \to 0}{\longrightarrow} H_{g,\rho_1}\left(s, d_{\mathrm{f}1}\right), \end{aligned} \tag{2.59}$$

assuming that

$$\begin{cases} b_{0,\rho_2}^{\mathrm{m}} \quad \left(\varepsilon\right) \underset{\varepsilon \to 0}{\longrightarrow} 1 \\ a_{2,\rho_2}^{\mathrm{m}} \quad \left(\varepsilon\right) \underset{\varepsilon \to 0}{\longrightarrow} 0. \end{cases}$$

This assumptions are only approximately satisfied, see Figures 2.12 and 2.13 and the use of the equivalent filter (2.58) causes too abrupt changes in the cost function 3.17, which may cause oscillations in the estimation of the fault parameters. We proposed to use instead

$$H_{\mathrm{g,eq}}\left(s, d_{\mathrm{f},1}, d_{\mathrm{f},2}\right) = \frac{b_{0,\rho_1}^{\mathrm{m}}\left(d_{\mathrm{f},1}\right) b_{0,\rho_2}^{\mathrm{m}}\left(d_{\mathrm{f},2}\right)}{1 + s a_{1,\mathrm{eq}}^{\mathrm{m}}\left(d_{\mathrm{f},1}, d_{\mathrm{f},2}\right)} \tag{2.60}$$

where the equivalent parameter $a_{1,\mathrm{eq}}^{\mathrm{m}}$ is adjusted so that the highest cut-off frequency between $\frac{1}{a_{1,\rho_1}\left(d_{\mathrm{f},1}\right)}$ and $\frac{1}{a_{2,\rho_2}\left(d_{\mathrm{f},2}\right)}$ is neglected, *i.e.*,

$$a_{1,\mathrm{eq}}^{\mathrm{m}} = \max\left\{a_{1,\rho_1}\left(d_{\mathrm{f},1}\right), a_{1,\rho_2}\left(d_{\mathrm{f},2}\right)\right\}. \tag{2.61}$$

The use of (2.61) ensures the continuity of the parameter a_1 at the cost of a less accurate filter.

An example of the fault identification algorithm that uses the proposed adaptive filter is detailed in Section 4.2.1.3. The proposed adaptive filter can be expanded to the case of transmission lines comprising more than two portions of significantly different soil resistivities by considering the smallest cut-off frequency. This approximation may however become less and less accurate as the number of portions to take into account increases.

2.3. Application to a mono-conductor line

2.3.1. Behavioral modeling in the mono-conductor case

In this section, the particular case of a monopolar configuration with a single conductor is addressed. A single conductor OHL can be seen as a special case of a multi-conductor line where the number of conductors is reduced to $n_c = 1$. The derivations in Section 2.1 thus still apply, but the modal transform is omitted. The same methodology as in Section 2.2 can be applied to represent the soil resistivity effects. The matrices are reduced to one-dimensional quantities, which allows to commute the different terms in the TW model, see (2.35) and (2.36). Consequently, the output $y^{m,g} = \left(v^{m,g}(t,\mathbf{p})^T, i^{m,g}(t,\mathbf{p})^T\right)^T$ of the combined model accounting for the effects of the soil resistivity is described in discrete time as

$$y^{m,g}(k,\mathbf{p}) = g(k,d_f) \otimes y^{m,0}(k,\mathbf{p}), \tag{2.62}$$

where $g(k)$ is a low pass filter corresponding to the transfer function $G(z) = \mathcal{Z}(g(k)) = \frac{B(z^{-1})}{A(z^{-1})}$. The output of the *physical model* neglecting the soil resistivity is $y^{m,0}(k,\mathbf{p})$. The ground filters can be tuned for different fault resistances as in Section 2.2.1. A first order filter is selected and the evolution of the filter coefficients a_1, b_0 as well as a pure delay n_d with the fault distance can be explicitly modeled as in (2.43) and (2.44), *i.e.*,

$$\begin{aligned} a_1^m(d_f) &= \alpha_{1,0} + \alpha_{1,1} d_f^{-1} + \alpha_{1,2} d_f^{-2} \\ b_0^m(d_f) &= \beta_{0,0} + \beta_{0,1} d_f^{-1} + \beta_{0,2} d_f^{-2} \\ n_d^m(d_f) &= \text{Round}\left(\nu_0 + \nu_1 d_f^1 + \nu_2 d_f^2\right) \end{aligned} \tag{2.63}$$

The parameters vectors $\boldsymbol{\alpha}_1 = (\alpha_{1,0}, \alpha_{1,1}, \alpha_{1,2})^T$, $\boldsymbol{\beta}_0 = (\beta_{0,0}, \beta_{0,1}, \beta_{0,2})^T$ and $\boldsymbol{\nu} = (\nu_0, \nu_1, \nu_2)^T$ has been adjusted by least-squares estimation. An illustration of the filter coefficients and their models is given in Section 2.3.2.1. For the first traveling wave, a comparison of the output of the complete model with that of the knowledge-based model is provided in Section 2.3.2.2.

The ground filter parameters a_1, b_0 are tuned using the first forward traveling wave after the fault. Nonetheless, it is also required to model the ground effect on the subsequent waves after a reflection at the fault location or at the remote station. To account for the effect of soil resistivity on such waves, the distance d in the parameters $a_1^m(d)$ and $b_0^m(d)$ of the model (2.62) has to represent the sum of the traveled distance. Consider, for instance, the first traveling wave from the fault to the station. This wave is reflected by the station, travels back to the fault and back again to the station after a new reflection. When it reaches again the station for the second time, it has traveled three times the fault distance. The combined

model output $\mathbf{y}^{\mathrm{m,g}}$ is thus obtained by filtering the knowledge-based model output $y^{\mathrm{m},0}$ with a filter g which parameters are $a_1^{\mathrm{m}}(3d_{\mathrm{f}})$ and $b_0^{\mathrm{m}}(3d_{\mathrm{f}})$. An example of this approach is provided in Section 2.3.2.3.

2.3.2. Evaluation of the model accuracy

This section evaluates the accuracy of the complete model of Section 2.3.1 in the case of a mono-conductor line whose detailed parameters can be found in A.3. The parameter fitting approach for the ground filters is illustrated in Section 2.3.2.1. The combined model output is compared against EMT simulation for the first TW in Section 2.3.2.2 and for subsequent TWs in Section 2.3.2.3.

2.3.2.1. Evaluation of the filter parameter fitting approach

Figure 2.15 describes the evolution of the coefficients a_1, b_0 and n_{d} of the behavioral model as estimated for different values of the fault distance d_{f} considering again $\rho = 100\,\Omega\mathrm{m}$. Here again, the estimation has been performed considering only the first traveling wave generated by a fault occurring at time $t_{\mathrm{f}} = 0$. The evolution with d_{f} of $a_1^{\mathrm{m}}(d_{\mathrm{f}})$, $b_0^{\mathrm{m}}(d_{\mathrm{f}})$, and $n_{\mathrm{d}}^{\mathrm{m}}(d_{\mathrm{f}})$ is also provided in Figure 2.15, showing an excellent match with the estimated values of a_1, b_0, and n_{d}. It can be observed that the inductive effect, represented by the parameter a_1, increases with the fault distance: for small values of d_{f}, as expected, a_1 is close to 0 and the physical model neglecting the soil resistivity has good modeling performances.

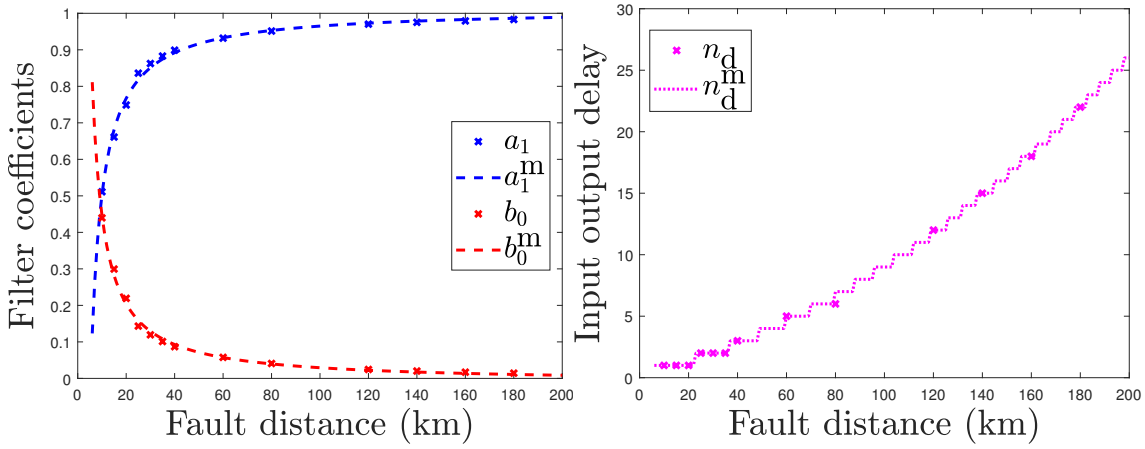


Figure 2.15.: Evolution of the estimated parameters a_1, b_0 (left) and n_{d} (right) of the behavioral model G as a function of d_{f} (crosses), compared to their modeled evolution using $a_1^{\mathrm{m}}(d_{\mathrm{f}})$ and $b_0^{\mathrm{m}}(d_{\mathrm{f}})$ (left) and $n_{\mathrm{d}}^{\mathrm{m}}$ (right) (dashed and dotted lines)

2.3.2.2. First traveling wave

Figure 2.16 represents an example of the evolutions of voltage and current for the first traveling wave generated by a fault occurring on line L_{13} at time $t_f = 0$, with a resistance of $20\,\Omega$ situated 50 km away from the station 3. The soil resistivity is taken as $\rho = 100\,\Omega$m, which is a low or average value according to [IEEE Substations Committee, 2012].

The estimation of the model parameters has been performed considering only the first traveling wave generated by a fault. It can be observed that the outputs of the combined model accounting for the soil resistivity ($\rho > 0$) are much closer to the outputs provided by EMTP-RV than the outputs of the knowledge-based model neglecting the soil resistivity ($\rho = 0$).

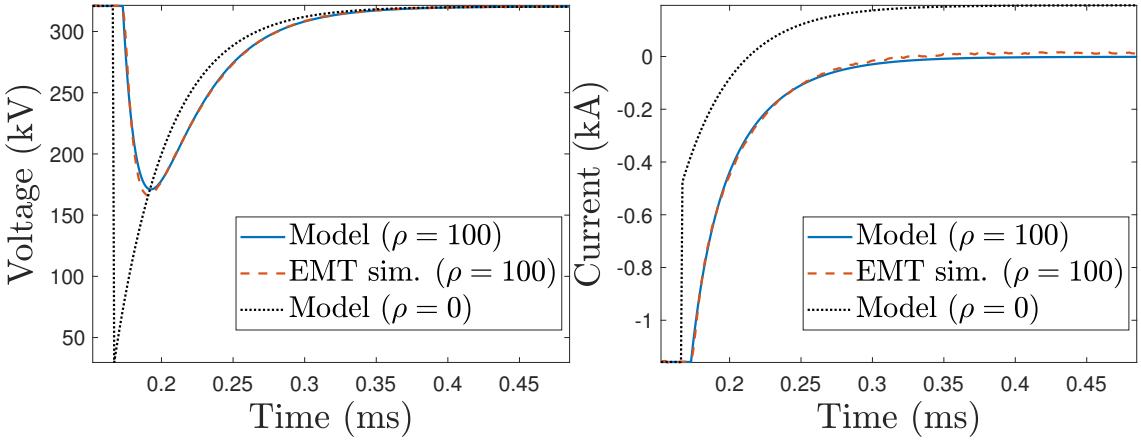


Figure 2.16.: Comparison of the voltage (left) and current (right) transient models, neglecting or accounting for the soil resistivity with the output of an EMT simulation software; the simulated fault occurred at $t_f = 0$ and 50 km from the station with a resistance of $20\,\Omega$; the soil resistivity is $\rho = 100\,\Omega$m.

2.3.2.3. Other traveling waves

As indicated in Section 2.3.1, the effect of soil resistivity on other traveling waves is taken into account by considering the distance d in the parameters $a_1^m(d)$ and $b_0^m(d)$ of the model (2.62) as the cumulative traveled distance .

Figure 2.17 shows the evolution of voltage and current at a station considering a fault occurring at time $t_f = 0$, with a resistance of $20\,\Omega$, and situated 30 km away from the station. One observes that the match with the EMT simulation is very good for the first traveling wave and still good for this wave after two additional reflections (at the station and at the fault).

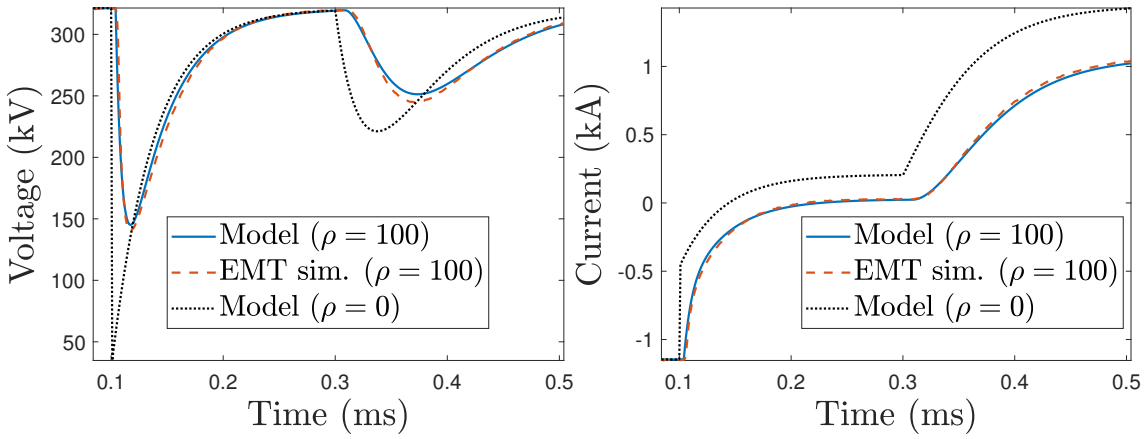


Figure 2.17.: Comparison of the voltage (left) and current (right) transient models, neglecting or accounting for the soil resistivity with the output of an EMT simulation software; the simulated fault is located at 30 km from the station and has a resistance of 20 Ω; the soil resistivity is $\rho = 100\,\Omega$m.

2.4. Extension to the case of hybrid lines

The section details the modeling of faults affecting hybrid transmission lines embedded in a MTDC grid with monopolar configuration. The main notations and the description of the grid as a graph is introduced in Section 2.4.1. The behavioral part of the model that takes into account the effect of the soil resistivity is described in Section 2.4.2. The parameters of the cable used for the EMT simulations are detailed in A.4.

2.4.1. Systematic modeling of traveling waves within a grid

In order to describe exhaustively the traveling waves generated by fault, the protected network is described as an undirected graph $\mathcal{G} = (\mathcal{Q}, \mathcal{E})$. This graph contains vertices $q \in \mathcal{Q}$ representing nodes between two or more line segments. Nodes can be bus-bars or junctions between overhead line and underground cable segments. A segment is represented by an edge $e \in \mathcal{E}$ of the graph. The edge between the nodes q_i and q_j is denoted e_{q_i,q_j}, or $e_{i,j}$ to lighten the notations. Since the graph is undirected, $e_{i,j} = e_{j,i}$. The length of the segment represented by the edge e_{ij} is d_{ij}.

A fault occurring at time t_f implies a modification of the graph $\mathcal{G}$. A node q_f is added to $\mathcal{Q}$ and the faulty edge $e_\mathrm{f} = e_{i,j} \in \mathcal{E}$ is replaced by the edges $e_{i,\mathrm{f}}$ and $e_{j,\mathrm{f}}$ of lengths $d_{\mathrm{f},i}$ and $d_{\mathrm{f},j}$ respectively. Formally, the graph $\mathcal{G}_\mathrm{f} = \left(\mathcal{V}_\mathrm{f}, \mathcal{E}_\mathrm{f}\right)$ accounting for the fault is defined as $\mathcal{E}_\mathrm{f} = \mathcal{E} \backslash \{e_\mathrm{f}\} \cup \left\{e_{i,\mathrm{f}}, e_{j,\mathrm{f}}\right\}$ and $\mathcal{Q}_\mathrm{f} = \mathcal{Q} \cup \{q_\mathrm{f}\}$. The fault can thus be characterized by the vector $\mathbf{p} = (t_\mathrm{f}, d_{\mathrm{f},i}, d_{\mathrm{f},j}, R_\mathrm{f}, e_\mathrm{f})$ of fault parameters, where R_f is the fault resistance between the transmission line and the ground. The two fault distance $(d_{\mathrm{f},i}, d_{\mathrm{f},j})$ are linked through the total length of the line $d_{\mathrm{f},i} + d_{\mathrm{f},j} = d_{i,j}$

which is known. Thus only one unknown fault distance is kept, for a fault located on the edge $e_{i,j}$, the fault distance is arbitrarily defined as

$$d_{\mathrm{f}} = \begin{cases} d_{\mathrm{f},i} \text{ if } i<j \\ d_{\mathrm{f},j} \text{ if } j<i \end{cases} . \tag{2.64}$$

Consider a node $q_{\mathrm{s}} \in \mathcal{Q}_{\mathrm{f}}$ at which voltage and current are observed. This node may, for instance, connect multiple transmission lines to a converter station. We aim to obtain a *physical* model of the traveling waves caused by the fault and that arrive at q_{s}. A TW is entirely determined by its *path*, *i.e.*, the sequence of nodes it has traveled through. Formally, all possible paths from q_{f} to q_{s} can be defined as

$$\mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} = \left\{ (q_{n_1}, .., q_{n_m}) \,|q_{n_1} = q_{\mathrm{f}}, q_{n_m} = q_{\mathrm{s}}, \left(q_{n_i}, q_{n_{i+1}}\right) \in \mathcal{E}_{\mathrm{f}}, m > 1 \right\} \tag{2.65}$$

A path $\pi \in \mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}$ can include several times the same node, including the faulty node q_{f} and the observation node q_{s}. Due to the reflections occurring at the junctions, a TW is indeed likely to pass several times via the same nodes. When a wave travels on an edge, considering the lossless approximation and constant distributed parameters, as in Section 2.1.1.2, only the propagation delay has to be taken into account. Consequently, when modeling traveling waves, one has to account for

- the different delays due to propagation along the edges,
- the effect of junctions on the incident wave.

Consider a path $\pi = (q_{n_1}, .., q_{n_m}) \in \mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}$ traveled by a given wave. The total propagation delay through the edges along π is

$$\tau_\pi(d_{\mathrm{f}}) = \sum_{i=1}^{m} \Delta t_{n_i,n_{i+1}} = \sum_{i=1}^{m-1} \frac{d_{n_i,n_{i+1}}}{c_{n_i,n_{i+1}}} \tag{2.66}$$

where $c_{n_i,n_{i+1}}$ is the wave propagation speed along the edge $e_{n_i,n_{i+1}} = \left(q_{n_i}, q_{n_{i+1}}\right)$, which depends on the propagation medium. The total delay thus depends on the fault distances $d_{\mathrm{f},i}$ or $d_{\mathrm{f},j}$ as at least the first traversed edge is necessarily connected to the fault.

At each junction along π, the voltage wave is subject to a transmission or reflection coefficient, depending on the propagation direction before and after the junction. The impact of reflections and transmissions at junctions is described by

$$V_{\pi,\mathrm{j}}(s, t_{\mathrm{f}}, R_{\mathrm{f}}) = \prod_{i=1}^{m} J_{e_{n_{i-1},n_i} \to q_{n_i}}(s, R_{\mathrm{f}}) \frac{\exp(-t_{\mathrm{f}} s)}{s} V_{\mathrm{bf}} \tag{2.67}$$

where the coefficient J is either a reflection or transmission (2.29) coefficient

$$J_{e_{n_{i-1},n_i} \to q_{n_i}} = \begin{cases} T_{e_{n_{i-1},n_i} \to q_{n_i}} \text{ if } n_{i-1} \neq n_{i+1} \\ K_{e_{n_{i-1},n_i} \leftarrow q_{n_i}} \text{ if } n_{i-1} = n_{i+1} \end{cases}$$

for $i = 2, \ldots, m-1$. The first term in the product (2.67) accounts for the initial surge at the fault location $J_{e_{n_0,n_i} \to q_{n_1}} = K_{q_{n_1},q_{n_2},\leftarrow q_{\mathrm{f}}}(R_{\mathrm{f}})$ and thus depends on

the fault resistance, see (2.32). The voltage at node q_s due to the arrival of an incident wave is obtained as the transmitted wave to the node q_s (2.28). This transmission coefficient is thus included as the last term in the product (2.67), hence $J_{e_{n_{m-1},n_m}\to q_m} = T_{e_{n_{m-1},n_m}\to q_\mathrm{s}}$. Consequently, for a given path π, considering the propagation delay (2.66) and the transmissions and reflections occurring along π via (2.67), one introduces the following *physical model*

$$V_\pi^0(s,\mathbf{p}) = \exp\left(-\tau_\pi(d_\mathrm{f})s\right) V_{\pi,\mathrm{j}}(s, t_\mathrm{f}R_\mathrm{f}) \tag{2.68}$$

of the wave.

The different paths taken by the TWs can be represented via a Bewley lattice diagram. Figure 2.18 illustrates such diagram on a point-to-point link consisting of an OHL and a cable segment. Even in this relatively simple case, the presence of the OHL-cable junction creates a large number of reflected TWs. Note also the propagation speed of the TWs in the underground part is slower than in the overhead line.

Contrary the case of overhead lines where the inverse Laplace transform (2.49) is used,the waveform in the time domain is here obtained through the inverse Fourier transform.

$$\begin{aligned} v_\pi^0(t,\mathbf{p}) &= \mathcal{F}^{-1}\left\{ V_\pi^0(s,\mathbf{p})\Big|_{s=j\omega} \right\} \\ &= \mathcal{F}^{-1}\left\{V_{\pi,\mathrm{j}}(\omega,R_\mathrm{f})\right\} \otimes \delta_{\tau_\pi(d_\mathrm{f})}(t) \\ &= v_{\pi,\mathrm{j}}(t,R_\mathrm{f}) \otimes \delta_{\tau_\pi}(d_\mathrm{f},t) \\ &= v_{\pi,\mathrm{j}}(t-\tau_\pi(d_\mathrm{f}),R_\mathrm{f}) \end{aligned} \tag{2.69}$$

where $\delta_{\tau_\pi(d_\mathrm{f})}(t) = \delta_0(t-\tau_\pi(d_\mathrm{f}))$ is the Dirac distribution corresponding to the propagation delay τ_π along the path. In practice, $\mathcal{F}^{-1}$ is computed numerically using the inverse discrete Fourier transform. Considering a sampling period T_s, the obtained discrete-time voltage model at time kT_s is thus written as $v_\pi^0(k,\mathbf{p})$.

2.4.2. Behavioral modeling of the current return path

The physical model developed in Sections 2.1 assumed the distributed line parameters are independent of the frequency. With this approximation, the distortions of the waves due to the soil resistivity (for overhead sections) and the screen resistance (for underground sections) cannot be described.

To account for such effects a behavioral approach is proposed in this section. The response for a voltage step propagating along a given edge e can be obtained using EMT simulation software, knowing the geometry of the transmission lines and the characteristics of the conductors. In particular, the step response depends on the length of the considered segment d_e and the value of the soil resistivity ρ_e, which is considered as a known and constant characteristic of the considered line. The

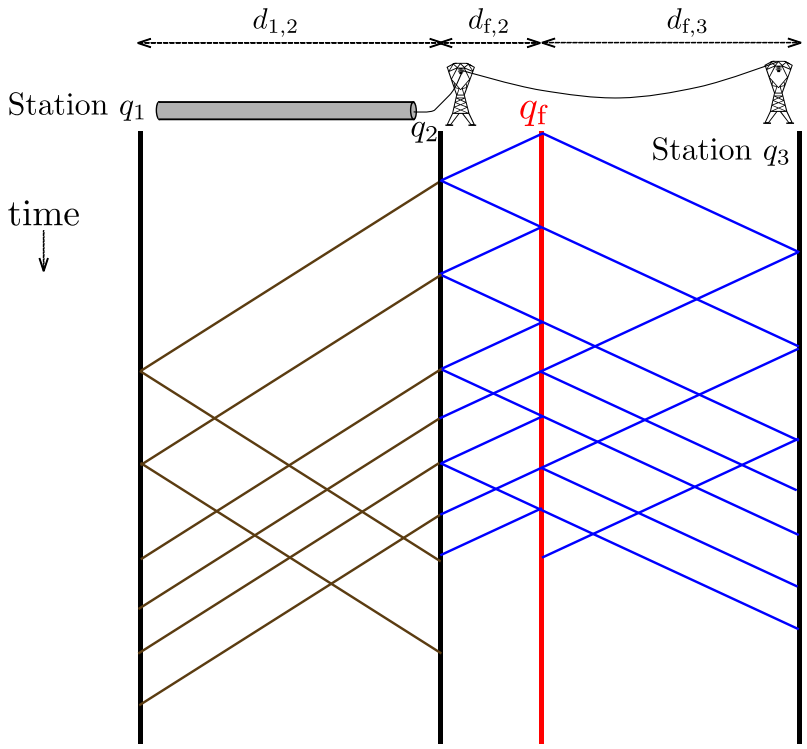


Figure 2.18.: Example of Bewley lattice diagram for a hybrid point-to-point link when a fault occurs as q_f located in an overhead portion of the line.

propagation delay is removed from the step responses as it is already accounted for by the propagation constant delay (2.66).

Assume that a set of known step responses $u_{d,\rho}(k)$ for various edge lengths $\{d_1, d_2, \ldots d_n\}$ is available for a given conductor and line geometry, as presented in Figure 2.19. The different step responses have smooth variations with respect to the line length. Thus, to obtain a step response u_d for any length d such that $d_i < d < d_{i+1}$, $i = 1, \ldots, n-1$, we propose an interpolation using the step responses at distances d_i and d_{i+1}

$$u_d(k) = \frac{u_{d_{i+1}}(k) - u_{d_i}(k)}{(d_{i+1} - d_i)}(d - d_i) + u_{d_i}(k). \tag{2.70}$$

The step response of a given edge e is differentiated to obtain the impulse response h_e

$$h_e(k) = \frac{u_{d,\rho}(k+1) - u_{d,\rho}(k)}{T_s}. \tag{2.71}$$

If the step response for the soil resistivity ρ_e of the edge e is unknown, it can be interpolated from step responses at known soil resistivities $\rho_1, \ldots, \rho_m$ in a similar fashion as (2.70).

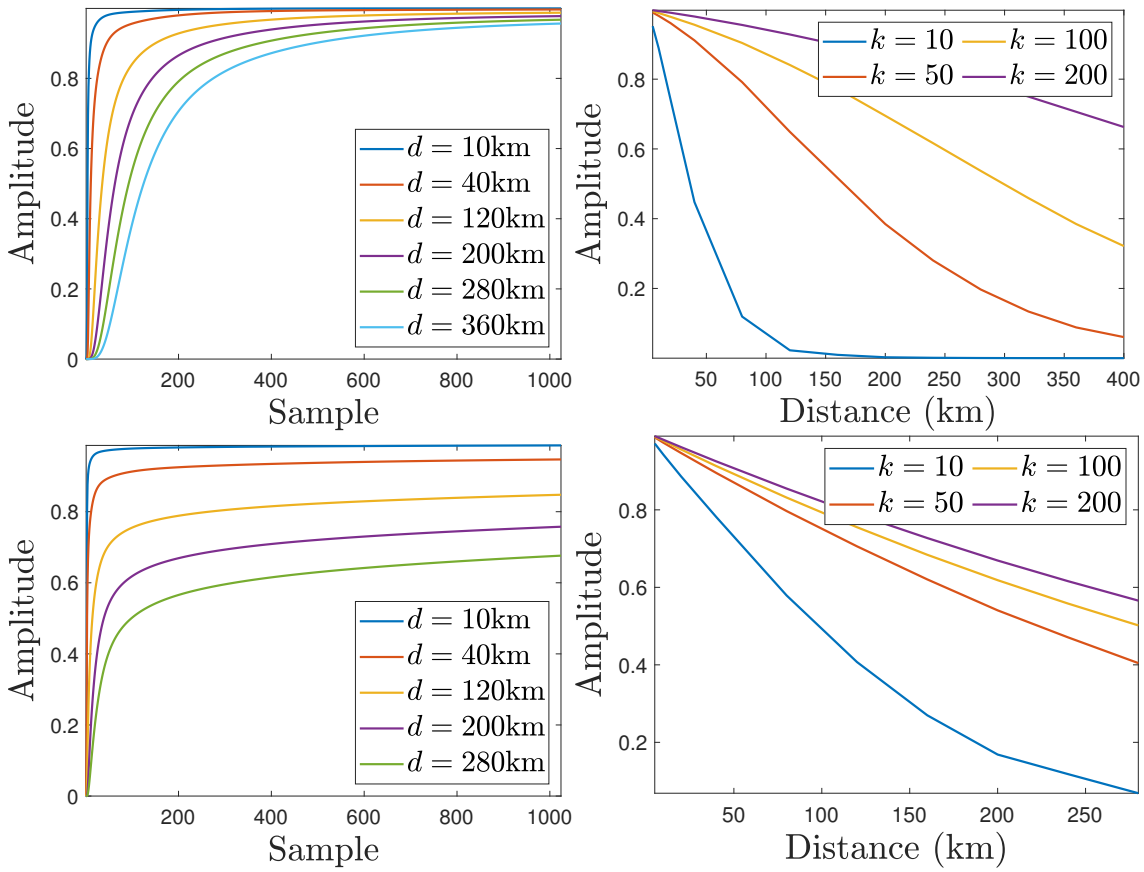


Figure 2.19.: Left: step response for overhead lines (top) and underground cables (bottom) of different lengths. Right: the evolution of the step responses at different samples k with the fault distance is detailed.

The evolution of a wave traveling through an edge e of length d_e is obtained as the output of the finite impulse response filter excited by the output of the physical model (2.68) of the edge e

$$v_e^{\mathrm{m}}(k,\mathbf{p}) = h_e\left(k, d_e\right) \otimes v_e^0\left(k,\mathbf{p}\right) \tag{2.72}$$

For a wave traveling through a path $\pi \in \mathcal{P}$ comprising several edges, the total voltage evolution at node q_{s} is obtained by cascading the impulse responses of the different edges, see for example Figure 2.20

$$v_\pi^{\mathrm{m}}\left(k,\mathbf{p}\right) = \underbrace{\bigotimes_{e\in\pi} h_e\left(k, d_e\right)}_{=h_\pi(k,d_{\mathrm{f}})} \otimes v_\pi^0(k,\mathbf{p}) \tag{2.73}$$

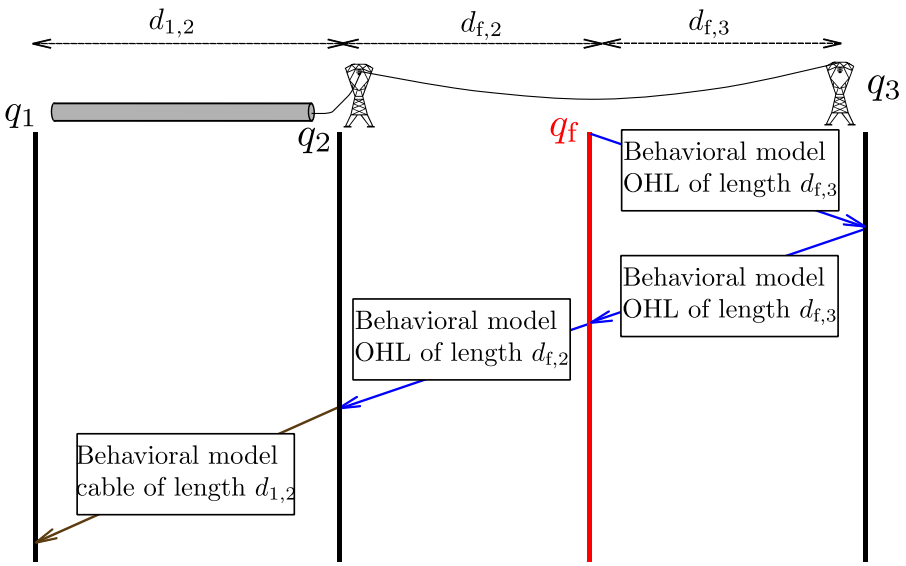


Figure 2.20.: Example of cascaded behavioral model to take into account segmented transmission lines

The total model of the voltage at the node of interest q_{s} is then the sum of all possible traveling waves between the faulty node q_{f} and q_{s}

$$v_{q_{\mathrm{s}}}^{\mathrm{m}}\left(\mathbf{p},k\right) = \sum_{\pi\in\mathcal{P}_{q_{\mathrm{f}}\to q_{\mathrm{s}}}} v_\pi^{\mathrm{m}}\left(\mathbf{p},k\right). \tag{2.74}$$

When considering a finite observation window of duration $\tau_{\max}$ after the occurrence of a fault, only a finite number of traveling waves may reach the node q_{s} within this time observation window. This reduces the set of paths to consider for simulation

$$\mathcal{P}_{q_{\mathrm{f}}\to q_{\mathrm{s}},\tau_{\max}} = \left\{\pi \in \mathcal{P}_{q_{\mathrm{f}}\to q_{\mathrm{s}}} | \tau_\pi < \tau_{\max}\right\}. \tag{2.75}$$

An alternative approach to limit the computational complexity is to simulate a maximum number$n_{\max}$ of TW and to consider as many paths.

2.4.3. Illustration of the modeling approach for hybrid lines

The meshed HVDC grid presented in Figure 1.4 is modified to include two hybrid lines, as presented in Figure 2.21. Two faults affecting line L_{14} are considered: fault F_1 affects the edge $e_{5,6}$ (cable section) at a distance $d_f = 15\,\text{km}$ from node q_5 with an resistance $R_f = 0.1\,\Omega$ and fault F_2 occurs on the edge $e_{6,7}$ (OHL section) at a distance $d_f = 180\,\text{km}$ from station q_1 with a resistance $R_f = 30\,\Omega$. The proposed approach for the modeling of the TW that arrive at the station q_1 is applied to the two fault cases.

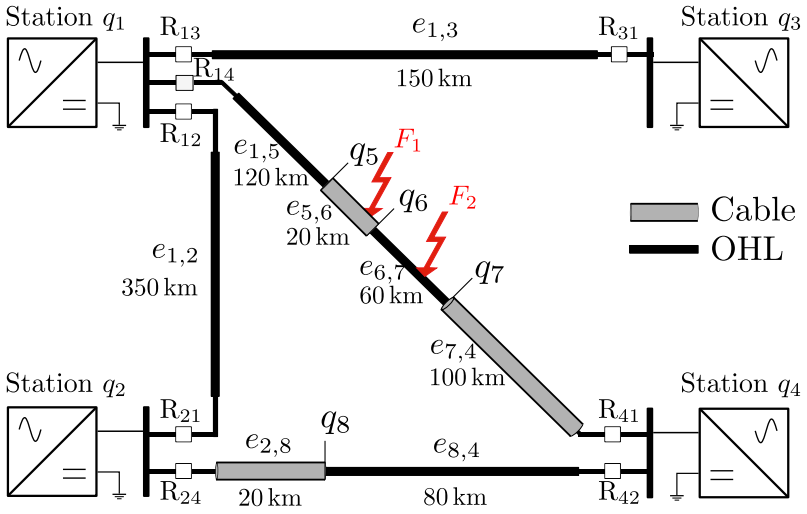


Figure 2.21.: HVDC grid comprising two hybrid lines.

The model output for the current and voltage and current transient evolution at relay R_{14} are compared with EMT data in Figure 2.23 for fault F_1 and in Figure 2.22 for fault F_2. The proposed model fits relatively well the EMT data despite the the numerous TW that must be accounted for during the considered 0.4 ms long measurement window.

2.5. Conclusion

In this chapter, we proposed a combined approach to model the transient behavior after a fault occurrence on a transmission line. The traveling wave theory and modal analysis, see Section 2.1 provide adequate tools for the description of the fault transientsNevertheless, the soil resistivity causes an important frequency dependent behavior of the line parameters. Explicit expressions of the voltage and current evolution can be obtained but they do not describe the distortion and attenuation due to the soil resistivity. We propose to supplement the *physical model* obtained from the traveling wave with a *behavioral approach* to represent such effects, see 2.2. Low pass filters are tuned so that the output of the combined model fits the

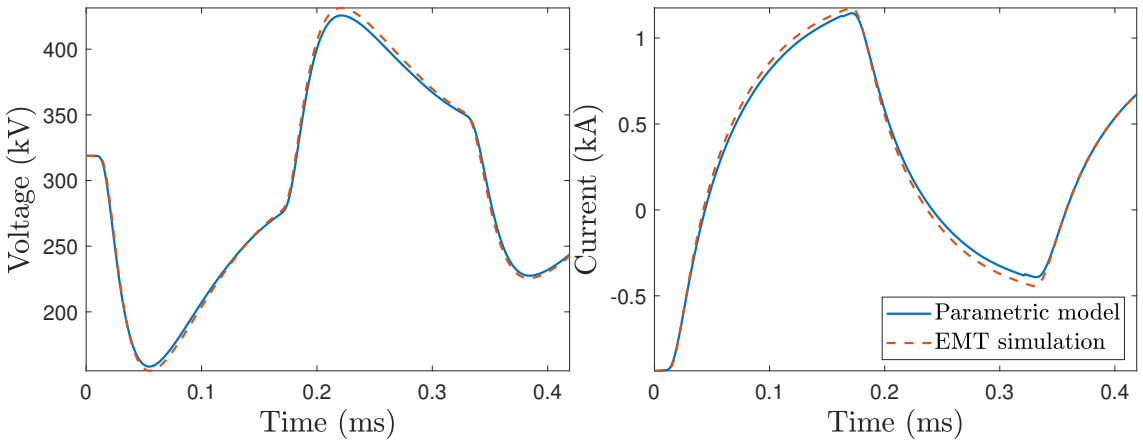


Figure 2.22.: Comparison of the voltage (left) and current (right) transient models with the output of an EMT simulation software; the simulated fault is located in a cable section at 135 km from the station 1 and has a resistance $R_f = 0.1\,\Omega$.

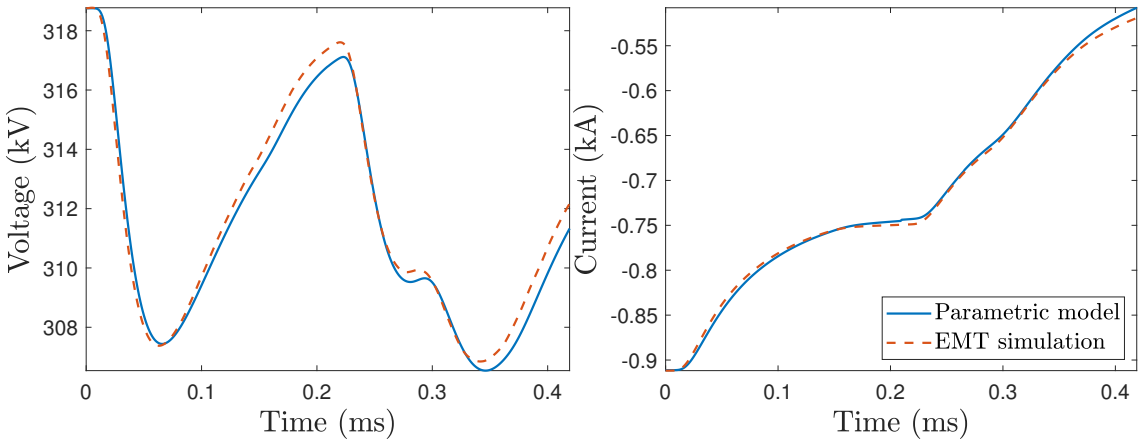


Figure 2.23.: Comparison of the voltage (left) and current (right) transient models with the output of an EMT simulation software; the simulated fault is located in an overhead section at 180 km from the station 1 and has a resistance of $30\,\Omega$.

TW observed with an EMT simulation software. A generic description of those filters was also obtained for any fault distance. This allows the combined model to depend explicitly on the fault parameters such as the fault resistance and distance. The proposed approach has been extended to hybrid transmission lines comprising portions of underground cables and overhead lines in Section 2.4. In this case, the multiple junctions within the line create a large number of transmitted and reflected waves. A more systematic description of the network as a graph allows to model any wave traveling through the grid.

3. Fault identification algorithm

In this chapter, the *parametric model* describing the transient evolution of the voltage and current after a fault developed in Chapter 2 is employed to identify the faulty line. Assuming the protected line is faulty, an estimate in the maximum likelihood sense of the unknown fault parameters is evaluated from the received measurements, as detailed in Section 3.1. Based on the results of the parameter estimation, a decision logic is used to confirm or deny that the line under protection is actually faulty. This faulty line identification step, described in Section 3.2, can also be adapted to the case when the line is composed of different segments such as hybrid lines comprising portions of overhead lines and underground cables. Implementation aspects of the proposed algorithm is detailed in Section 3.3.

3.1. Estimation of the fault parameters

In this section, the estimation of the fault parameters is detailed. This step corresponds to the *parameter tuning* block of Figure 1.9 introduced in Section 1.3 and uses the parametric model described in 2.

3.1.1. Overview

Consider a fault occurring at some time instant t_f on a given line e of length d connecting the substations q and q'. The transmission line e is composed of n_c conductors. The distances between the fault and the stations q and q' are $d_{\mathrm{f},q}$ and $d_{\mathrm{f},q'} = d - d_{\mathrm{f},q}$. The fault is assumed to be characterized by its pole-to-ground or pole-to-pole resistance R_f depending on the type of fault. A fault can hence be described by the vector of the fault parameters $\mathbf{p} = (e, d_{\mathrm{f},q}, R_\mathrm{f}, t_\mathrm{f})^\mathrm{T}$.

Each relay must first determine whether the grid behaves normally or not and the suspected type of fault (pole-to-pole, pole-to-ground), using for instance tools presented in [Azad et al., 2015, Torwelle et al., 2020]. An abnormal behavior is detected when the rate of change of the voltage (ROCOV) goes above a predetermined threshold. A fault is then suspected, and each relay must confirm or refute that a fault occurred on their protected line. Thus, the line e in $\mathbf{p}$ is fixed during the fault identification process. Furthermore $d_{\mathrm{f},q}$ and t_f are linked through the detection time

of the first TW at the relay located at node q,

$$t_{\mathrm{d},q} = t_{\mathrm{f}} + \frac{d_{\mathrm{f},q}}{c_{\mathrm{w}}}, \tag{3.1}$$

which can be measured. In (3.1), c_{w} is the propagation speed of the TW. Thus the vector of the fault parameters can be reduced to $\mathbf{p} = (d_{\mathrm{f},q}, R_{\mathrm{f}})^{\mathrm{T}}$.

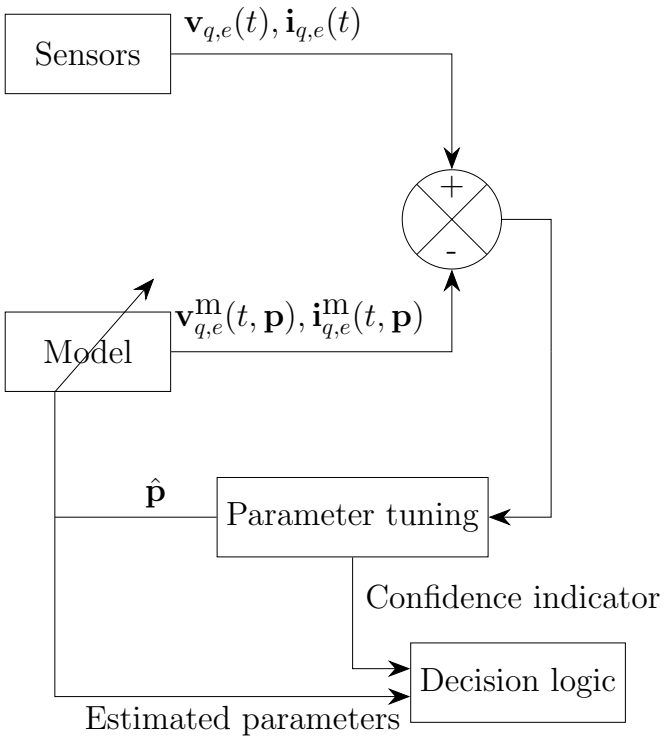


Figure 3.1.: Components of the fault identification algorithm based on the estimation of the fault parameters.

The principle of the fault identification algorithm based on the estimation of the fault parameters is reminded in Figure 3.1. The relay at substation q monitoring line e acquires voltage and current measurements $(\mathbf{v}_{q,e}(t), \mathbf{i}_{q,e}(t))$ at a frequency f_{s} from the sensors located at the end of e, where the voltage and current vectors $\mathbf{v}_{q,e} = \left(v^1_{q,e}, \ldots, v^{n_{\mathrm{c}}}_{q,e}\right)^T$ and $\mathbf{i}_{q,e} = \left(i^1_{q,e}, \ldots, i^{n_{\mathrm{c}}}_{q,e}\right)^T$ gather the data for all the n_{c} conductors. The output at time t of the model for all the conductors is denoted

$$\mathbf{y}^{\mathrm{m}}_{q,e} = \left(\mathbf{v}^{\mathrm{m}}_{q,e}(t,\mathbf{p})^T, \mathbf{i}^{\mathrm{m}}_{q,e}(t,\mathbf{p})^T\right)^T \tag{3.2}$$

The output of the model and the measurements from the sensors are used to evaluate an estimate $\hat{\mathbf{p}}$ of the vector of the fault parameters. The value of the estimated

parameters as well as an approximation of the confidence region are employed in a decision logic to confirm or refute that the fault occurred on line e. As the considered line e and station q are fixed during the parameter estimation process, they are omitted in the rest of the following sections to lighten the notations.

3.1.2. Maximum Likelihood estimation

This section briefly presents the main concepts associated to the maximum likelihood estimation. Further details can be found in [Walter and Pronzato, 1997]. Consider that an unusual behavior is detected at time $t_{\mathrm{d},q}$ by a relay q monitoring a line e. Assume that n voltage and current measurements $\mathbf{y}(t) = \left(\mathbf{v}(t)^T, \mathbf{i}(t)^T\right)^T$ have been obtained at time t_k, $k = 1, \ldots, n$, with $t_1 \leqslant t_{\mathrm{d},q} < t_n$. Note that for n measurement instants, the number of associated samples is actually $2 \times n_{\mathrm{c}} \times n$ to account for the voltage and current measurements on all conductors. In this chapter, the *number of observations* n always refer to the number of time instants considered.

For a given value of $\mathbf{p}$, one may evaluate the combined model outputs

$$\mathbf{y}^{\mathrm{m}}(\mathbf{p}, t) = \left(\mathbf{v}^{\mathrm{m}}(\mathbf{p}, t)^T \;\; \mathbf{i}^{\mathrm{m}}(\mathbf{p}, t)^T\right)^T$$

and evaluate the output error for a value $\mathbf{p}$ of the vector of parameters

$$\mathbf{e}_{\mathrm{y}}(t_k, \mathbf{p}) = \mathbf{y}(t_k) - \mathbf{y}^{\mathrm{m}}(t_k, \mathbf{p}), \;\; k = 1, \ldots, n. \tag{3.3}$$

Assume that, for the true value of the vector of fault parameters $\mathbf{p}^*$, the observed data satisfy

$$\mathbf{y}(t_k) = \mathbf{y}^{\mathrm{m}}(t_k, \mathbf{p}^*) + \boldsymbol{\varepsilon}(t_k), \;\; k = 1, \ldots, n \tag{3.4}$$

where $\boldsymbol{\varepsilon}(t_k)$ accounts for both the measurement and model noises. We assume $\boldsymbol{\varepsilon}(t_k)$ are a sequence of stationary random vectors that follow a zero-mean Gaussian distribution with *a priori* unknown covariance matrix $[\Sigma]$.

Each measurement $\mathbf{y}(t_k)$ can thus be seen as a random variable with a normal distribution whose likelihood $\mathcal{L}$ is a function of the observed data $\mathbf{y}$, the parameter vector $\mathbf{p}$, and the covariance $[\Sigma]$

$$\mathcal{L}(\mathbf{y}(t_k), \mathbf{p}, [\Sigma]) \tag{3.5}$$

$$= \frac{1}{\sqrt{(2\pi)^2 \det[\Sigma]}} \exp\left(-\frac{1}{2}\left(\mathbf{y}(t_k) - \mathbf{y}^{\mathrm{m}}(t_k, \mathbf{p})\right)^T [\Sigma]^{-1} \left(\mathbf{y}(t_k) - \mathbf{y}^{\mathrm{m}}(t_k, \mathbf{p})\right)\right). \tag{3.6}$$

As the noise $\boldsymbol{\varepsilon}(t_k)$ are considered stationary and identically distributed, the likelihood of n observations is

$$\begin{aligned}&\mathcal{L}\left(\mathbf{y}\left(t_1\right),\ldots,\mathbf{y}\left(t_n\right),\mathbf{p},[\Sigma]\right)\\&=\prod_{k=1}^{n}\frac{1}{\sqrt{\left(2\pi\right)^2\det\left[\Sigma\right]}}\exp\left(-\frac{1}{2}\left(\mathbf{y}\left(t_k\right)-\mathbf{y}^{\mathrm{m}}\left(t_k,\mathbf{p}\right)\right)^T\left[\Sigma\right]^{-1}\left(\mathbf{y}\left(t_k\right)-\mathbf{y}^{\mathrm{m}}\left(t_k,\mathbf{p}\right)\right)\right)\end{aligned}\tag{3.7}$$

$$=\left(\left(2\pi\right)^2\det\left[\Sigma\right]\right)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}\sum_{k=1}^{n}\left(\mathbf{y}\left(t_k\right)-\mathbf{y}^{\mathrm{m}}\left(t_k,\mathbf{p}\right)\right)^T\left[\Sigma\right]^{-1}\left(\mathbf{y}\left(t_k\right)-\mathbf{y}^{\mathrm{m}}\left(t_k,\mathbf{p}\right)\right)\right)\tag{3.8}$$

Introducing the vector of the residuals for the n received measurements $\mathbf{f}^{(n)}\left(\mathbf{p}\right)\in\mathbb{R}^{2\times n\times n_c}$,

$$\mathbf{f}^{(n)}\left(\mathbf{p}\right)=\left[\left(\mathbf{y}\left(t_1\right)-\mathbf{y}^{\mathrm{m}}\left(\mathbf{p},t_1\right)\right)^T,\ldots,\left(\mathbf{y}\left(t_n\right)-\mathbf{y}^{\mathrm{m}}\left(\mathbf{p},t_n\right)^T\right)^T,\tag{3.9}$$

the likelihood (3.8) can be rewritten as

$$\mathcal{L}\left(\mathbf{y}\left(t_1\right),\ldots,\mathbf{y}\left(t_n\right),\mathbf{p},[\Sigma]\right)=\left(\left(2\pi\right)^2\det\left[\Sigma\right]\right)^{-\frac{n}{2}}\exp\left(-\frac{1}{2}\mathbf{f}^{(n)}\left(\mathbf{p}\right)^T\left[\Sigma^{(n)}\right]^{-1}\mathbf{f}^{(n)}\left(\mathbf{p}\right)\right),\tag{3.10}$$

where the covariance matrix $\left[\Sigma^{(n)}\right]$ is a block-diagonal matrix, each block being equal to the covariance matrix $[\Sigma]$.

The maximum likelihood estimate of $\mathbf{p}$ and $[\Sigma]$ from n observations is

$$\left(\widehat{\mathbf{p}},\left[\widehat{\Sigma}\right]\right)=\arg\max_{\mathbf{p},[\Sigma]}\mathcal{L}\left(\mathbf{y}\left(t_1\right),\ldots,\mathbf{y}\left(t_n\right),\mathbf{p},[\Sigma]\right).\tag{3.11}$$

As the logarithm is a monotonically increasing function, The ML estimate also maximizes the log-likelihood,

$$\left(\widehat{\mathbf{p}},\left[\widehat{\Sigma}\right]\right)=\arg\max_{\mathbf{p},[\Sigma]}\log\mathcal{L}\left(\mathbf{y}\left(t_1\right),\ldots,\mathbf{y}\left(t_n\right),\mathbf{p}\right).\tag{3.12}$$

here

$$\log\mathcal{L}\left(\mathbf{y}\left(t_1\right),\ldots,\mathbf{y}\left(t_n\right),\mathbf{p},[\Sigma]\right)=-\frac{n}{2}\log\left(\left(2\pi\right)^2\det\left[\Sigma\right]\right)-\frac{1}{2}\mathbf{f}^{(n)}\left(\mathbf{p}\right)^T\left[\Sigma^{(n)}\right]^{-1}\mathbf{f}^{(n)}\left(\mathbf{p}\right).\tag{3.13}$$

The minimization of (3.13) leads to [Walter and Pronzato, 1997]

$$\widehat{\mathbf{p}}=\arg\min_{\mathbf{p}}\ \log\det\sum_{k=1}^{n}\left(\mathbf{y}\left(t_k\right)-\mathbf{y}^{\mathrm{m}}\left(\mathbf{p},t_k\right)\right)\left(\mathbf{y}\left(t_k\right)-\mathbf{y}^{\mathrm{m}}\left(\mathbf{p},t_k\right)\right)^T\tag{3.14}$$

and

$$\left[\widehat{\Sigma}\right] = \frac{1}{n}\sum_{k=1}^{n}\left(\mathbf{y}\left(t_k\right) - \mathbf{y}^{\mathrm{m}}\left(\widehat{\mathbf{p}}, t_k\right)\right)\left(\mathbf{y}\left(t_k\right) - \mathbf{y}^{\mathrm{m}}\left(\widehat{\mathbf{p}}, t_k\right)\right)^T. \tag{3.15}$$

One can further assume that the measurement noise are independent random variables and that the model noise can be neglected. The covariance matrix $[\Sigma]$ is then diagonal. Its components may be obtained from the sensor characteristics. The ML estimation (3.12) then aims at evaluating

$$\widehat{\mathbf{p}} = \arg\min_{\mathbf{p}} c^{(n)}(\mathbf{p}), \tag{3.16}$$

where the cost function is

$$c^{(n)}(\mathbf{p}) = \mathbf{f}^{(n)}(\mathbf{p})^T \left[W^{(n)}\right] \mathbf{f}^{(n)}\left(\mathbf{p}\right). \tag{3.17}$$

The evaluation of (3.16) corresponds to a weighted non-linear least squares problem where the weight matrix is

$$\left[W^{(n)}\right] = \left[\Sigma^{(n)}\right]^{-1} = \begin{pmatrix} \lambda_{\mathrm{v}}^1 & 0 & & & & 0 \\ 0 & \ddots & & & & \\ & & \lambda_{\mathrm{v}}^{n_{\mathrm{c}}} & & & \\ & & & \lambda_{\mathrm{i}}^1 & & \\ & & & & \ddots & 0 \\ 0 & & & & 0 & \lambda_{\mathrm{i}}^{n_{\mathrm{c}}} \end{pmatrix} \tag{3.18}$$

and the weight coefficients corresponding to the voltage and current on conductor ℓ are $\lambda_{\mathrm{v}}^{\ell}$ and $\lambda_{\mathrm{i}}^{\ell}$ respectively.
The minimization of (3.14) may be performed using gradient descent as the second order derivatives of the log det criterion are expensive to obtain. Instead, the weighted least-squares problem (3.17) allows one to employ second-order methods such as Gauss-Newton algorithm which generally offer faster convergence.

The difference between the log-det and least-squares criteria is observed on the level curves of the two cost functions on a particular example in Figure 3.2. The vector of the true fault parameters $d_{\mathrm{f}}^* = 90\,\mathrm{km}$, $R_{\mathrm{f}}^* = 80\,\Omega$ is indicated. The cost functions are computed in the bipolar case on 107 samples and have been normalized so that the minimum of the cost is, in absolute value, 1. The minimization of the least-squares criterion may induce a higher bias in the estimated distance than the log-det criterion, though the minimum of the two cost functions does not correspond to the true fault parameters. We propose to employ the least-squares criterion as its minimization allows to employ more efficient optimization techniques and the additional bias seems moderate. Under the assumptions that for the true value of the fault parameters $\mathbf{p}^*$, the measurement noise (3.4) is an i.i.d. random variable, as well as more technical conditions, see [Walter, 2014], the ML estimate is *consistent, i.e.,* it converges in probability to the true value of the parameter $\mathbf{p}^*$

$$\forall \delta > 0,\ \mathbb{P}\left(\|\mathbf{p}^* - \widehat{\mathbf{p}}\| \geq \delta\right) \underset{n\to\infty}{\longrightarrow} 0. \tag{3.19}$$

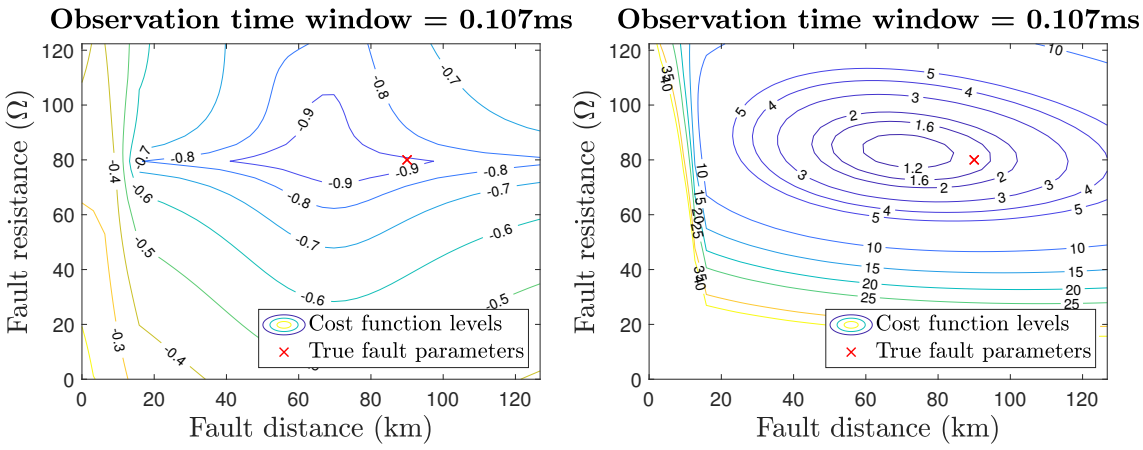


Figure 3.2.: Contour plot of the log-det (3.14) (left) and least squares (3.17) (right) criteria on a particular example.

The ML estimate is also *asymptotically Gaussian* and *unbiased*, as it converges in distribution (d) to a normally distributed random variable,

$$\widehat{\mathbf{p}} \xrightarrow[n\to\infty]{d} \mathcal{N}\left(\mathbf{p}^*, [\Sigma_{\mathrm{ML}}]\right) \tag{3.20}$$

where the covariance matrix $[\Sigma_{\mathrm{ML}}]$ corresponds to the inverse of Fisher information, which is exploited in Section 3.2 for the identification of the faulty line.

In order to solve the minimization problem (3.17) the Levenberg-Marquardt algorithm is employed, see Section 3.1.3. The inherent issues of solving a non-linear problem with a local optimization algorithm such as local minima and initial point choice are discussed in Section 3.3 as well as in the simulations presented in Chapter 4.

3.1.3. Levenberg-Marquardt algorithm

This section reminds the principle of the Levenberg-Marquardt algorithm [Levenberg, 1944, Marquardt, 1963], used for the minimization of the cost function (3.17). More details about optimization tools can be found in [Walter, 2014].

Consider a given number of observations n and an initialization value for the vector of the fault parameters $\mathbf{p}_0^{(n)}$. The ML estimate (3.17) can be computed iteratively

$$\mathbf{p}_{k+1}^{(n)} = \mathbf{p}_k^{(n)} + \boldsymbol{\delta}_k^{(n)}, \tag{3.21}$$

where $\boldsymbol{\delta}_k^{(n)}$ is a correction term such that the cost function is ensured to decrease, $c^{(n)}(\mathbf{p}_{k+1}^{(n)}) < c^{(n)}(\mathbf{p}_k^{(n)})$. In the Levenberg-Marquardt algorithm, the correction term is

$$\boldsymbol{\delta}_k^{(n)} = -\left(\left[J_{\mathrm{f}}^{(n)}\right]^T \left[W^{(n)}\right] \left[J_{\mathrm{f}}^{(n)}\right] + \mu_k^{(n)} [D]\right)^{-1} \left[J_{\mathrm{f}}^{(n)}\right]^T \left[W^{(n)}\right] \mathbf{f}^{(n)}, \tag{3.22}$$

where $\left[J_{\mathrm{f}}^{(n)}\right] = \frac{\partial \mathbf{f}^{(n)}(\mathbf{p})}{\partial \mathbf{p}^T}\Big|_{\mathbf{p}=\mathbf{p}_k^{(n)}}$ is the Jacobian matrix of $\mathbf{f}^{(n)}$ with respect to $\mathbf{p}$ evaluated at $\mathbf{p}_k^{(n)}$. A damping term $\mu_k^{(n)}\,[D]$ is added so that

- When μ_k is large the algorithm behaves as a gradient descent with small step,
- When μ_k is small the algorithm behaves as a Gauss-Newton and progresses fast.

The damping matrix $[D]$ can be chosen as the diagonal elements of $\left[J_{\mathrm{f}}^{(n)}\right]^T \left[W^{(n)}\right] \left[J_{\mathrm{f}}^{(n)}\right]$ to compensate for the different scales of the parameters. The damping matrix can also be adapted at each step, as suggested in [Transtrum and Sethna, 2012], so that the i^{th} diagonal element is

$$[D]_{i,i} \leftarrow \max\left([D]_{i,i}\,; \left(\left[J_{\mathrm{f}}^{(n)}\right]^T \left[W^{(n)}\right] \left[J_{\mathrm{f}}^{(n)}\right]\right)_{i,i}\right).$$

A simpler yet effective choice is the identity matrix, $[D] = [I_{2n_{\mathrm{c}}\times n}]$.

The damping parameter μ_k should then be decreased as the estimate gets closer to the vector of the true fault parameters. The usual tuning strategy is chosen,

1. Compute $\mathbf{p}_{k+1}^{(n)}$ and the associated cost $c^{(n)}(\mathbf{p}_{k+1}^{(n)})$.
2. If $c^{(n)}(\mathbf{p}_{k+1}^{(n)}) < c^{(n)}(\mathbf{p}_k^{(n)})$, accept the iteration $k+1$ and decrease the damping parameter by some coefficient $\mu_{\downarrow}$.
3. If $c^{(n)}(\mathbf{p}_{k+1}^{(n)}) \geq c^{(n)}(\mathbf{p}_k^{(n)})$, reject the iteration $k+1$ and increase the damping parameter by some coefficient $\mu_{\uparrow}$. Go back to step 1 with the updated value of μ_k.

Step 3 can be repeated a maximum number of times, after what the iteration $k+1$ is accepted whatever the cost evolution. The tuning of the initial damping values μ_0 as well as the coefficients μ_{down} and μ_{up} are indicated in Section 3.3.

The computation of the Jacobian matrix requires the evaluation of the gradient of the voltage and current model output with respect to the fault distance and resistance. In most cases, the structure of the models developed in Chapter 2 allows an explicit computation of the derivatives, which is addressed in the Appendix C. In the remaining cases, a finite difference approximation may be used.

As indicated in Section 1.3 new data can be added to the cost function (3.17) so as to improve the estimate. Consider that Δn new samples are added after κ iterations of the LM algorithm. The estimation of the fault parameters proceeds with the new data set starting from the previously obtained estimate. Formally

$$\mathbf{p}_0^{(n+\Delta n)} = \mathbf{p}_{\kappa}^{(n)}.$$

In practice, the number of iterations per data set κ and the number of additional measurements Δn required to update the cost function may be tuned to reach a compromise between the speed and accuracy of the estimation process. For each

obtained estimate $\mathbf{p}_k^{(n)}$, the algorithm can test the stopping conditions that evaluate whether the fault is confirmed to be on the protected line or not,as detailed in Section 3.2. The maximum duration of the measurement window that can be used by the optimization algorithm is fixed to $\tau_{\max}$. The maximum number of iterations performed by the optimization algorithm in case the maximum number of measurement points are used is

$$f_s \frac{\tau_{\max}}{\Delta n} \kappa \tag{3.23}$$

depends on the sampling frequency f_s as well as $\tau_{\max}$, κ, and Δn whose settings are discussed in Section 3.3.

3.2. Faulty line identification

This section presents the decision logic that allows the algorithm to conclude whether a fault actually occurred on the protected line or not. This step corresponds to the *decision logic* block of Figure 1.9 introduced in Section 1.3.

After each iteration of the estimation, the algorithm determines whether it has obtained a satisfying estimate of the fault parameters with a good level of confidence, indicating that the monitored line is faulty. The algorithm may also stop once it has performed a maximum number of iterations, corresponding to a maximum measurement window. In the latter case, the fault is considered to be outside the protected line, or to be non-existent.
To determine whether the estimate is consistent with the hypothesis that the monitored line is faulty, two tests are considered. First, a *validity* test determines whether or not the value of the estimated parameter vector is included in some domain of interest. Second, an *accuracy* test determines if the confidence region associated to the estimate is small enough. In this section the number n of observed data is omitted to lighten the notations.

First, the estimated parameters must belong to a certain protection zone $\mathcal{D}_{\mathbf{p}}$ which represents plausible values for the fault parameters. $\mathcal{D}_{\mathbf{p}}$ may be defined as

$$\mathcal{D}_{\mathbf{p}} = \{(d_f\,, R_f) | d_{f,\min} \leq d_f \leq d_{f,\max}, R_{f,\min} \leq R_f < R_{f,\max}\} \tag{3.24}$$

where $(d_{\min}, d_{\max})$ defines the portion of the line actually monitored by the relay and $(R_{f,\min}, R_{f,\max})$ the range of fault resistance that requires fast decision, since a high value of R_f corresponds to a non-critical fault for which more time is available to take action, as investigated in [Auran et al., 2017]. The characteristics of the validity domain $\mathcal{D}_{\mathbf{p}}$ are detailed in Section 3.3.3.

Second, in the accuracy test, for each estimate $\mathbf{p}_k$ of $\mathbf{p}^*$ belonging to $\mathcal{D}_{\mathbf{p}}$, an associated confidence region is evaluated. Various approaches for this evaluation are available, see, *e.g.*, [Kay, 1993, Walter and Pronzato, 1997].

As mentioned in (3.20), under some assumptions [Walter and Pronzato, 1997], the ML estimate $\mathbf{p}$ converges in distribution to a normal distribution such that

$$(\mathbf{p} - \mathbf{p}^*) \xrightarrow{d} \mathcal{N}(0, [\Sigma_{\mathrm{ML}}])$$

Where the covariance matrix

$$[\Sigma_{\mathrm{ML}}] = \mathbb{E}_{\mathbf{y}|\mathbf{p}^*}\left\{(\mathbf{p}_k - \mathbf{p}^*)(\mathbf{p}_k - \mathbf{p}^*)^T\right\} \tag{3.25}$$

is the inverse of the Fisher information matrix

$$[\Sigma_{\mathrm{ML}}] = [F_I]^{-1} \tag{3.26}$$

defined as

$$[F_I,] = -\mathbb{E}_{\mathbf{y}|\mathbf{p}^*}\left\{\frac{\partial^2 \ln \mathcal{L}(\mathbf{y}(t_1), \ldots, \mathbf{y}(t_n), \mathbf{p})}{\partial \mathbf{p} \partial \mathbf{p}^T}\right\} \tag{3.27}$$

$$= \mathbb{E}_{\mathbf{y}|\mathbf{p}^*}\left\{\frac{\partial \ln \mathcal{L}(\mathbf{y}(t_1), \ldots, \mathbf{y}(t_n), \mathbf{p})}{\partial \mathbf{p}} \frac{\partial \ln \mathcal{L}(\mathbf{y}(t_1), \ldots, \mathbf{y}(t_n), \mathbf{p})}{\partial \mathbf{p}^T}\right\} \tag{3.28}$$

The proposed evaluation of the Fisher information matrix relies on assumptions that are in general not satisfied in practice. In particular, it is assumed the model (or characterization) error is zero (or at least be small compared to the measurement noise). Moreover, the measurement noise due to the sensors may not be independent and identically distributed with zero-mean as bias related to long-term dynamic such the temperature can affect the sensors, as detailed in Section 5.4. Nevertheless, the Fisher information gives a reasonable and fast evaluation of the estimation uncertainty.

The log likelihood of n observations is

$$\log \mathcal{L}(\mathbf{y}(t_1), \ldots, \mathbf{y}(t_n), \mathbf{p}) = -\frac{n}{2} \ln\left((2\pi)^2 \det\left[\Sigma^{(n)}\right]\right) - \frac{1}{2}\mathbf{f}^{(n)}(\mathbf{p})^T \left[\Sigma^{(n)}\right]^{-1} \mathbf{f}^{(n)}(\mathbf{p}) \tag{3.29}$$

Assuming the covariance matrix does not depend on $\mathbf{p}$, one has

$$[F_I] = \left[J^{(n)}(\mathbf{p}^*)\right]^T \left[\Sigma^{(n)}\right]^{-1} \left[J^{(n)}(\mathbf{p}^*)\right]. \tag{3.30}$$

As the true vector of the fault parameters is not available, $\mathbf{p}^*$ is replaced by the estimate $\mathbf{p}_k^{(n)}$ in (3.30),

When $\left[\Sigma^{(n)}\right]$ is not known, it can be replaced using an estimate of the covariance matrix $[\Sigma]$ as in (3.15)

$$\left[\widehat{\Sigma}\right] = \frac{1}{n}\sum_{k=1}^{n}(\mathbf{y}(t_k) - \mathbf{y}^{\mathrm{m}}(\mathbf{p}, t_k))(\mathbf{y}(t_k) - \mathbf{y}^{\mathrm{m}}(\mathbf{p}, t_k))^T \tag{3.31}$$

As the error has been assumed stationary, an estimate of the total covariance matrix is

$$\left[\widehat{\Sigma}^{(n)}\right] = \begin{pmatrix} \left[\widehat{\Sigma}\right] & 0 & \dots & 0 \\ 0 & \left[\widehat{\Sigma}\right] & & \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & \left[\widehat{\Sigma}\right] \end{pmatrix} \tag{3.32}$$

The covariance matrix $\left[\widehat{\Sigma}\right]$ can be considered diagonal if the measurement errors of the different voltage and current sensors at time t_k are independent of each other. In that case one has

$$\left[\widehat{\Sigma}\right] = \begin{pmatrix} \widehat{\sigma}^2_{\mathrm{v},1} & 0 & & & & 0 \\ 0 & \ddots & & & & \\ & & \widehat{\sigma}^2_{\mathrm{v},n_\mathrm{c}} & & & \\ & & & \widehat{\sigma}^2_{\mathrm{i},1} & & \\ & & & & \ddots & 0 \\ 0 & & & & 0 & \widehat{\sigma}^2_{\mathrm{i},n_\mathrm{c}} \end{pmatrix} \tag{3.33}$$

where the variances are estimated for each conductor $\ell = 1, \dots, n_\mathrm{c}$ by [Walter and Pronzato, 1997]

$$\begin{aligned} \widehat{\sigma}^2_{\mathrm{v},\ell} &= \frac{1}{n - n_\mathrm{p}} \sum_{k=1}^{n} \left(v^\ell\left(t_k\right) - v^{\ell,\mathrm{m}}\left(t_k, \mathbf{p}_k\right) \right)^2 \\ \widehat{\sigma}^2_{\mathrm{i},\ell} &= \frac{1}{n - n_\mathrm{p}} \sum_{k=1}^{n} \left(i^\ell\left(t_k\right) - i^{\ell,\mathrm{m}}\left(t_k, \mathbf{p}_k\right) \right)^2 . \end{aligned} \tag{3.34}$$

For the example depicted in Figure 3.2, the estimate of the covariance matrix at the minimum of the cost function (3.31) is

$$\begin{aligned} \left[\widehat{\Sigma}\right] &= \frac{1}{n} \sum_{k=1}^{n} \left(\mathbf{y}\left(t_k\right) - \mathbf{y}^\mathrm{m}\left(\widehat{\mathbf{p}}, t_k\right) \right) \left(\mathbf{y}\left(t_k\right) - \mathbf{y}^\mathrm{m}\left(\widehat{\mathbf{p}}, t_k\right) \right)^T \\ &= \begin{pmatrix} 0.28 & -0.0038 & -0.15 & 0.022 \\ -0.0038 & 0.067 & -0.0005 & -0.038 \\ -0.15 & -0.0005 & 0.088 & -0.011 \\ 0.022 & -0.038 & -0.011 & 0.023 \end{pmatrix} \times 10^{-3}. \end{aligned} \tag{3.35}$$

The residuals are arranged such that

$$\left(\mathbf{y}\left(t_k\right) - \mathbf{y}^\mathrm{m}\left(\mathbf{p}, t_k\right) \right) = \begin{pmatrix} v^\mathrm{g}\left(t_k\right) - v^{\mathrm{m,g}}\left(\mathbf{p}, t_k\right) \\ v^\mathrm{a}\left(t_k\right) - v^{\mathrm{m,a}}\left(\mathbf{p}, t_k\right) \\ i^\mathrm{g}\left(t_k\right) - i^{\mathrm{m,g}}\left(\mathbf{p}, t_k\right) \\ i^\mathrm{a}\left(t_k\right) - i^{\mathrm{m,a}}\left(\mathbf{p}, t_k\right) \end{pmatrix}$$

where upper-scripts g and a stand for the ground and aerial modes, respectively. The estimate of the covariance matrix (3.35) suggests that the residuals of the voltage and current of different modes are relatively uncorrelated, as the entries $\left[\widehat{\Sigma}\right]_{2,1}$ and $\left[\widehat{\Sigma}\right]_{4,3}$ are relatively small compared to the diagonal elements. Similarly, the cross-correlation of the residuals between the voltage and current of different modes, given by $\left[\widehat{\Sigma}\right]_{4,1}$ and $\left[\widehat{\Sigma}\right]_{3,2}$ is small compared to the diagonal elements. On the opposite, a significant correlation appears between the residuals of the ground modes of the current and voltage as well as between the aerial modes of the current and voltage, as shown by the entries $\left[\widehat{\Sigma}\right]_{3,1}$ and $\left[\widehat{\Sigma}\right]_{4,2}$. With the analysis of (3.35), the hypothesis that the residuals, and thus the noises, are uncorrelated may not be justified for all the considered quantities. A diagonal estimate such as (3.33) of the covariance matrix was nevertheless used as it renders the computation of the Fisher information matrix easier.

Specific confidence regions can be computed using the fact that $\frac{1}{n_\mathrm{p}}(\mathbf{p}-\mathbf{p}_k^{(n)})^T\,[F_I]\,(\mathbf{p}-\mathbf{p}_k^{(n)})$ follows approximately a Fisher-Snedecor distribution $F(n_\mathrm{p}, 2n-n_\mathrm{p})$, [Walter and Pronzato, 1997]. Assuming that the number of observations n is large compared to the number of estimated parameters n_p, the $(1-\alpha)\%$ confidence region can be approximated by the ellipsoid

$$\mathcal{R}^\alpha(\mathbf{p}) = \left\{\mathbf{p}\in\mathbb{R}^{n_\mathrm{p}} \mid (\mathbf{p}-\mathbf{p}_k^{(n)})^T\,[F_I,]^{-1}\left(\mathbf{p}_k^{(n)}\right)(\mathbf{p}-\mathbf{p}_k) \leq \chi^2_{n_\mathrm{p}}(1-\alpha)\right\}, \tag{3.36}$$

where $\chi^2_{n_\mathrm{p}}(1-\alpha)$ is the value that a random variable distributed according to a chi-square distribution with n_p degrees of freedom has a probability $1-\alpha$ to be larger than.

The volume of the $(1-\alpha)\%$ confidence ellipsoids $\mathcal{R}^\alpha(\mathbf{p})$ can then be used to determine the accuracy of the current estimate $\mathbf{p}_k$ of $\mathbf{p}^*$. The estimation algorithm is deemed to have obtained an accurate estimate of the fault parameters if the volume of $\mathcal{R}^\alpha(\mathbf{p})$ is below a certain threshold $c_{1-\alpha}$

$$\mathrm{vol}\left(\mathcal{R}^\alpha(\mathbf{p})\right) \leqslant c_{1-\alpha} \tag{3.37}$$

The threshold $c_{1-\alpha}$ has to be tuned so as to ensure both . The estimation algorithm has to correctly identify all faults occurring on the protected line while rejecting faults occurring elsewhere. Tuning is done heuristically considering a large number of simulated fault cases, especially limit cases where faults occur on an adjacent line, since they are harder to distinguish from faults occurring at the remote end of the protected line.

When both validity and accuracy tests are satisfied, the estimation algorithm stops and the fault is considered to be in the protected zone. If one of the conditions is not met, Δn new measurement points are added when available and the estimation process resumes from the last parameter estimate. If the maximum length of the

measurement window $\tau_{\max}$ has been reached and the stopping condition has not been satisfied, the protected line is considered healthy.

It may happen that several fault identification algorithms are launched in parallel at the same relay. Consider for instance the case of an hybrid line L comprising segments of overhead lines and underground cables, as presented in Section 2.4. To simplify the notations, we write $e_1, \ldots, e_m$ the different edges of the line and $e_{\mathrm{f}} \in e_1, \ldots, e_m$ the faulty edge. In this case the vector of the fault parameters to estimate also includes the faulty edge e_{f}. Thus, m fault parameter estimation algorithms are launched in parallel, each one corresponding to an hypothesis on the faulty segment e_{f}. The hypothesis $\mathcal{H}_k$ corresponds to a fault located in the edge $e_k \in L$. This allows to further reduce the unknown parameter to $\mathbf{p} = (d_{\mathrm{f}}, R_{\mathrm{f}})$, as in the case of non hybrid-lines. For each hypothesis, an estimate of the fault parameters can thus be evaluated as presented in Section 3.1.3.

For each hypothesis $\mathcal{H}_k$, the algorithm can confirm or deny that the fault is located on the segment e_k using the validity (3.24) and accuracy tests (3.37).

It may happen that, after considering n measurements, more than one hypothesis $\mathcal{H}_k$ that segment e_k is faulty are deemed valid. Assuming there is a single fault, the algorithm must decide which segment is actually faulty among those which satisfied the decision logic. The hypothesis with the smallest cost (3.17) is chosen, *i.e.*, the estimated faulty line is

$$\widehat{e}_{\mathrm{f}} = \min_{e_k} \{ c^{(n)}(\mathbf{p}, e_k) | \text{fault is identified on } e_k \} \tag{3.38}$$

If the fault is never identified on any of the segments $(e_1, \ldots, e_m)$ after reaching the maximum length of the measurement window, the algorithm concludes the line L is healthy. The fault may either be located elsewhere in the grid or be non existent.

3.3. Implementation and tuning considerations

This section presents further details on the implementation and tuning of the fault identification algorithm for which a pseudo-code is provided in Table 3.1. The fault identification algorithm uses an increasing amount of measurements as they arrive at the relay.

3.3.1. Adjustment of the detection time

The detection time at sub-station q, $t_{\mathrm{d},q}$ is the arrival time of the first traveling wave from the fault. As seen in (3.1), it is related to the fault instant t_{f} by

$$t_{\mathrm{d},q} = t_{\mathrm{f}} + \frac{d_{\mathrm{f},q}}{c_{\mathrm{w}}}$$

Algorithm 3.1 Pseudo-code of the fault identification algorithm

1: Input: n_0, $\mathbf{p}_0$, $n_{\max}$, κ
2: Output: fault in protected zone, fault parameters
3: **if** Unusual behavior detected **then**
4: Collect $n = n_0$ current and voltage measurements $\mathbf{y}(t_1), \ldots, \mathbf{y}(t_n)$
5: Initialize $\mathbf{p}_0^{(n)} = \mathbf{p}_0$
6: **while** $\tau < \tau_{\max}$ **do**
7: **for** $k = 1 : \kappa$ **do**
8: Compute cost $c^{(n)}(\mathbf{p}_k^{(n)})$ and Jacobian $\left[J_{\mathrm{f}}^{(n)}(\mathbf{p}_k^{(n)})\right]$ using (3.17)
9: Compute $\boldsymbol{\delta}_k^{(n)}$ using (3.22)
10: Update $\mathbf{p}_{k+1}^{(n)} = \mathbf{p}_k^{(n)} + \boldsymbol{\delta}_k^{(n)}$
11: Compute cost $c^{(n)}(\mathbf{p}_{k+1}^{(n)})$
12: **if** $c^{(n)}(\mathbf{p}_{k+1}^{(n)}) < c^{(n)}(\mathbf{p}_k^{(n)})$ **then**
13: Accept $\mathbf{p}_{k+1}^{(n)}$ and decrease μ
14: **else**
15: Reject $\mathbf{p}_{k+1}^{(n)}$, increase μ and go back to line 9
16: **end if**
17: Compute Jacobian $\left[J_{\mathrm{f}}^{(n)}(\mathbf{p}_{k+1}^{(n)})\right]$ and confidence region $\mathcal{R}^{(\alpha)}(\mathbf{p}_{k+1}^{(n)})$
18: **if** *stop_cond*$(\mathbf{p}_{k+1}^{(n)}, \mathcal{R}^{(\alpha)}(\mathbf{p}_{k+1}^{(n)}))$ =true **then**
19: Fault is internal **return** (true, $\mathbf{p}_{k+1}^{(n)}$)
20: **end if**
21: **end for**
22: Collect Δn new measurements $\mathbf{y}(t_{n+1}), \ldots, \mathbf{y}(t_{n+\Delta n})$
23: Set new initial point $\mathbf{p}_0^{(n+\Delta n)} = \mathbf{p}_\kappa^{(n)}$
24: $n \leftarrow n + \Delta n$
25: **end while**
26: Fault is external **return** (false, $\emptyset$)
27: **end if**

where c_{w} is the wave speed. Considering that $t_{\mathrm{d},q}$ can be accurately measured, the time instant can be deduced from the fault distance and thus is not a free unknown parameter. This allows to reduce the vector of the fault parameters to $\mathbf{p} = (R_{\mathrm{f}}, d_{\mathrm{f}})$ in Section 3.1.1. In practice, a first estimate of the detection time is obtained as the first instant for which the voltage derivative goes above a threshold, *i.e.*,

$$t_{\mathrm{d},0} = \min_{t_j} \left\{ t_j : \left| \frac{\Delta v(t_j)}{\Delta t} \right| > \Delta v_{\mathrm{th}} \right\}. \tag{3.39}$$

A Savitzky-Golay filter [Savitzky and Golay, 1964] may be used to compute a numerical derivative of the voltage that is less sensitive to noise. Nevertheless, depending on the initial slope of the wave and the tuning of the threshold, the detection time may be slightly erroneous. Even though the difference is only of a few time

steps, it may induce an important error in the cost function (3.17) and may degrade the estimation of the remaining parameters. In order to fine tune the detection time, a separated least squares approach is adopted. Consider a quadratic cost that represents the mismatch between the modeled voltage and the received measurements,

$$c_{\mathrm{d}}\left(t, t_{\mathrm{d}}\right) = \frac{1}{2}\left(v^{\mathrm{m}}\left(t, t_{\mathrm{d}}\right) - v(t)\right)^2$$

The detection time is estimated using a gradient descent algorithm, *i.e.*,
$t_{\mathrm{d},k+1} = t_{\mathrm{d},k} - \lambda_k \left.\frac{\partial c_{\mathrm{d}}}{\partial t_{\mathrm{d}}}\right|_{t_{\mathrm{d},k}}$.

The line search term λ_k can be adjusted using the steepest descend method. An alternative approach can be to force the fault detection instant to be shifted forward or backward by only one time step per iteration. This prevents the detection time to oscillate around its true value, at the expense of a slower convergence of the detection time. This is acceptable because the first estimate of the detection time (3.39) is usually only a few time steps away from its true value. Illustrations of the estimation of the detection time are provided in Chapter 4.

3.3.2. Implementation details

The model employed in Sections 3.1 and 3.2 is the *parametric model* $\left(\mathbf{v}^{\mathrm{m}}\left(t, \mathbf{p}\right), \mathbf{i}^{\mathrm{m}}\left(t, \mathbf{p}\right)\right)$ developed in Chapter 2. For the purpose of parameter estimation, per unit (p.u.) values are used for the voltage and current quantities. This allows the voltage and current to have variations of the same order of magnitude. In the case of a multiconductor transmission line, one can employ either the modal or phase quantities. The former are preferred as modal quantities present variations of similar amplitudes whereas for phase quantities the faulty pole experiences larger variations than the healthy pole.

The fault identification algorithm, as presented in 3.1, requires for a given number of measurements (n) the evaluation of 2κ cost functions and Jacobian matrices (with the exception of the loop for the tuning of the damping parameter μ). However, this can be reduced as the cost $c^{(n)}\left(\mathbf{p}_{k+1}^{(n)}\right)$ evaluated in line 11 can be employed to evaluate the cost at the next iteration $c^{(n)}\left(\mathbf{p}_{k+1}^{(n)}\right)$, at line 8. The same applies to the computations of the Jacobian matrices at lines 17 and 8. This reduces the number of evaluations of $c^{(n)}$ and $\left[J_{\mathrm{f}}^{(n)}\right]$ that must be performed for n measurements to $\kappa + 1$.

Consider now the arrival of Δn new measurements. The cost $c^{(n+\Delta n)}\left(\mathbf{p}_0^{(n+\Delta n)}\right)$ and Jacobian $\left[J_{\mathrm{f}}^{(n+\Delta n)}\left(\mathbf{p}_0^{(n+\Delta n)}\right)\right]$ should be evaluated on the total measurement set (line 8). However, those quantities are already known for the n previously available

measurements, *i.e.*,

$$c^{(n)}\left(\mathbf{p}_0^{(n+\Delta n)}\right) = c^{(n)}\left(\mathbf{p}_\kappa^{(n)}\right)$$

and

$$\left[J_\mathrm{f}^{(n)}\left(\mathbf{p}_0^{(n+\Delta n)}\right)\right] = J_\mathrm{f}^{(n)}\left(\mathbf{p}_\kappa^{(n)}\right).$$

This is possible as the initial vector for the updated data set $\mathbf{p}_0^{(n+\Delta n)}$ is the last obtained estimate from the previous data set $\mathbf{p}_\kappa^{(n)}$. Thus, the cost and Jacobian need only to be evaluated on the Δn new measurements (line 8). An implementation that takes advantage of this feature allows to reduce the complexity when considering a large amount of data as some of the computations need only to be performed on the last Δn points.

3.3.3. Tuning of the parameters of the algorithm

The estimation algorithm as well as the identification tests require the tuning of multiple parameters introduced in Sections 3.1 and 3.2 and recalled in the Algorithm 3.1. Those parameters are gathered in the Table 3.1 with their typical value. The different settings are further discussed in this section.

Parameter	Meaning	Chosen value
τ_max	Maximum duration of the measurement window	1 ms
Δn	Size of the batch of additional samples	10
κ	Number of iterations before processing of an additional batch of Δn samples.	1
f_s	Sampling frequency of the measurements	1 MHz
$\mathbf{p}_0$	Initial value of the fault parameter vector.	$(5\,\Omega,\ 5\,\mathrm{km})$
$\lambda_\mathrm{v}^1, \lambda_\mathrm{v}^2, \lambda_\mathrm{i}^1, \lambda_\mathrm{i}^2$	Relative weights used in the cost function (3.17)	$(0.2, 0.8, 0.2, 0.8)$
$\mu_0,\ \mu_\uparrow,\ \mu_\downarrow$	Initial value, increasing and decreasing factors for the damping parameter μ	$(10^1,\ 2,\ 5)$
c_{95}	Threshold for the 95% confidence region	10
$d_\mathrm{f,min}, d_\mathrm{f,max}$	Minimum and maximum plausible values for d_f	$(0,\ 0.9d)$
$R_\mathrm{f,min}, R_\mathrm{f,max}$	Minimum and maximum plausible values for R_f	$(0,\ 200\,\Omega)$

Table 3.1.: Parameters of the fault identification algorithm and their typical value.

3.3.3.1. Iterations / added samples trade off

The maximum number of iterations (3.23) that can be performed by the algorithm depends on $\tau_{\max}$, $\Delta n, \kappa$ and of the sampling frequency f_s. The latter is considered as a fixed parameter $f_s = 1\,\text{MHz}$ depending on the sensor characteristics. The influence of f_s is studied along with the sensor characteristics in Section 5.4.2.

The maximum duration of the measurement window $\tau_{\max}$ can be considered as an external requirement on the maximum time available to identify the faults. Another approach is to consider whether enough information is contained in a time window $\tau_{\max}$ to allow the algorithm to conclude. The arrival time difference between different waves contains decisive information on the fault distance. Thus a possible indicator for the window length to consider is the arrival time of the second TW, either due to a reflection at the fault location or at the remote station. Its arrival delay (with respect to the arrival of the first wave) is, depending on its the fault distance d_f,

$$\min\left(\frac{2d_f}{c_w}, \frac{2(d-d_f)}{c_w}\right) \leq \frac{d}{c_w}$$

Thus, a measurement window of length

$$\tau_{\max} = \frac{d}{c_w} \tag{3.40}$$

allows to always take into account the second TW. For a line 300 km long, the maximum observation window is thus of 1 ms which is compatible with the requirements on the fault identification speed. In practice, it is observed the fault identification algorithm generally requires much less data to identify internal faults.

The tuning of the number of added samples Δn and the number κ of iterations until a new batch of samples is taken into account results in a trade-off between the speed and the accuracy of the algorithm. In the perspective of a real-time implementation, the algorithm should be able to perform κ iterations in $\frac{\Delta n}{f_s}$. Thus a large κ or small Δn would slow down the algorithm. On the other hand, a small κ prevents the optimization algorithm to reach the minimum of the cost function. The selected compromise is $\Delta n = 10$ and $\kappa = 1$.

3.3.3.2. Levenberg-Marquardt tuning

The Levenberg-Marquardt presented in Section 3.1.3 requires to specify several parameters such as

- the initial point $\mathbf{p_0}$,
- the cost function weights λ_v, λ_i,
- the damping parameters μ_0 and their multiplying factors $\mu_\uparrow, \mu_\downarrow$.

The setting of the initial point $\mathbf{p_0} = (R_{\mathrm{f},0},\, d_{\mathrm{f},0})$ should allow the algorithm to converge to the true value of fault parameter vector as fast as possible, for any possible fault cases. The faults that should be identified the fastest (leading to the fastest current rise at the station) are the closest ones with a low impedance. Initializing the algorithm in this area speeds up the convergence for these severe faults.
In addition, the initial distance and number of points should not lead to the computation of a second wave due to the reflection at the fault location. This is preferable as such reflected waves increase the computational complexity of the model so that their evaluation should be avoided when possible. Thus it should hold

$$\underbrace{\frac{n_0}{f_\mathrm{s}}}_{\text{Initial window length}} < \underbrace{\frac{2d_{\mathrm{f},0}}{c_\mathrm{w}}}_{\text{Delay of the modeled second wave}}.$$

The initial number of points n_0 is typically set to Δn. We selected a close distance low impedance initial point such as $(R_{\mathrm{f},0},\, d_{\mathrm{f},0}) = (5\,\Omega, 5\,\mathrm{km})$. However, it has been observed that the choice of the initial point only has a small impact on the convergence of the estimation algorithm. A more detailed analysis of the impact of the initial point on the performances of the algorithm is provided in Appendix D.

The parameters $\lambda_\mathrm{v}^1, \ldots, \lambda_\mathrm{v}^{n_\mathrm{c}}, \lambda_\mathrm{i}^1, \ldots, \lambda_\mathrm{i}^{n_\mathrm{c}}$ of the weight matrix (3.18) define the relative importance of the voltage and current in the minimization of the cost function (3.17). In the case of a multi-conductor, different weights can also be applied to the different modes. The settings of the weights can be guided by the estimate of the covariance matrix (3.15). Considering the numerical example (3.35) the ratios between the diagonal entries are

$$\begin{aligned} \frac{\left[\widehat{\Sigma}\right]_{3,3}}{\left[\widehat{\Sigma}\right]_{4,4}} &\simeq 3.8 \qquad \frac{\left[\widehat{\Sigma}\right]_{1,1}}{\left[\widehat{\Sigma}\right]_{2,2}} \simeq 4.1 \\ \frac{\left[\widehat{\Sigma}\right]_{3,3}}{\left[\widehat{\Sigma}\right]_{1,1}} &\simeq 3.1 \qquad \frac{\left[\widehat{\Sigma}\right]_{2,2}}{\left[\widehat{\Sigma}\right]_{4,4}} \simeq 2.9 \end{aligned} \tag{3.41}$$

which suggests to employ a ratio of 4 between the ground and aerial weights and a ratio of 3 between the current and voltage weights, *i.e.*,

$$\begin{aligned} \frac{\lambda_\mathrm{i}^2}{\lambda_\mathrm{i}^1} &\simeq 3.8 \qquad \frac{\lambda_\mathrm{v}^2}{\lambda_\mathrm{v}^1} \simeq 4.1 \\ \frac{\lambda_\mathrm{i}^1}{\lambda_\mathrm{v}^1} &\simeq 3.1 \qquad \frac{\lambda_\mathrm{i}^2}{\lambda_\mathrm{v}^2} \simeq 2.9 \end{aligned} \tag{3.42}$$

In practice, the same weights were applied to the current and voltage quantities, only using different weights for the different modes. The weight matrix employed

in the bipolar case is thus

$$[W] = \begin{pmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.8 \end{pmatrix}.$$

The larger weights on the aerial modes may be explained as this part is modeled using only the physical approach whereas the ground modes employ a behavioral part more prone to model errors.

In the mono-conductor case the differences between the ground and aerial modes no longer apply. A ratio close to the one suggested in (3.42) showed to be appropriate, leading to $\lambda_{\mathrm{v}} = 0.2$, $\lambda_{\mathrm{i}} = 0.8$.

The tuning of the damping parameter μ follows the guidelines for the direct method such as presented in [Transtrum and Sethna, 2012]. In particular, the increasing factor $\mu_{\uparrow}$ is set to be less than the decreasing factor $\mu_{\downarrow}$, *e.g.*, $\mu_{\uparrow} = 2$, $\mu_{\downarrow} = 5$. The initial value μ_0 should prevent that large steps are performed with few available data, we selected $\mu_0 = 10^1$.

3.3.3.3. Identification tests tuning

The two tests of the decision logic of Section 3.2, namely the validation and accuracy tests, must be tuned adequately to ensure the correct distinction between internal faults and external faults. The validity domain (3.24)

$$\mathcal{D}_{\mathbf{p}} = \{(d_{\mathrm{f}}\,, R_{\mathrm{f}}) | d_{\mathrm{f,min}} \leq d_{\mathrm{f}} \leq d_{\mathrm{f,max}}, R_{\mathrm{f,min}} \leq R_{\mathrm{f}} < R_{\mathrm{f,max}}\} \tag{3.43}$$

defines the faults that should actually be identified by the algorithm. The minimum values $d_{\mathrm{f,min}} = 0$ and $R_{\mathrm{f,min}} = 0$ simply reflect that the fault distance and resistance should be positive.

The maximum fault resistance $R_{\mathrm{f,max}}$ defines the highest expected fault impedance that should be considered. While high fault impedance may occur, they do not lead to an important fault current and thus do not require fast action. On the other hand, setting a maximum value allows to identify as external many faults occurring elsewhere in the grid. A value $R_{\mathrm{f,max}} = 200\,\Omega$ is typically used for overhead lines whereas $R_{\mathrm{f,max}} = 10\,\Omega$ is preferred for underground cables where fault resistances are generally very low.

The maximum fault distance $d_{\mathrm{f,max}}$ is set as a proportion of the protected line length d. Idealy $d_{\mathrm{f,max}} = d$ which means 100 % of the line is within the protection zone of the relay. In practice, $d_{\mathrm{f,max}} = 0.8d$ or $d_{\mathrm{f,max}} = 0.9d$ is beneficial as many external faults can be interpreted as internal faults occurring at the remote end of the protected line. Hence, to renounce to protect the remote 10 % or 20 % of the line allows

one to considerably improve the security against external faults of the algorithm. Those remote faults would however be detected by the opposite relay protecting the same line. This other relay can control the inter-tripping of the first relay. As the current rise is not that important for remote faults, the recourse to communication is acceptable in such cases. The system integration of the fault identification algorithm is more extensively investigated in Chapter 5.

The accuracy test requires the setting of the value α for which the confidence region is evaluated and the threshold $c_{1-\alpha}$ on the confidence region size (3.37). The setting of the two values are related, we first fixed $\alpha = 0.05$, which leads to 95% confidence regions. The tuning of the threshold can be done heuristically considering a large amount of faults. Alternatively, one can focus on the critical cases that are the remote internal faults of high impedance and the external faults of low impedance occurring close the extremity of a line. Simulation results presented in Chapter 4 show however that the algorithm is quite robust to the chosen value of the threshold. This is also confirmed by the fact that the same threshold can be applied to relays protecting different lines in the network.

3.4. Conclusion

In this chapter, the fault identification algorithm based on the estimation of the fault parameters is presented. The use of one of the parametric models from Chapter 2 is crucial as it allows to simulate easily the voltage and current evolution for different fault characteristics. Assuming the line it protects is faulty, each relay estimates the fault parameters using the measurements that arrive successively from the sensors. The algorithm can thus be started as soon as the first data are available, and more and more measurements are taken into account when available in order to improve the estimate of the fault parameters.
The maximum likelihood estimate is computed using a Levenber-Marquardt optimization algorithm. The results of the parameter estimation allow the algorithm to confirm or deny that the protected line is faulty. In particular, an approximation of the 95% confidence region of the estimated fault parameters is employed to assess the estimation accuracy. The behavior of the estimation algorithm is illustrated in Chapter 4 on particular examples and the proposed fault identification is evaluated on extensive simulations.

4. Simulation results

The simulation results of the fault identification algorithm detailed in Chapter 3 are presented in this chapter. The methodology adopted for the simulations is detailed in Section 4.1. The three main application cases of the proposed approach are treated successively, the rigid bipole in Section 4.2, the monopole in Section 4.3 and the hybrid lines in Section 4.4.

4.1. Simulation methodology

In order to test the fault identification algorithm presented in Chapter 3 a Matlab implementation has been performed.. The transient behavior of the grid in case of a fault is reproduced using the EMT simulation software [Mahseredjian et al., 1993] EMTP-RV. The EMT simulations are always performed with a time step of 1 μs. The tests are performed open loop: the grid is first entirely simulated in EMT software, and the obtained data are then given to the fault identification algorithm in Matlab. The self-protection of the MMC stations is however included in the EMT models [Auran, 2017]. The outcome of the identification hence does not impact the behavior of the grid. In particular, the possible tripping of the circuit breakers is not simulated. This is not damageable since the fault identification algorithm only uses transient data obtained in two milliseconds after the fault occurrence.

The proposed grid for the simulation is a four station meshed grid adapted from the PROMOTIoN project, see [PROMOTIoN WP4, 2018]. A general representation of the grid is depicted in Figure 4.1. Each of the four transmission lines is protected by two relays located at the extremity of each line. Each relay controls the opening of the DC line breakers located at the end of each line, which are not represented in Figure 4.1 for readability. While DC reactors are generally considered as part of the breakers, they are omitted in the implemented test grid as they are not required by the fault identification algorithm. The compliance of this hypothesis with a full-selective fault clearing strategy is addressed in Chapter 5. The transmission lines can be monoconductor or multiconductor overhead lines as well as hybrid lines comprising also portions of underground cables, depending on the considered application case. The characteristics of the three different considered type of lines: OHL mono-conductor, OHL bipolar and hybrid lines, are detailed in Appendix A. The transmission lines are simulated using the wideband model and The MMC sub-stations are simulated with the model: MMC arm equivalent switching function

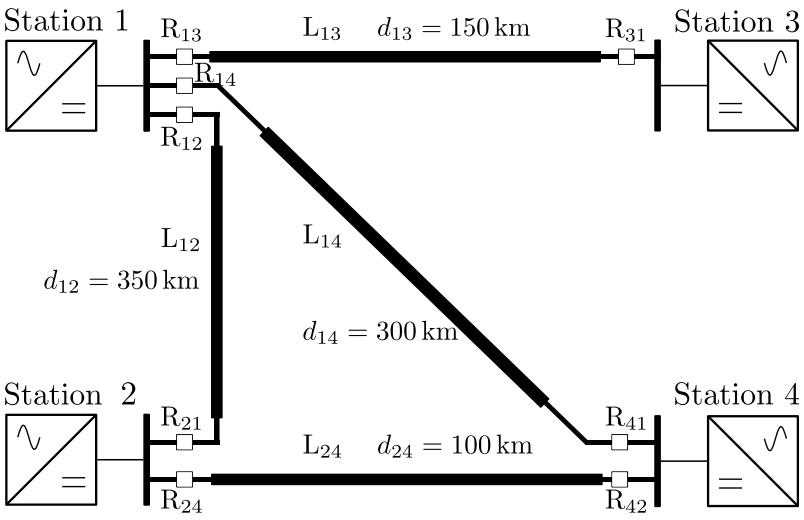


Figure 4.1.: General scheme of the 4 station meshed grid simulated in EMT software to test the proposed fault identification algorithm. .

[Saad et al., 2014] (also known as model 3) whose parameters are also provided in given in Appendix A. In the bipolar scheme, each converter station is composed of two identical MMC grounded by a small resistor ($0.5\,\Omega$).
In this chapter, we assume ideal measurements are available for the current and voltage at each extremity of the lines. The sampling frequency of the measurements $f_0 = 1\,\text{MHz}$ is identical to the simulation step used for the EMT simulations. In Section 5.4, a detailed sensor model is introduced and the impact of more realistic measurement errors on the identification algorithm is investigated.

Various indicators are employed to evaluate the performances of the fault identification algorithm. We evaluate at a given relay,

- the *dependability*, *i.e.*, the ability of the algorithm to identify all the faults occurring on the protected line
- the *security*, *i.e.*, the ability to not to identify faults when the protected line is healthy [Van Hertem et al., 2016].

Reliability refers to the addition of both properties. Assuming the computing can be performed in real time, the protection speed reduces to the length of the measurement window required for the fault identification.The validity of this assumption is questioned in Section 4.2.2.3 where the computing times are discussed.

Though the main purpose of the algorithm is the reliable identification of faults, the accuracy of the estimated fault parameters is also analyzed using relative errors. The relative error of the fault distance with respect to the protected line length d_e is employed, *i..e*,

$$\varepsilon_{\text{d}} = \left(\widehat{d}_{\text{f}} - d_{\text{f}}^{*}\right)/d_e. \tag{4.1}$$

The relative error for the fault resistance is given with respect to the base DC impedance ($Z_{\text{base}} = 102\,\Omega$),

$$\varepsilon_{\text{R}} = \left(\widehat{R}_{\text{f}} - R_{\text{f}}^{*}\right) / Z_{\text{base}}. \tag{4.2}$$

4.2. Bipolar configuration

This section presents the results of the proposed fault identification algorithm in the case of a rigid bipole configuration. The behavior of the fault identification algorithm is first illustrated on particular examples in Section 4.2.1, where internal and external fault cases are both detailed. More extensive simulations are presented in Section 4.2.2 as well as a comparison with a state of the art method. A sensitivity analysis is then conducted in Section 4.2.3, where the impact of a mismatch in the soil resistivity is evaluated.

4.2.1. Illustrative examples

Two particular examples are detailed in this section. The fault identification of internal faults is first detailed in Section 4.2.1.1. The behavior of the algorithm when the network is affected by disturbances not causing a fault is addressed in Section 4.2.1.2 considering the tripping of a line in the network. A transmission line composed of two portions of significantly different soil resistivities is considered in Section 4.2.1.3 and the adaptive ground filter proposed in Section 2.2.2.2 is applied.

4.2.1.1. Selective fault identification

Consider a pole-to-ground fault occurring on the line L_{12} within the 4-station grid of Figure 4.1. The fault is located at $d_{\text{f}}^{*} = 140\,\text{km}$ from station 2 and has an impedance of $R_{\text{f}}^{*} = 80\,\Omega$. The ability of the algorithm to identify the fault at the relay R_{21} is studied, though the fault identification algorithm is triggered at all relays of the grid using a ROCOV criterion.
After the fault is suspected at the relay R_{21} the fault parameters are estimated using the first available measurements from the voltage and current sensors. The estimation algorithm stops and identifies the fault as internal after having performed 8 iterations, corresponding to an observation window of $0.87\,\mu s$. The evolution of the estimated parameters as well as the contour of the cost function are displayed in Figure 4.2 at iterations 2, 4, 6, and 8. The estimated fault parameters after the 8^{th} iteration are $\widehat{R}_{\text{f}} = 89\,\Omega$ and $\widehat{d}_{\text{f}} = 126\,\text{km}$, which represents a relative error of $\varepsilon_{\text{R}} = 9\%$ for the fault resistance and $\varepsilon_{\text{d}} = 4\%$ for the fault distance, see (4.1), (4.2). While the estimated parameters get closer to the minimum of the cost function, the confidence ellipse becomes smaller. The area of the confidence region eventually goes below

the predetermined threshold after the 8$^{\text{th}}$ iteration, see Figure 4.4. As the estimated fault parameters also satisfy the validity test ($0 \leq \widehat{d}_{\text{f}} \leq d_{\max} = 0.9 d_{12} = 315\,\text{km}$ and $0 \leq \widehat{R}_{\text{f}} \leq R_{\max} = 250\,\Omega$), see (3.24), the fault is identified on the line L_{12} by the relay R_{21}.

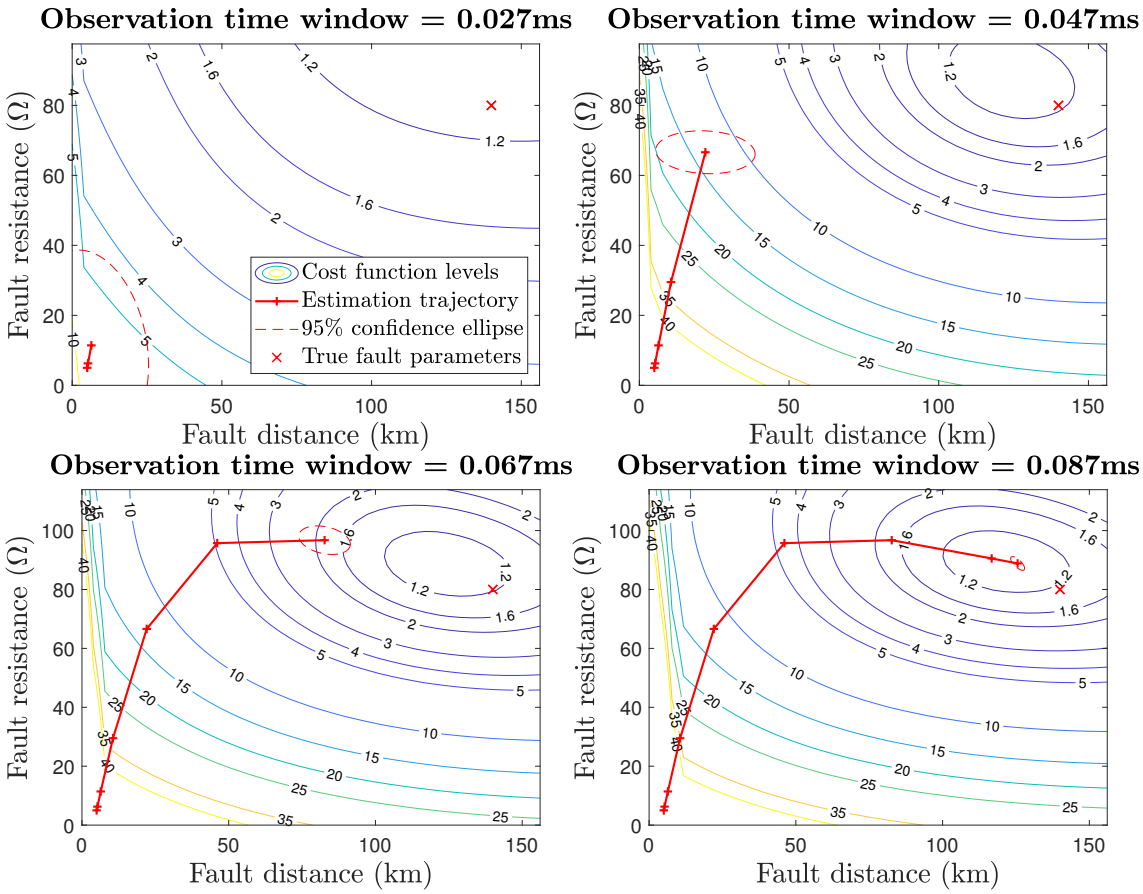


Figure 4.2.: Evolution of the estimate of the fault parameters and contour plots of the cost function at iterations 2, 4, 6, and 18. The pole-to-ground fault is located at $d_{\text{f}}^{*} = 140\,\text{km}$ from the station with an impedance of $R_{\text{f}}^{*} = 80\,\Omega$. The estimated fault parameters after 8 iterations are $\widehat{d}_{\text{f}} = 126\,\text{km}$ and $\widehat{R}_{\text{f}} = 89\,\Omega$.

The evolution of the estimate of the fault parameters at each step is plotted in Figure 4.3. The estimate of the resistance converges slightly earlier then that of the distance. This may be due to the absence of information on the fault distance contained in the aerial mode, which arrives before the ground mode. The estimate of the detection time (Figure 4.3, right) only evolves during the 5 first iterations, after which it remains constant.

The EMT simulated voltages and currents are compared in phase and modal domain with the output of the parametric model using the obtained parameters $\widehat{R}_{\text{f}} = 89\,\Omega$ and $\widehat{d}_{\text{f}} = 126\,\text{km}$, see Figure 4.5 and Figure 4.6, respectively. Only the first TW is represented as it was not required to wait for the subsequent waves for the algorithm

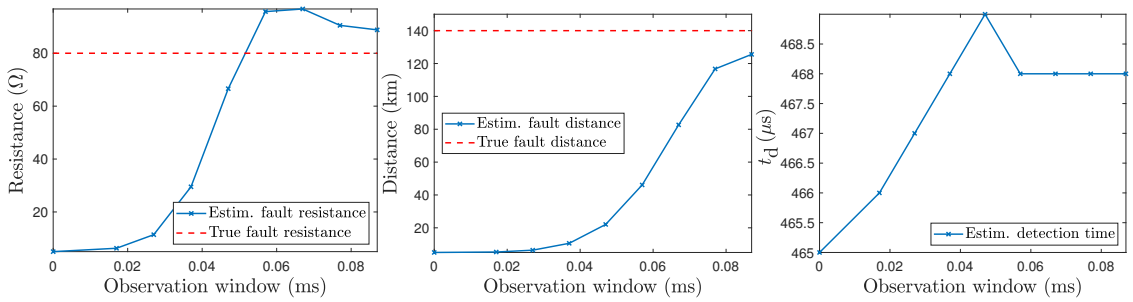


Figure 4.3.: Evolution of the estimated fault parameters at each iteration of the optimization algorithm. The estimated fault resistance (left) and distance (middle) are compared with the true values. The detection time (right) is adjusted so that the voltage wavefront of the EMT data matches the one of the first modeled TW, as detailed in Section 3.3.1.

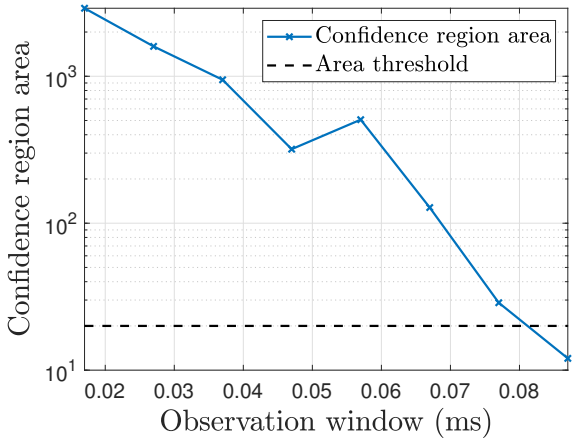


Figure 4.4.: Evolution of the 95% confidence area at each iteration step of the optimization algorithm. The fault identification stops when the area goes below the threshold (here $c_{95} = 20$).

to identify the fault.
Despite a mismatch with the EMT simulation, the parametric model captures most of the features of the voltage and current behavior. In particular, the characteristics of the different modes, which are difficult to distinguish in Figure 4.5, are well depicted in Figure 4.6. Especially, the difference in the arrival times of the ground and aerial modes, which depends on the traveled distance, improves the estimation of the fault distance. In the general case it is not possible to observe separately the ground and aerial modes at the station since they are combined through the reflection matrix (2.29). In this case however, the particular configuration of the 2-conductor rigid bipole implies the station does not interfere with the propagation modes for the first traveling wave[1]. This explains why one can clearly see the ground and aerial modes in Figure 4.5.

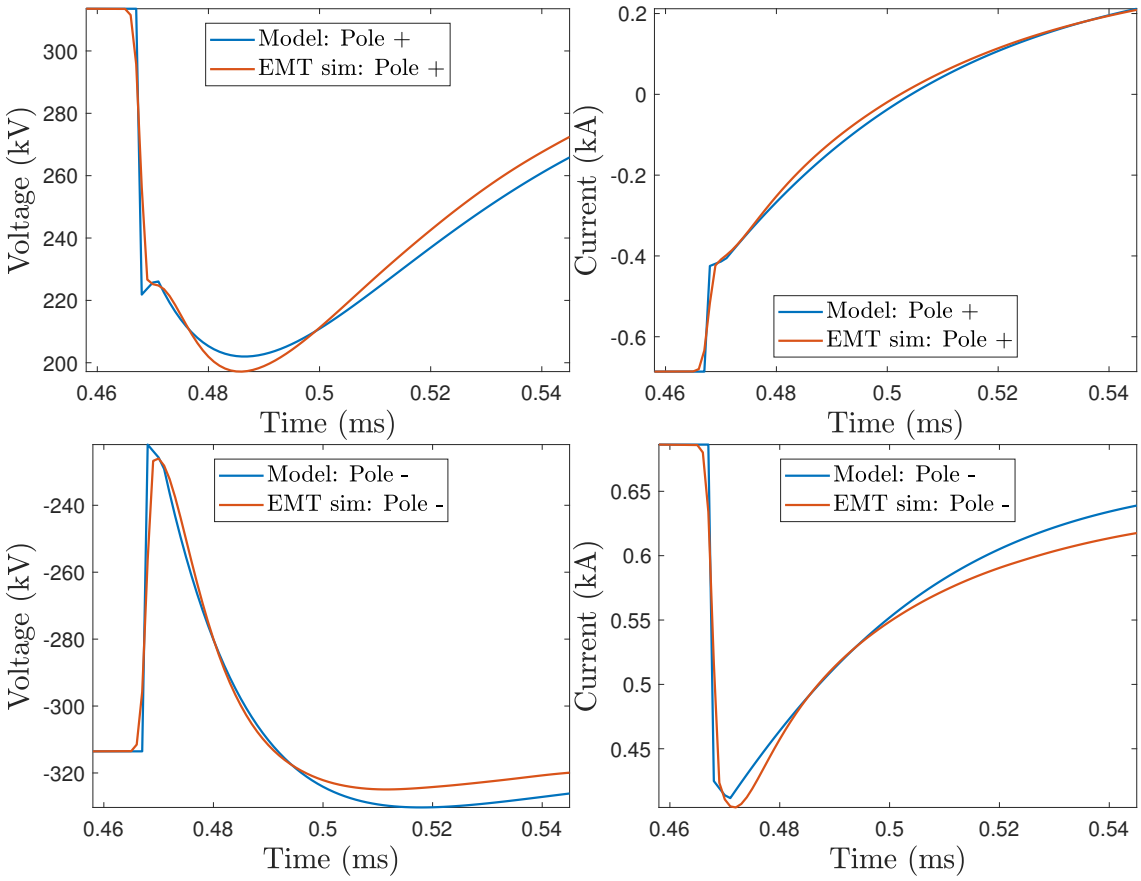


Figure 4.5.: Comparison of the modeled and simulated positive (top) and negative (bottom) pole voltages (left) and current (right) at the relay R_{12}. The fault parameters used in the parametric model are those obtained after 8 iterations: $\widehat{R}_{\mathrm{f}} = 89\,\Omega$ and $\widehat{d}_{\mathrm{f}} = 126\,\mathrm{km}$.

[1]Formally the reflection matrix at the station can be diagonalized by the same transformation matrix $[T_V]$ than the propagation matrix. This is a particular case of the rigid bipole and, for instance, would not apply to a 3 conductor system.

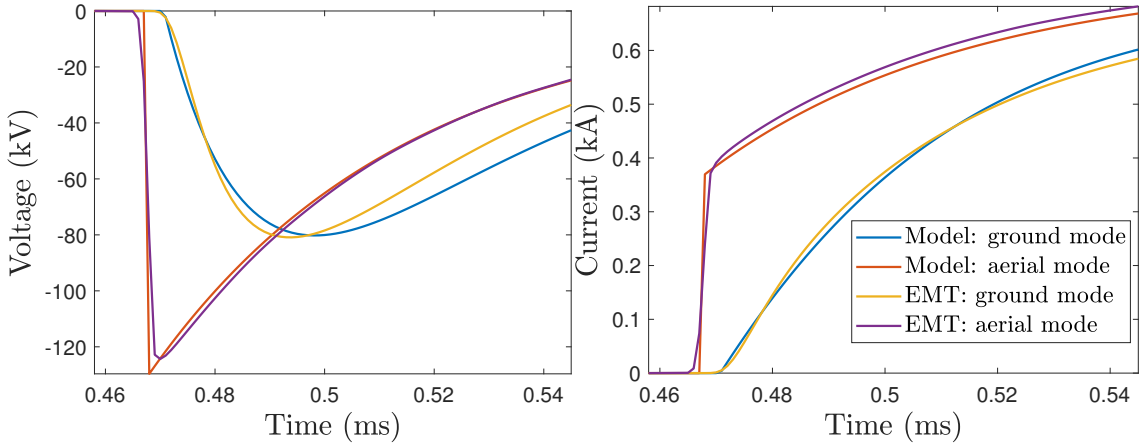


Figure 4.6.: Comparison of the modeled and simulated modal voltage (left) and current (right) at the relay R_{12}. The fault parameters used in the parametric model are the ones obtain after 8 iterations: $\widehat{R}_{\mathrm{f}} = 89\,\Omega$ and $\widehat{d}_{\mathrm{f}} = 126\,\mathrm{km}$.

The example presented in this section is representative of the behavior of the algorithm in most cases of internal faults. The extensive simulations performed in Section 4.2.2 show the algorithm presents similar performance on a wide range of fault cases.

The considered fault which affects the line L_{12} also triggers the fault identification algorithm at the other relays of the grid which detect an abnormal behavior. The two relays that protect the faulty line should identify the fault (dependability) whereas the remaining six relays of the grid should not identify the fault (security). The evolution of the accuracy criterion is compared for all the relays in Figure 4.7. The evolution of the area of the confidence ellipsoid $\mathrm{vol}\,(\mathcal{R}^{\alpha}(\mathbf{p}))$ used in the accuracy test is plotted until the maximum measurement window is reached, even if the algorithm can stop as soon as the fault is identified. Note that the algorithm implemented at the different relays run with different time window lengths. As the latter depends on the total length of the monitored line (3.40), the algorithm at relay R_{31} which monitors line L_{14} of length $d_{13} = 300\,\mathrm{km}$ uses a measurement window twice as long as the algorithm at relay R_{13} which monitors line L_{13} of length $d_{13} = 150\,\mathrm{km}$.

As the fault is close to station 2, it is first detected at the relays R_{21} (monitoring line L_{12}) and R_{24} (monitoring line L_{24}). As detailed above, the accuracy test is satisfied at relay R_{21} after 8 iterations and the fault is identified on line L_{12}. At relay R_{24} the area of the confidence region never goes below the threshold, indicating that the fault is not on the line L_{24}. At $t \simeq 0.7\,\mathrm{ms}$ the fault is successively detected by the relays R_{12}, R_{13} and R_{14} located at station 1 and by the relays R_{41} and R_{42} located at station 4. The relay R_{12} correctly identifies the fault as internal after few iterations. All other relays conclude the fault is external as the area of the confidence region associated to their own fault parameter estimate stays above the threshold. After 1.2 ms the first TW reaches the station 3, triggering the algorithm at relay

R_{31}. At R_{31} the fault is also identified as external even though the accuracy test is satisfied. The validity test is, indeed, not satisfied at this relay as the estimated fault parameters go out of the protected zone (3.43). This is evidenced in Figure 4.8 where the fault distance and resistance estimated by the algorithm at relay R_{31} are plotted. In less than 10 steps, $\widehat{d}_{\mathrm{f}} > 0.9 d_{13} = 135\,\mathrm{km}$ and $\widehat{R}_{\mathrm{f}} > R_{\max} = 250\,\Omega$, which prevents the algorithm to erroneously identify the fault as internal.

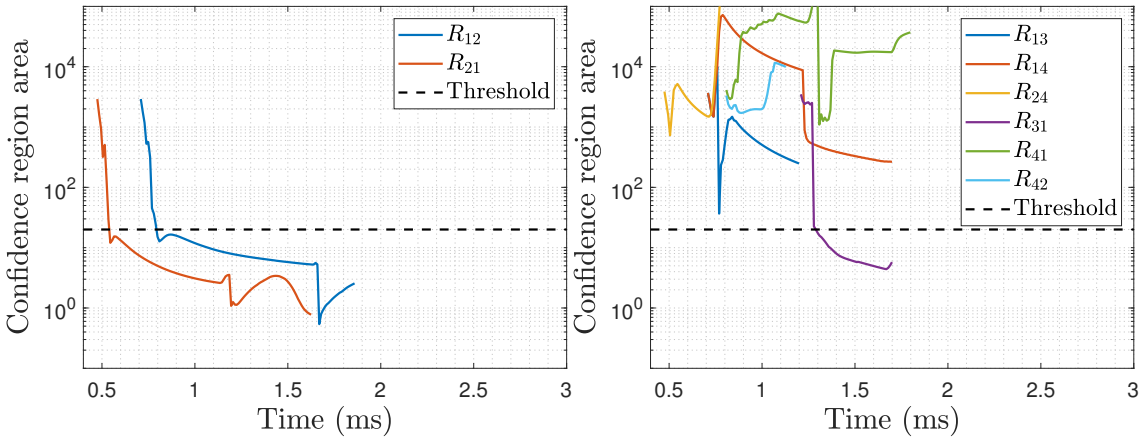


Figure 4.7.: Evolution of the accuracy criterion at the two relays (R_{12} and R_{21}) that protect the faulty line L_{12} (left) and at the remaining six relays of the grid. R_{12} and R_{21} identify the fault as internal using a short time window. The relay R_{31} correctly identifies the fault as external despite the small confidence region because the estimated fault parameters are not within the validity domain.

This example illustrates how the method identifies internal faults using very few measurements while rejecting faults occurring on neighboring lines. Furthermore, the method is inherently directional as the direction of the current is embedded in the parametric model. This is confirmed in Figure 4.7 where the relays R_{24}, R_{13}, and R_{14}, for which the fault is a reverse fault, have particularly large confidence areas.

The conclusions of this section are confirmed in Section 4.2.2 where extensive simulations are performed on a wider range of fault cases.

4.2.1.2. Response to a non faulty event

In this section, we investigate the behavior of the algorithm in when the grid may experience transient phenomena without being affected by a fault. Such situations may occur if one of the line within the grid is disconnected, or when the power reference of one or several converter stations is changed.

In this section we focus on the first example where one of the line of the grid is disconnected. In the grid of Figure 4.1 we consider the positive pole of the line L_{13} between the stations 1 and 3 is disconnected. The behavior of the grid in this

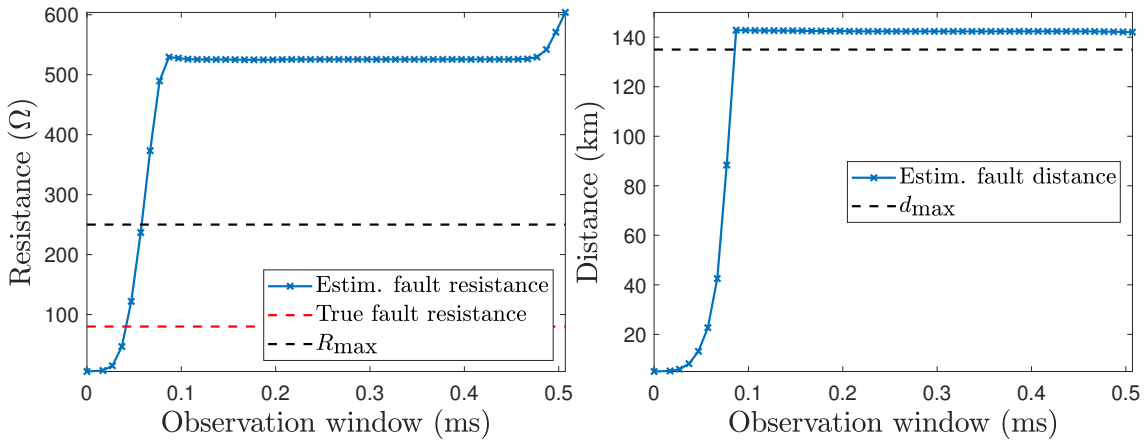


Figure 4.8.: Evolution of the estimated fault resistance (left) and distance (right) by the identification algorithm at relay R_{31}. The estimated fault parameters go out of the validity domain in less than 10 steps which prevents the identification of the fault as internal despite the accuracy test being satisfied.

situation is simulated with EMT software. The algorithms protecting the remaining lines may detect an abnormal behavior[2] but should remain secure with respect to such disturbances. We consider the behavior of the algorithm at relay R_{12}. The evolution of the estimated parameters as well as the size of the confidence region are plotted in Figure 4.9. The estimated fault resistance diverges towards high values and the confidence region remains large, leading the algorithm to conclude the line it protects, L_{12}, is healthy, after considering a measurement window of length $\tau_{\max} = 1.167\,\text{ms}$. The current and voltage evolutions obtained from the EMT simulation are compared with the output of the parametric model in Figure 4.10 using the estimated fault parameters. The transient behavior of the tripping event is significantly different from the one due to a fault, in particular considering the decrease of the current.

This particular example shows how non-faulty events may be handled by the identification algorithm. On a particular example corresponding to the tripping of a conductor, the method showed to be secure as the transient behavior after the tripping significantly differs from the one after an internal fault. This should however be studied for various disturbance cases.

4.2.1.3. Non-uniform soil resistivity

This section provides an example of the adaptation of the algorithm, and in particular the parametric model, to the case of a transmission line with non-homogeneous

[2]Such events may not be detected depending on the tuning of the ROCOV algorithm responsible for the abnormal behavior detection. For the sake of the study, we consider here the ROCOV is sufficiently sensitive to detect the considered events.

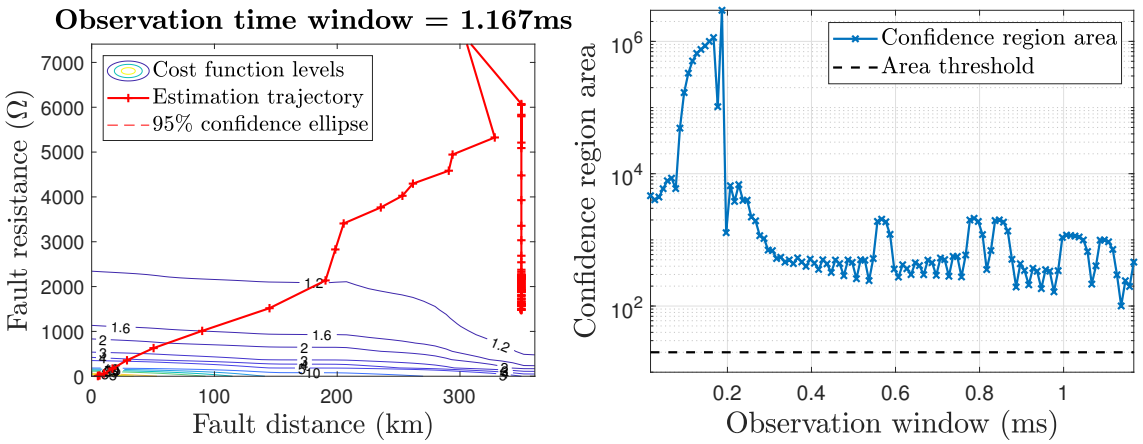


Figure 4.9.: Evolution of the estimated fault parameters and levels of the cost function (left) as well as the size of the corresponding confidence region (right).

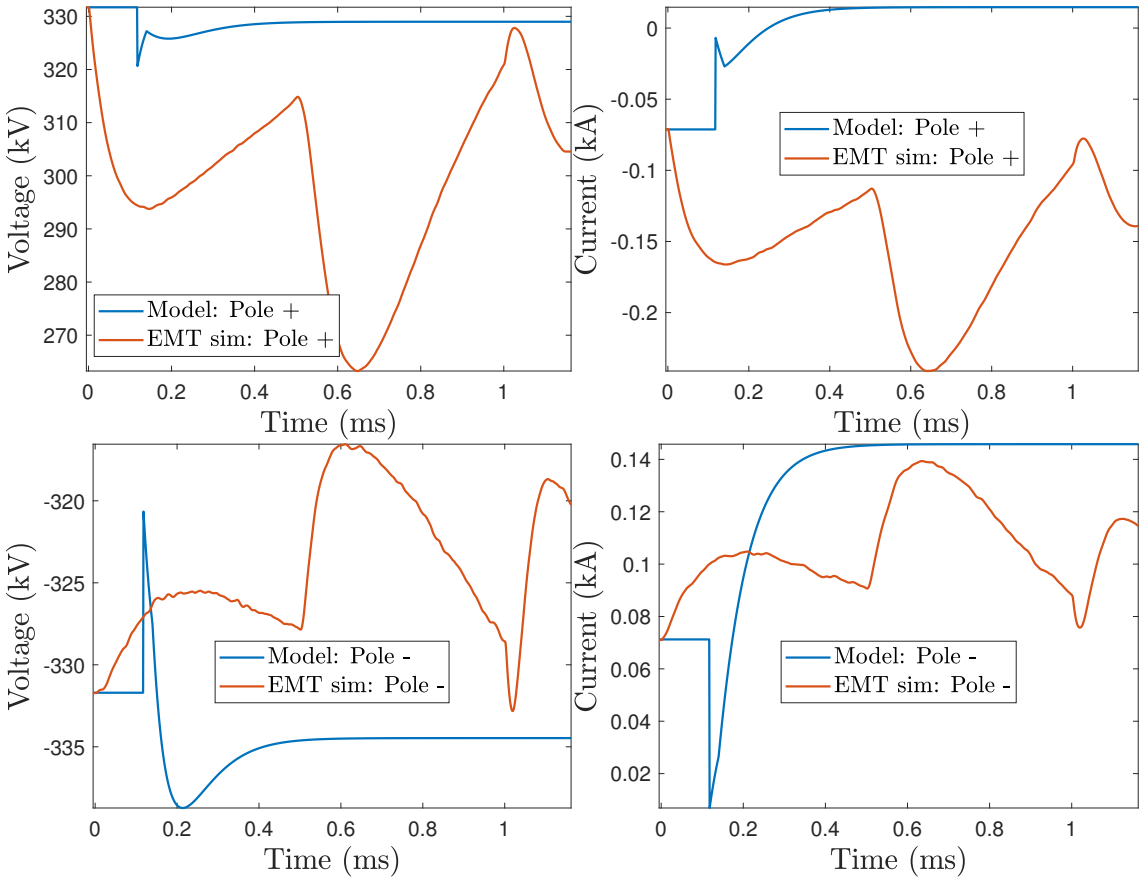


Figure 4.10.: Comparison of the modeled and simulated positive (top) and negative (bottom) pole voltage (left) and current (right) at the relay R_{12}. The parametric model describes the transient evolution of the voltage and current after a fault on the protected line whereas the EMT simulations correspond the tripping of line the positive pole of the line L_{13}.

soil resistivity, as presented in Section 2.2.2.2. In the 4-station grid of Figure 4.11, the line L_{14} consists of two portions of lengths $d_{14,1} = 70\,\text{km}$ and $d_{14,2} = 230\,\text{km}$ and respective soil resistivity $\rho_1 = 50\,\Omega\text{m}$ and $\rho_2 = 5000\,\Omega\text{m}$.

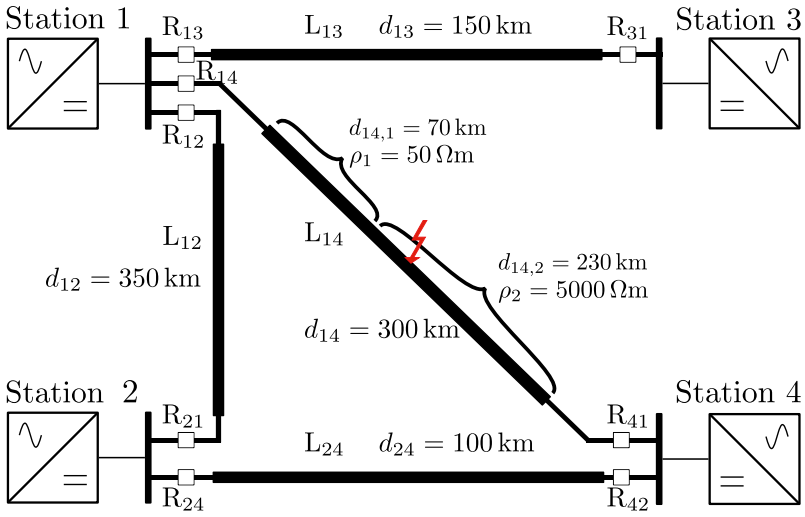


Figure 4.11.: Adapted 4 station meshed where line L_{14} comprises two portions with significantly different soil resistivity. A fault is simulated at $d_\text{f} = 100\,\text{km}$ from station 1 and the behavior of the algorithm is detailed at relay R_{14}.

A fault is simulated at $d_\text{f}^* = 100\,\text{km}$ from station 1 with an impedance of $R_\text{f}^* = 65\,\Omega$. The fault is thus located in the portion with the soil resistivity $\rho = 5000$. We consider the behavior of the algorithm at relay R_{14}. The parameter estimation stops after 9 iterations corresponding to a measurement window of $97\,\mu s$ and the fault is identified as internal. The obtained estimated fault parameters are $\widehat{R}_\text{f} = 71\,\Omega$ and $\widehat{d}_\text{f} = 95\,\text{km}$. The evolution of the estimated parameter as well as the contour plots of the cost function are plotted in Figure 4.12 at iterations 5 and 9. The discontinuity in the cost function contours due to the adaptation of the ground filter can be noticed for $d_\text{f} = 70\,\text{km}$.

The phase voltages and currents provided by the parametric model are compared with the EMT simulation in Figure 4.13 using the estimated fault parameters obtained after 11 iterations $\widehat{R}_\text{f} = 71\,\Omega$ and $\widehat{d}_\text{f} = 95\,\text{km}$. The voltage and current are however represented over a $300\,\mu\text{s}$ long window. The discontinuity in the soil resistivity only causes a very small reflected wave which can be noticed at about $t = 550\,\mu\text{s}$.

Besides the detailed fault case (F_1), two additional faults were simulated on the same non-homogeneous line to further validate the proposed approach. Fault F_2 is located at 50 km from station 1 and has an impedance of $120\,\Omega$. Fault F_3 is located at 150 km from station 1 and has an impedance of $35\,\Omega$. The three faults are correctly identified by the two relays that protect L_{14} and are never identified by the 6 other relays of the grid. No dependability nor security failures were thus observed

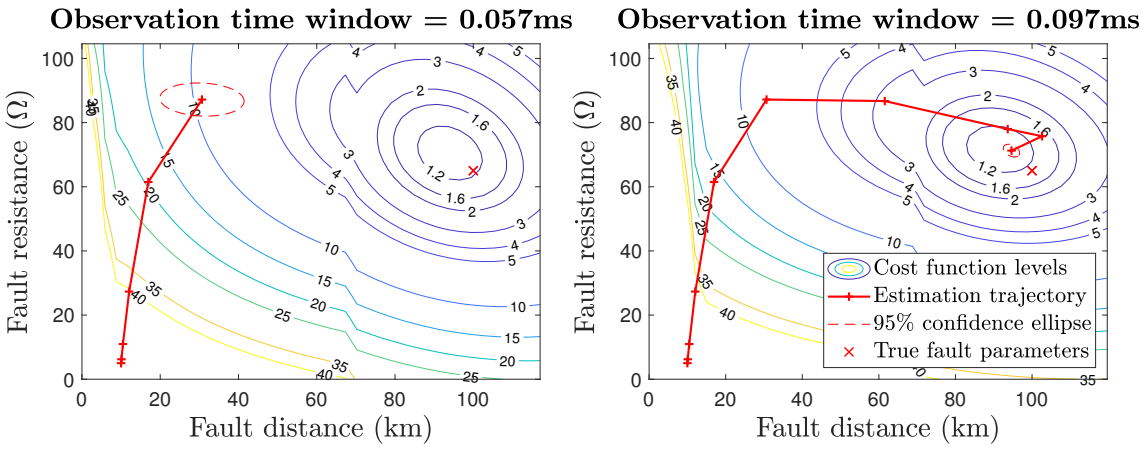


Figure 4.12.: Evolution of the estimated fault parameters and levels of the cost function at iterations 5 and 11. The pole-to-ground fault is located at $d_f = 100\,\text{km}$ from the station 1 with an impedance of $R_f = 65\,\Omega$. The estimated fault parameters after 11 iterations are $\widehat{R}_f = 71\,\Omega$ and $\widehat{d}_f = 95\,\text{km}$.

in those cases. For relays R_{14} and R_{41}, the performance of the fault identification using the adaptive filter is compared with the original approach where the soil resistivity is supposed uniform in Table 4.1. When using a uniform ground filter the soil resistivity is set according to the closest portion to the relay ($\rho_1 = 50\,\Omega$ for R_{14} and $\rho_2 = 5000\,\Omega$ for R_{41}). Only the situations for which the first wave travels through both portions of L_{14} are investigated, *i.e.*, F_2 seen from R_{41} and F_1, F_3 seen from R_{14}

The use of the adaptive filter generally improves the accuracy of the estimated fault parameters as well as the length of the observation window for faults F_1 and F_3. Nevertheless, the fault identification remains able to identify all faults assuming a uniform resistivity along the line. For fault F_2 seen from R_{41}, the two approaches lead to identical results. This is because the ground effect of the portion with $\rho_2 = 5000\,\Omega\text{m}$ is largely predominant in this case and corresponds both to the uniform and adaptive ground filters employed at R_{41}. By contrast, for F_1 and F_3 the effect of the closest portion to R_{14} for which $\rho_1 = 50\,\Omega\text{m}$ is negligible and the ground filtering must take into account the soil resistivity of the second portion.
The promising results obtained in this section for the adaptive ground filter proposed in Section 2.2.2.2 may be further evaluated on a broader range of faults cases. The applicability of the method on more complex configurations with more discontinuities also needs to be investigated.

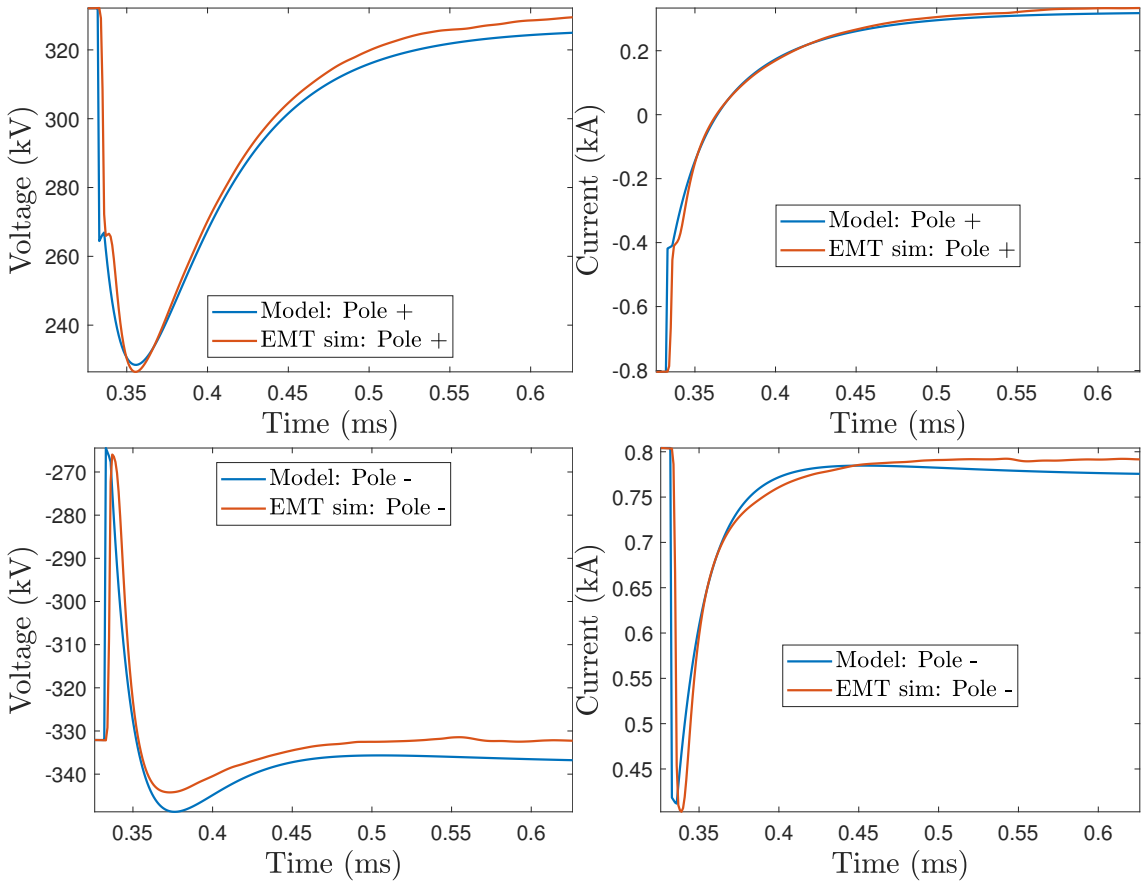


Figure 4.13.: Comparison of the modeled and simulated positive (top) and negative (bottom) pole voltage (left) and current (right) at the relay R_{14}. The different quantities are represented on a measurement window of 300 μs longer than the 97 μs. required to identify the fault.

Table 4.1.: Comparison of the adaptive ground filter approach with the uniform soil approximation when faults occurring on non-uniform transmission lines. The three fault cases F_1, F_2, and F_3 are successively compared.

(a) Fault F_1 and identification at relay R_{14}

Indicator	ε_R $(\%Z_{base})$	ε_d $(\%d_{14})$	Window length (μs)
Adaptive ground filter	6.1	1.8	97
Uniform ground ($\rho = 50\,\Omega$m)	11.1	1.1	107

(b) Fault F_2 and identification at relay R_{41}

Indicator	ε_R $(\%Z_{base})$	ε_d $(\%d_{14})$	Window length (μs)
Adaptive ground filter	14.9	7.9	117
Uniform ground ($\rho = 5000\,\Omega$m)	14.9	7.9	117

(c) Fault F_3 and identification at relay R_{14}

Indicator	ε_R $(\%Z_{base})$	ε_d $(\%d_{14})$	Window length (μs)
Adaptive ground filter	4.7	3.9	107
Uniform ground ($\rho = 50\,\Omega$m)	0	13.7	177

4.2.2. Extensive simulations and comparison with existing method

To evaluate the proposed approach on a wider range of fault cases, we performed extensive simulations on the 4-station grid presented in Figure 4.1. Multiple pole-to-ground faults were simulated on the four lines of the grid with varying fault distance d_f and fault resistance R_f, such that, for each line $e \in (L_{12}, L_{13}, L_{14}, L_{24})$, $R_f \in (0, 10, 40, 80, 120, 160)\,\Omega$ and $d_f \in (0.01, 0.1, 0.2, 0.4, 0.6, 0.8, 0.9, 0.99) \times d_e$. A total of 192 fault cases were thus simulated. For each of them an abnormal behavior is detected at all the 8 relays and the fault identification is started. This represents 1536 executions of the identification algorithm. The ability of the proposed approach to reliable identify faults using a short measurement window is compared with an existing method in Section 4.2.2.1. The accuracy of the estimated fault parameters is detailed in Section 4.2.2.2. Finally, computing time of the algorithm is analyzed in Section 4.2.2.3.

4.2.2.1. Fault identification - comparison with existing method

We compare the proposed approach with a reference fault identification algorithm presented in [Zhang et al., 2020]. We adapted the reference method to our test grid which does not include DC reactors at the end of each line, contrary to the grid used in [Zhang et al., 2020]. In this method, the authors propose a behavioral model of

the ground mode first current TW

$$i_{\text{ground}}(t) = -a \exp(-bt) + c$$

where the parameters a, b, c are estimated when a fault is suspected at the relay. The fault is identified on the protected line if the estimated parameter $\widehat{b}$ goes above some threshold, *i.e.*,

$$\widehat{b} > b_{\text{th}}. \tag{4.3}$$

The tuning of the reference method are adjusted to the 4-station grid used for the extensive simulations. The maximum value of the estimated parameter b for internal fault is computed as suggested by the authors of [Zhang et al., 2020] using the approximated speed of the ground mode c_{g} and the total length d_e of the protected line e„

$$b_{\text{se}} = \frac{c_{\text{g}}}{d_e}.$$

The threshold on the parameter b is then defined as $b_{\text{th}} = k_{\text{rel}} b_{\text{se}}$, where k_{rel} is a reliability factor set to $k_{\text{rel}} = 3$.

To analyze and compare the results of the two approaches, we define the dependability rate for line e as the proportion of faults that are correctly identified on the faulty line

$$r_{\text{d}}(e) = \frac{\#\text{Faults correctly identified on line } e}{\#\text{Faults that occured on line } e}.$$

Similarity the security rate for line e is the rate of faults that where correctly identified to be outside line e

$$r_{\text{s}}(e) = 1 - \frac{\#\text{Faults wrongly identified on line } e}{\#\text{Faults that occured outside line } e}.$$

These indicators[3] are given for the proposed approach and for [Zhang et al., 2020] in Table 4.2 for faults affecting the different lines. The proposed approach shows a slightly better dependability, except for faults occurring on the shortest line L_{24} (100 km long) for which [Zhang et al., 2020] performs particularly well. Furthermore, the proposed method proves to be much more secure against external faults, which may be related to the difficulty to extend the reference method to a grid without large DC reactors placed at the end of the lines.

The evolution of the dependability rate with the fault distance and resistance is detailed in Figure 4.14, showing the fault resistance only has a limited impact on the dependability. The influence of the fault distance resembles the feature of a

Table 4.2.: Fault identification results of the proposed approach for a bipolar configuration. Various faults are simulated on the four lines of the test grid.

Indicator	Approach	L_{12}	L_{13}	L_{14}	L_{24}
Dependability	Proposed	**92%**	**91%**	**92%**	91%
	From [Zhang et al., 2020]	90%	88%	82%	**95%**
Security	Proposed	**99%**	**99%**	**99%**	**99%**
	From [Zhang et al., 2020]	85%	99%	85%	93%

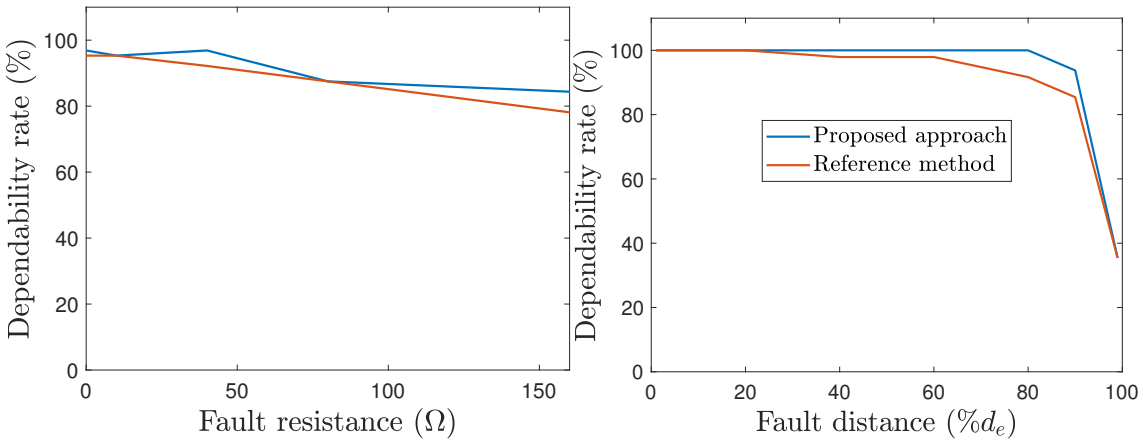


Figure 4.14.: Average dependability rate with the fault resistance and distance.

distance protection: dependability failures only occur for faults located near the remote end of the protected line, in this case $d_{\mathrm{f}}^{*} \geq 0.9 d_e$.

The length of the observation window used for the identification of internal faults is displayed for all the fault cases in Figure 4.15. The measurement window is less than 200 μs long in 93% of the fault cases for the proposed approach and only in 46% of the fault cases for the reference method. Specifically for the proposed approach, all faults occurring within the monitored portion of the line e, *i.e.*, $d_{\mathrm{f}}^{*} \leq 0.9 d_e$ are identified using a window of less than 0.25 ms. The proposed approach, indeed, stops as soon as the fault is identified whereas in [Zhang et al., 2020] the algorithm waits for the arrival time of the second TW or, if possible, for the availability of 1 ms long data window. As implemented in Matlab, the proposed approach has an average computing time of 138 ms to identify internal faults whereas [Zhang et al., 2020] requires 9 ms on average in. The computing time of the proposed approach is further discussed in Section (4.2.2.3).

4.2.2.2. Accuracy of the estimated fault parameters

In addition to the reliable identification of internal faults, the approach proposed produces an estimate of the fault parameters. For internal faults which are correctly

[3]The proposed metrics are also known as sensitivity and specificity [Fawcett, 2006].

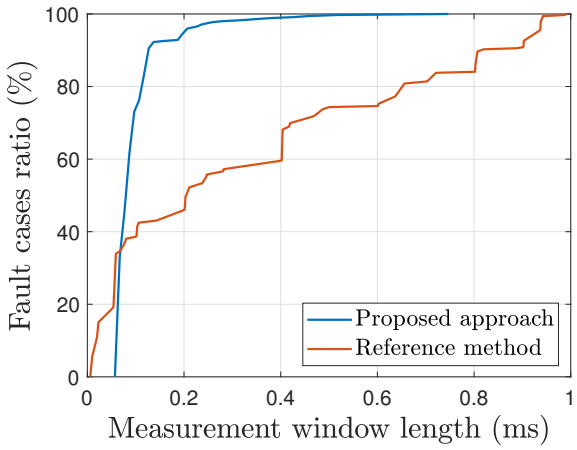


Figure 4.15.: Length of the observation window required for the identification of internal faults for the proposed approach and for the reference approach from [Zhang et al., 2020].

identified, the accuracy of the estimated fault parameters is compared to the true value of the fault parameters, see Figure 4.15. The error rate for the fault distance (4.1) is given as a percentage of the total line length and the error rate for the fault resistance (4.2) is given as a percentage of the base DC impedance. For 66% of the faults identified as internal, the relative error is less than 13% for both the fault distance and resistance. This is acceptable as the primary purpose of the algorithm is to reliably identify as fast as possible the faults that occur on the protected line. The algorithm generally stops before it has converged to the minimum of the cost function, see Figure 4.2. This allows to identify the fault while using very few measurements but also limits the precision of the estimated parameters.

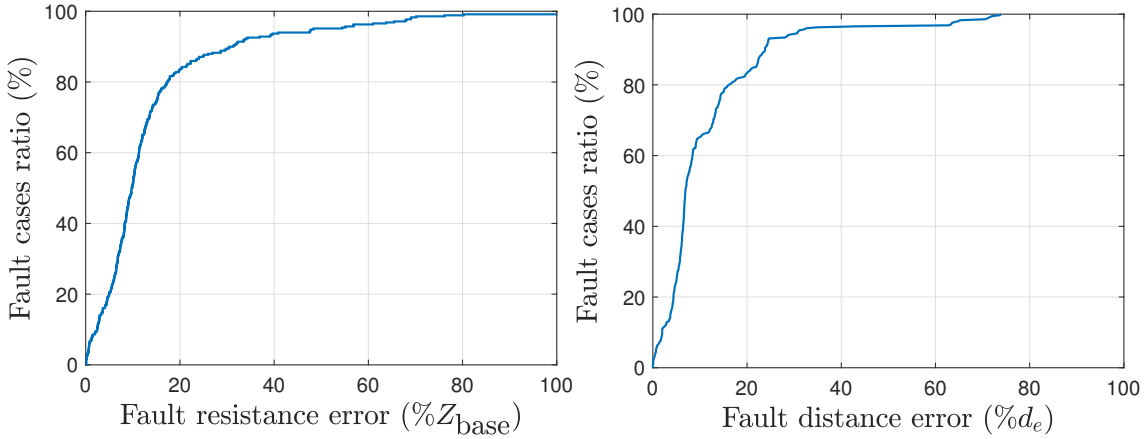


Figure 4.16.: Cumulative distribution function of the relative errors for the estimated fault resistance (left) and distance (right) using the proposed approach. For 66% of the faults identified as internal, the relative error is less than 13% for both the fault distance and resistance.

The estimated fault parameters can be used to assess the severity of the fault. For the least severe faults, *e.g.,* remote or high-impedance faults, an ultra-fast decision to trip the breakers may not be required. In such cases, the available time allows to wait for additional information from remote relays, making the fault identification results more reliable. This possibility is detailed in the primary sequence of the proposed full selective fault clearing strategy, see Section 5.3.

4.2.2.3. Computing time performances

The performances regarding the computing times of the fault identification algorithm as implemented in Matlab is further detailed in this section for the bipolar configuration. As the time constraints are mostly stringent for internal faults, we only consider here the computing times required to identify internal faults. Considering the extensive simulations performed in Section 4.2.2, the cumulative distribution of the computing times are plotted for the identified fault cases in Figure 4.17. For 70% of the faults identified as internal, the computing time is less than 50 ms. A real time objective at 0.5 ms is indicted as it corresponds to the maximum length of the observation window required for the identification of most faults, see Figure 4.15. The obtained computing times is thus about 100 times more than the real-time objective of 0.5 ms for the fault identification.

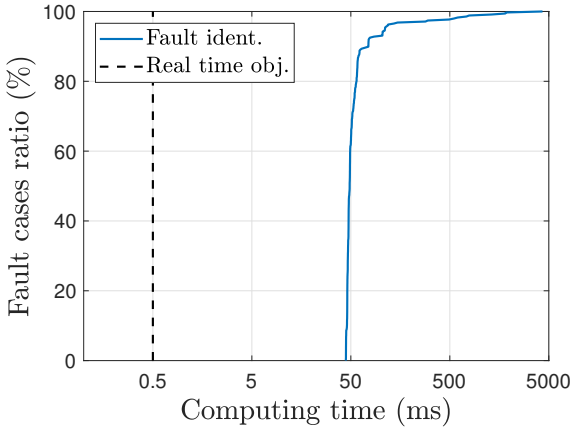


Figure 4.17.: Cumulative distribution function of the computing times (log-scale) for fault identification algorithm obtained with the extensive simulations.

Most of the computing time of the algorithm lies in the evaluation of the cost function (3.17) and its partial derivatives with respect to the fault parameters. This amounts to the evaluation of the different voltage waves and their derivatives. These computations are investigated apart from the rest of the algorithm using measurement widows of lengths 0.5 ms and 1 ms long and random fault parameters. The obtained execution times for one evaluation of the cost function and its partial derivatives are plotted in Figure 4.18. The first bars to the left correspond to faults for which the evaluation of only one wave is required, which are faults occurring

relatively far from the extremities of the line. The bars on the right correspond to faults that require the evaluation of two or three waves. Longer observation windows are more likely to require the computation of two or three waves, which explains that more fault cases are located to the right for the longest observation window. The length of the window in itself explains the shift of the different bars to the right but has a less significant impact than the number of waves that must be computed. The performance of the developed model to compute the evolution

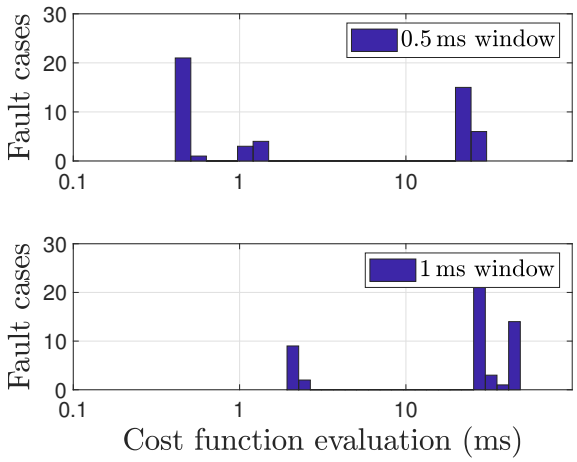


Figure 4.18.: Time required to evaluate once the cost function and its derivatives considering an observation window 0.5 ms (top) and 1 ms (bottom) long.

of the voltage in the first millisecond after the fault can be compared with existing tools, in particular with EMT software. A plain EMT computation requires the simulation of the grid on hundreds of milliseconds before the fault. To compare only the evaluation of the voltage in the first millisecond after the fault, a reduced EMT model is implemented. As a comparison point, the results from [Guo et al., 2015] reduced to a observation window of 1 ms, are also indicated. In [Guo et al., 2015], a sensitivity-based approach allows one to compute the traveling waves for any fault resistance, assuming the remaining fault conditions (in particular the fault distance) are known. The results show the proposed parametric model performs well compared to existing methods, at the cost of a reduced accuracy.

Table 4.3.: Comparison of the computation times to evaluate the voltage and current evolution on an observation window of 1 ms.

Model	Computing times (ms)
Approach from [Guo et al., 2015]	6-12
Simplified EMT	150
Proposed model	2-20

The obtained computing times, despite presenting performances in line with existing tools, are still about 500 times too slow compared to the real time objectives. The

improvement of the computing times is a challenge, nevertheless some possibilities are mentioned in Section 6.3.

4.2.3. Erroneous soil resistivity

In this section we investigate the consequence of using erroneous value of the soil resistivity in the ground filter H_g different from the value used in the EMT simulations. Though the resistivity is a property that depends on the nature of the ground, it is prone to variations due to changes in environmental conditions such as the temperature or the humidity, see Figure 4.19. More specifically, the deeper layers (at a deeper layer than 1 m) are less prone to changes than the surface layers [Coelho et al., 2015, Afa and Anaele, 2010].
As soil resistivity impacts essentially the ground mode which does not propagates in the shallow layers, only the soil resistivity variations at deeper layers is of interest. Hence we consider that for a known resistivity ρ_0, the actual value may variate within $[0.5\rho_0, 2\rho_0]$.

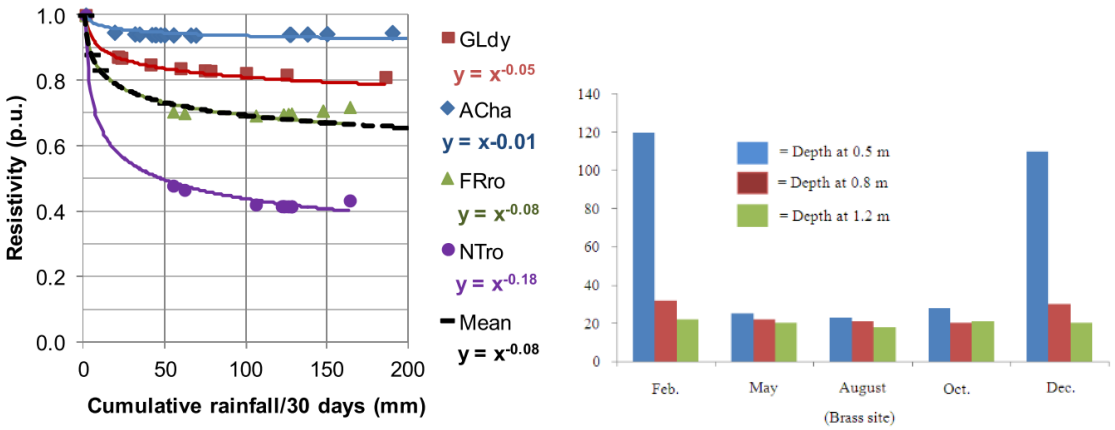


Figure 4.19.: Left: Variability of the soil resistivity (in p.u.) depending on the cumulative rainfall for rods between 1 m and 2 m deep, from [Coelho et al., 2015]. Right: Seasonal evolution of the soil resistivity (in Ωm) for rods 0.5 (blue), 0.8 (red), and 1.2 (green) meters deep, from [Afa and Anaele, 2010].

We performed extensive simulations as in Section 4.2.2 using the simulation data from EMT software with an actual resistivity of $\rho_0 = 100\,\Omega$m. The extensive simulations are performed with three different hypothesizes for the soil resistivity: $\rho = 50\,\Omega$m (underestimated value), $\rho = 100\,\Omega$m (actual value) and $\rho = 200\,\Omega$m (overestimated value). The performance indicators introduced in Section 4.2.2 are compared considering these three cases. The ability of the algorithm to identify only internal faults is evaluated in Table 4.4. An underestimation of the soil resistivity ($\rho < \rho_0$) causes more dependability failures while the security is almost not affected.

Indicator	ρ modeled (Ωm)	L_{12}	L_{13}	L_{14}	L_{24}
Dependability	$\rho = 50 < \rho_0$	84%	73%	88%	67%
	$\rho = 100 = \rho_0$	**92**%	91%	**92**%	91%
	$\rho = 200 > \rho_0$	**92**%	**93**%	**92**%	**92**%
Security	$\rho = 50 < \rho_0$	**100**%	**99**%	**100**%	**99**%
	$\rho = 100 = \rho_0$	99%	**99**%	99%	**99**%
	$\rho = 200 > \rho_0$	98%	**99**%	99%	98%

Table 4.4.: Influence of an erroneous modeled soil resistivity on the dependability and security of the fault identification algorithm.

The impact on the required observation window length is depicted in Figure 4.20 where the cumulative distributions of the observation window lengths are compared for the three values of ρ. More than 90% of the fault cases are identified with an observation window less than 0.4 ms long. Nevertheless, an underestimated value of ρ leads to a slight increase of the required observation window length.

The precision of the estimated fault parameters is only marginally affected by an inaccurate ρ, as shown in Figure 4.21. Considering $\rho = 50 < \rho_0$ leads to similar - or even improved - precision whereas $\rho = 200 > \rho_0$ decreases the accuracy of the estimated fault distance.

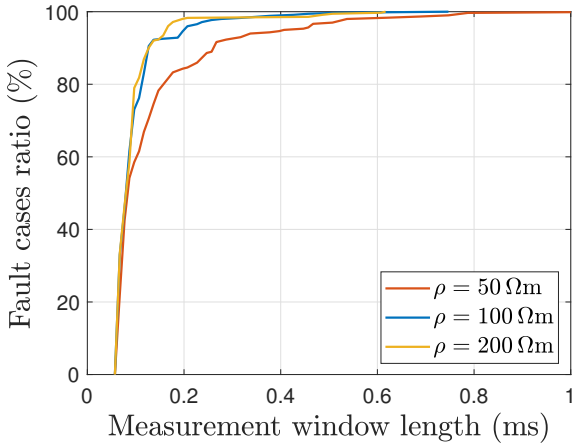


Figure 4.20.: Influence of an erroneous soil resistivity on the required length of the observation window to identify internal faults. The cases where $\rho < \rho_0$ (red) and $\rho > \rho_0$ (yellow) are compared with the nominal case $\rho = \rho_0$ (blue).

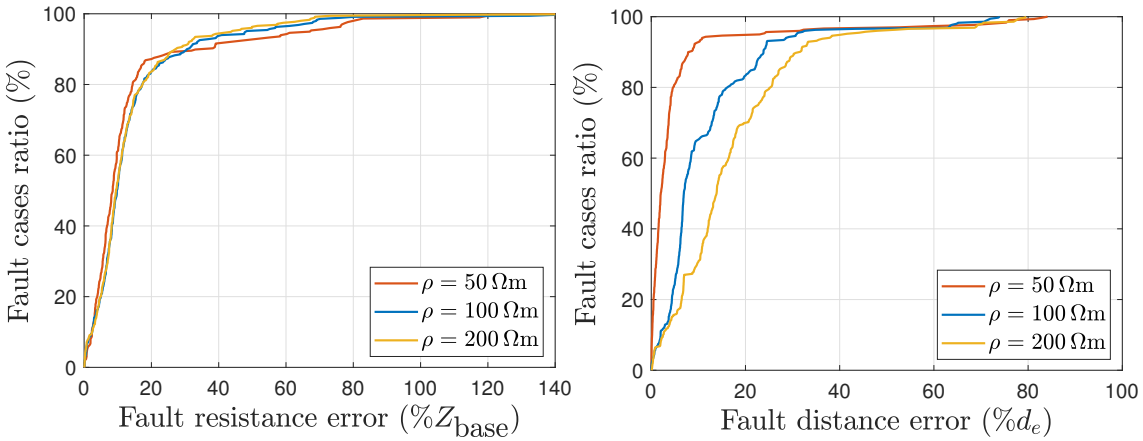


Figure 4.21.: Cumulative distribution function of the fault resistance error (left) and distance (right) error for the three values of ρ. The precision of the estimated fault resistance is barely impacted by the considered soil resistivity. An overestimated ρ leads to significantly less accurate fault distance. An underestimated ρ improves the precision of the estimated distance.

A better insight on the impact of the modeled soil resistivity ρ on the algorithm can be obtained by observing that the distortion of the TW increases both with ρ and the estimated fault distance $\widehat{d}_{\text{f}}$. Consider a fault occurring on a line with an actual soil resistivity ρ_0 that triggers an identification algorithm that uses a model tuned with a soil resistivity $\rho < \rho_0$. The algorithm is likely to overestimate the fault distance d_{f} in order to compensate for the observed distortion. Overestimated fault distances are more likely to exceed the maximum fault distance $d_{\text{f,max}}$, eventually causing dependability failures, see Table 4.4. Nevertheless, as the estimated fault distance is usually lower than the true value when $\rho = \rho_0$, $\widehat{d}_{\text{f}} < d_{\text{f}}^*$, a higher $\widehat{d}_{\text{f}}$ is likely to reduce the fault distance error, as seen in Figure 4.21 for $\rho = 50\,\Omega$m. Likewise, an assumed $\rho > \rho_0$ induces a lower $\widehat{d}_{\text{f}}$, which reduces the possibility of dependability failures. Though security failures could be expected in this situation, they are not observed in Table 4.4 which shows the security of the algorithm is preserved. Consequently, the precision of d_{f} is decreased in Figure 4.21. Note that the estimated fault resistance is barely impacted as its estimation also benefits from the aerial mode which is not impacted by the soil resistivity.

4.3. Application to the mono-conductor case

The simulation results of the algorithm applied to the asymmetric monopole are presented in this section. As the OHL consist of a single conductor, the modal analysis is not required to establish the parametric model, see Section 2.3. Compared to the bipole, the model is reduced to the ground mode and the estimation may not benefit from the aerial mode. In addition, the ground wire was omitted for the monopole

which removes one potential source of model error. The parameter used for the EMT grid are detailed in Appendix A. An illustrative example is first presented in Section 4.3.1 before performing more extensive simulations in Section 4.3.2.

4.3.1. Illustrative example

Consider a pole-to-ground fault occurring at $t_f = 0$ between stations 1 and 4, *i.e.*, on line L_{14} of the grid in Figure 4.1. The fault is located at $d_f^* = 30\,\text{km}$ from station 4 with $R_f^* = 60\,\Omega$. Once an abnormal behavior is detected at relay R_{41}, the identification algorithm is started. Its behavior can be analyzed by plotting the contour of the cost function to minimize at each iteration as well as the trajectory of the estimate $\left(\widehat{d}_f, \widehat{R}_f\right)$ of the fault parameter vector, see Figure 4.22. The 95% confidence ellipse of the estimated parameter vector is also displayed at each step. It can be observed that the minimum of the cost function gets closer to (d_f^*, R_f^*) and that the cost contours concentrate around (d_f^*, R_f^*). The estimate $\left(\widehat{d}_f, \widehat{R}_f\right)$ also gets closer (d_f^*, R_f^*) and the size of the confidence ellipsoid reduces. The estimation algorithm stops and correctly identifies the fault on the line after 6 iterations, requiring only measurements obtained in a time window of $66\,\mu\text{s}$. The estimated parameters are $\widehat{d}_f = 26\,\text{km}$ and $\widehat{R}_f = 61\,\Omega$ when the algorithm stops. Nevertheless, the stopping conditions (see Section 3.2) may be satisfied before the minimum of the cost function is reached. This allows ultra-fast identification of the fault, which is the main objective of the protection algorithm, even though it limits the accuracy of the estimated parameters.

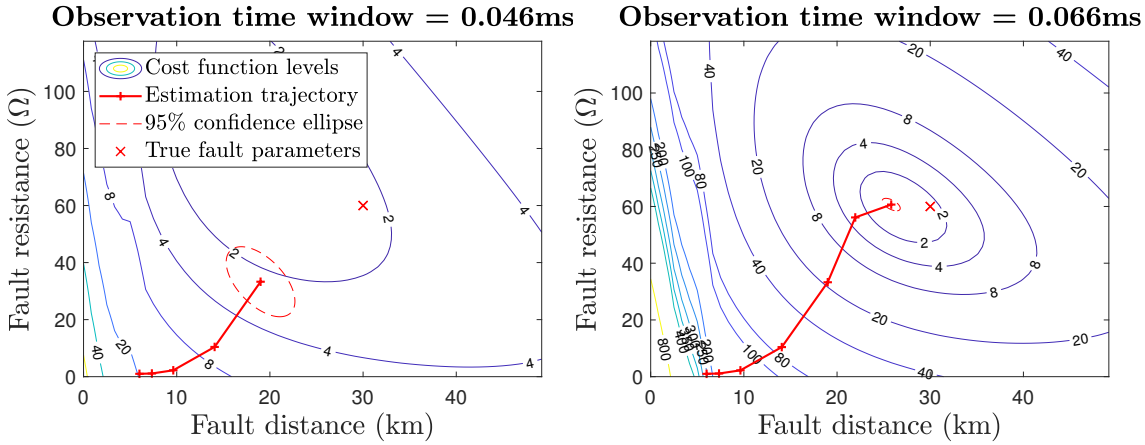


Figure 4.22.: Fault at $d_f^* = 30\,\text{km}$ from station 4 (close fault) with an impedance of $R_f^* = 60\,\Omega$: evolution of the contour plot of the cost function and estimated parameters at iterations 4 and 6.

The voltage and current measurements simulated with the EMT software are compared with the output of the parametric model using the estimated fault parameters $\widehat{d}_f = 26\,\text{km}$ and $\widehat{R}_f = 61\,\Omega$ in Figure 4.23. One sees that only the first traveling wave

has to be observed to get a sufficiently accurate estimate of the fault parameters for an accurate fault identification.

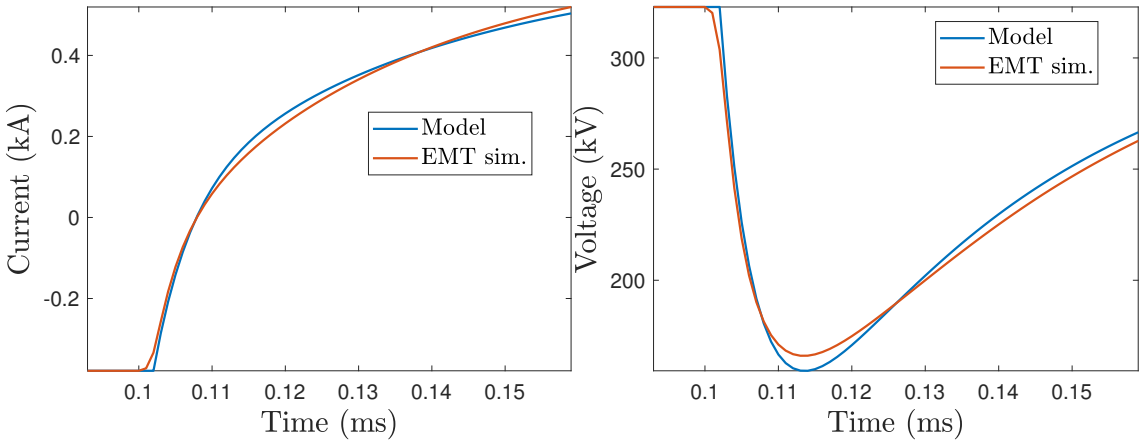


Figure 4.23.: Simulated current and voltage measurements compared to the parametric model outputs for $\widehat{d}_{\mathrm{f}} = 26\,\mathrm{km}$ and $\widehat{R}_{\mathrm{f}} = 21\,\Omega$.

The fault affecting the line L_{14} also triggers the fault identification algorithm at the other relays of the grid which detect an abnormal behavior. To analyze the behavior of the identification algorithm at the different relays, the evolution of the accuracy criterion is compared for all the relays in Figure 4.24. The evolution of the area of the confidence ellipsoid $\mathrm{vol}\left(\mathcal{R}^{\alpha}(\mathbf{p})\right)$ used in the accuracy criterion is plotted until the maximum measurement window is reached for the sake of comparison even if the algorithm stops as soon as the fault is identified. Since the fault is close to station 4, it is first detected at relays R_{41} (monitoring line L_{14}) and R_{42} (monitoring line L_{24}). The accuracy test is satisfied at relay R_{41} after 6 iterations and the fault is identified on line L_{14}, as shown previously. It can be seen that after reaching a minimum at 0.7 ms, the confidence region area starts increasing and reaches higher values after 1 ms. This is due to the reduced number of waves included in the model which limits its time validity. At relay R_{42} the area of the confidence region never goes below the threshold, indicating that the fault is not on line L_{24}. At $t = 0.6\,\mathrm{ms}$ the fault is detected at the relays R_{24} and R_{21} for which the fault is also identified as external. After 1 ms the first TW reaches the station 1, triggering the algorithms at relays R_{12}, R_{13} and R_{14}. The fault is again correctly identified on the line L_{14} after few iterations at R_{14}. The relays R_{12} and R_{13} identify the fault as external as their confidence region area remains above the threshold. The relay R_{31} is the last to detectthe abnormal behavior as it is located the farthest from the fault location. At this relay, the confidence region area reaches the threshold. However, the corresponding estimated fault parameters are not within the validity domain (3.24) (in particular $\widehat{R}_{\mathrm{f}} \geq 1000\,\Omega \gg R_{\max} = 200\,\Omega$) which prevents it from identifying the fault as internal.

This example illustrates that the method is able to identify internal faults using very few measurements while rejecting faults occurring on neighboring lines. Those

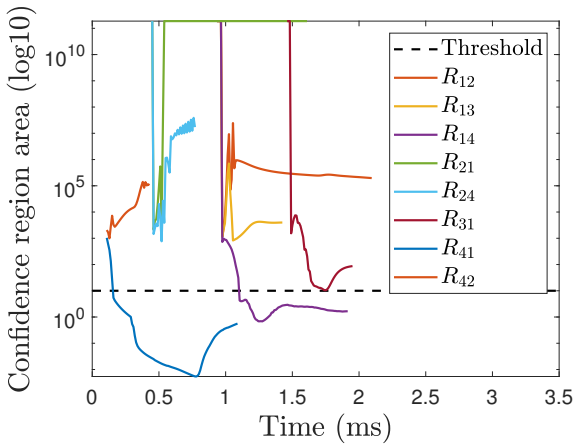


Figure 4.24.: Evolution of the accuracy criterion at the 8 relays of the grid. The 2 relays that protect the faulty line L_{14} identify the fault as internal using a short time window. The relay R_{31} correctly identifies the fault as external despite the small confidence region because the estimated fault parameters are not within the validity domain.

conclusions are confirmed in the next section on a wider range of fault cases.

4.3.2. Extensive simulations

More extensive simulations have been performed to analyze the behavior of the algorithm over a wider range of fault characteristics. We simulated different single fault scenarios with parameters $(R_\mathrm{f}, d_\mathrm{f})$ affecting the line L_{14} between stations 1 and 4, with $R_\mathrm{f} \in \{0, 5, 15, 25, 40, 60\}\,\Omega$ and $d_\mathrm{f} \in \{20, 50, 80, 100, 185, 240, 270, 280\}\,\mathrm{km}$, see Figure 4.1. The fault distance d_f is taken with respect to station 1. For all of the 48 values of the parameter vector, an abnormal behavior is always detected by the 8 relays of the grid which trigger the identification algorithms at different time instants. The results at relays R_{14} and R_{41} allow to evaluate the dependability of the proposed algorithm, i.e. its ability to identify all the faults occurring on the protected link. The outcomes at the 6 other relays are used to verify the security of the algorithm with respect to external faults.

Figure 4.25 displays the results of the identification algorithm implemented at the relays R_{14} and R_{41} for the considered fault distance and fault resistance. The green area corresponds to fault cases correctly identified by the proposed algorithms and the orange area indicates faults that were not identified on line L_{14}. The 3 non-identified faults lie outside of the protection zone (3.24) of the relays, *i.e.*, $d_\mathrm{f}^* \geq 0.9 d_{14} = 270\,\mathrm{km}$. This dependability analysis shows that the algorithm successfully identifies all faults within its protected zone, though some faults occurring in the protected line are not identified. Using the performance indicators introduced in Section 4.2.2 for the bipolar configuration, namely the dependability and security

rates, for the faults affecting L_{14} one has

$$r_{\mathrm{d}}\left(L_{14}\right) = 97\,\%$$
$$r_{\mathrm{s}}\left(L_{14}\right) = 100\,\%.$$

The security rate $r_{\mathrm{s}}\left(L_{14}\right) = 100\,\%$ indicates that for the considered fault cases, no security failures occur, *i.e.*, the algorithm running at relays R_{12}, R_{21}, R_{13}, R_{31}, R_{24}, and R_{42} never identify the faults affecting line L_{14} as an internal fault.

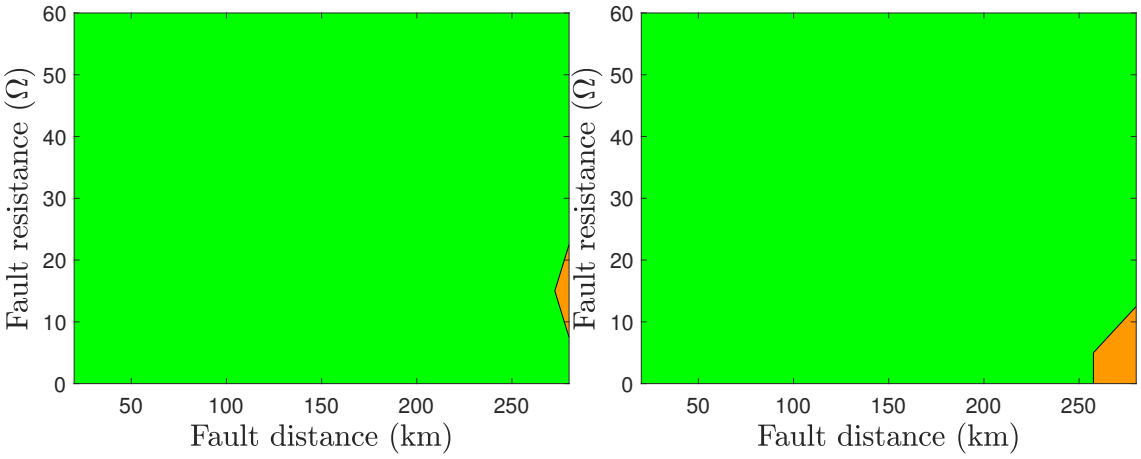


Figure 4.25.: Dependability analysis at relays R_{14} (left) and R_{41} (right): The green area corresponds to faults correctly identified on line L_{14}; the orange area indicates faults that were not identified on line L_{14}. The fault distance is taken with respect to the considered relay, *i.e.*, R_{14} (left) and R_{41} (right).

The cumulative distribution function of the duration of the measurement window required for the identification of all the fault cases at relays R_{41} and R_{14} is showed in Figure 4.26. The measurement window is always less than 0.7 ms and less than 200 μs for 90% the fault cases.

The relative precision of the estimated distance for the different fault cases is summarized and compared with several state-of-the-art methods in the Table 4.5. Though the precision of the proposed approach is lower, it is able to identify the faulty line using a measurement window 20 times smaller than the other methods. This makes it suitable as a primary fault identification algorithm for selective protection strategies. For the proposed approach the cumulative distribution of the error for the different fault cases is detailed in Figure 4.27. For 80% of the fault cases, the distance error is less than 3.3% of d_{14} (*i.e.* less than 10 km) and the resistance error is less than 23% of Z_{base} (*i.e.* less than 23 Ω).

The extensive simulations in the case of mono-conductor configuration confirmed the conclusions obtained in the bipolar case in Section 4.2. The slightly better performances compared to the bipole configuration, in particular regarding the accuracy of the estimated parameters, may be due to the simpler (mono-conductor) model

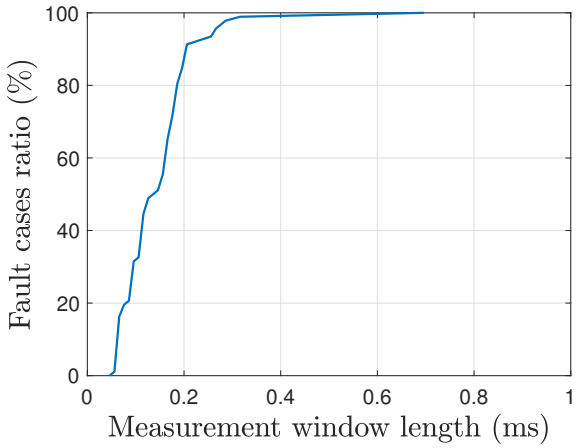


Figure 4.26.: Duration of the observation window required for faults identified at at relays R_{14} and R_{41}. The considered window length is always less than 0.7 ms and less than 200 μs for 90% the fault cases.

Table 4.5.: Comparison between the proposed scheme and other presented schemes.

Reference	Avg. observation	Distance error	Fault identification
[Suonan et al., 2010]	13 ms	0.02 %	No
[Farshad and Sadeh, 2013]	10 ms	0.3 %	No
[Abu-Elanien et al., 2016]	10 ms	0.4 %	Yes
[Ali Al Hage et al., 2016]	5.3 ms	2.3 %	Yes
Proposed algorithm	0.13 ms	2 %	Yes

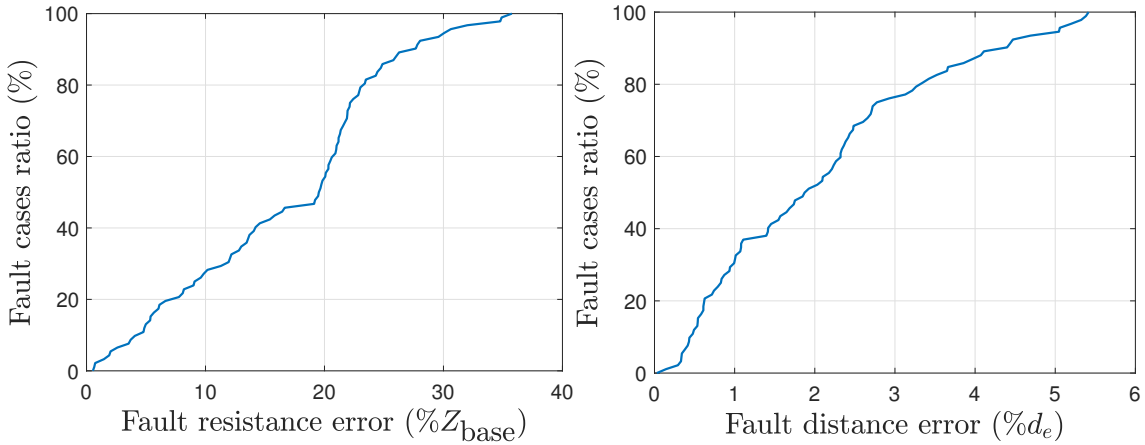


Figure 4.27.: Cumulative distribution function of the relative errors in the estimated fault resistance (left) and distance (right) for the faults identified at relays R_{41} and R_{14}.

employed as well as the absence of ground wire in the EMT grid. The good security of the algorithm with respect to external faults has been confirmed. While more than 90 % of internal faults are successfully identified, the dependability failures only occur for remote faults that lie outside of the protected zone of the relay.

4.4. Hybrid lines

This section presents an illustrative example of the proposed approach in the case of a grid comprising mixed lines with segments of overhead lines and underground cables. The diagram of the considered 4-station meshed grid with hybrid lines is reminded in Figure 4.28.

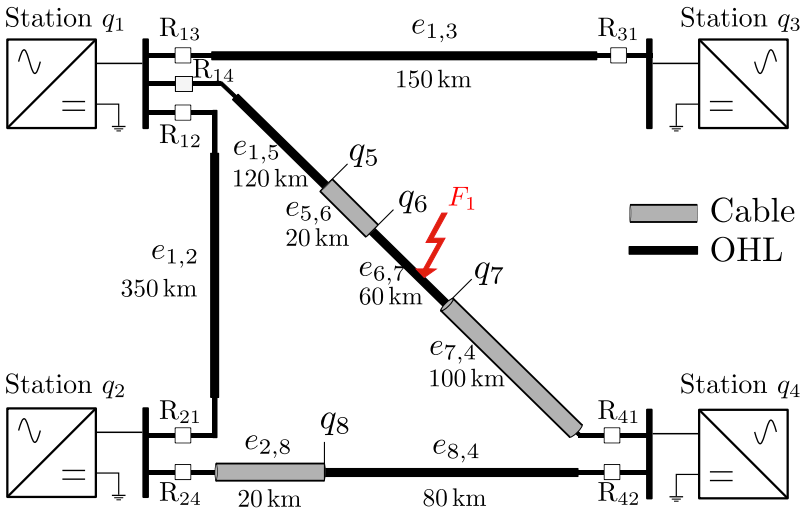


Figure 4.28.: 4-station meshed grid with hybrid lines.

In Figure 4.28, consider the fault F_1 affecting the line between station q_1 and q_4, on the edge $e_{6,7}$ corresponding to an overhead line section. The fault is located at $d_{\mathrm{f}}^{*} = 50\,\mathrm{km}$ from the node q_6 and has an impedance $R_{\mathrm{f}}^{*} = 70\,\Omega$. The behavior of the algorithm at the relay R_{14} is detailed. The algorithm starts four different parameter estimation algorithms in parallel, each corresponding to a hypothesis on the faulty segment, as described in Section 3.2. For each algorithm, the description of the network as a graph proposed in Section 2.4.1 is adopted. The area of the 95% confidence for the estimated parameters for the four different hypothesis are compared in Figure 4.29 (left). Three different hypothesizes satisfy the accuracy test as their confidence region area goes below the threshold. Nevertheless, the validity test is not satisfied for the hypothesizes corresponding to the edges $e_{5,6}$ and $e_{7,4}$, as their estimated fault resistance exceeds the maximum resistance $R_{\max} = 5\,\Omega$ for the cables. On the contrary, the estimated fault parameters assuming that the faulty edge is $e_{6,7}$ satisfy the validity test as the estimated resistance stays below the maximum fault resistance for an overhead section.

Thus, the fault is correctly identified after 16 iterations on the edge $e_{6,7}$ when the confidence region area for this hypothesis goes below the threshold. The estimated fault parameters after 16 iterations, with a measurement window of 160 μs are $\widehat{R}_{\mathrm{f}} = 47\,\Omega$ and $\widehat{d}_{\mathrm{f}} = 54\,\mathrm{km}$.

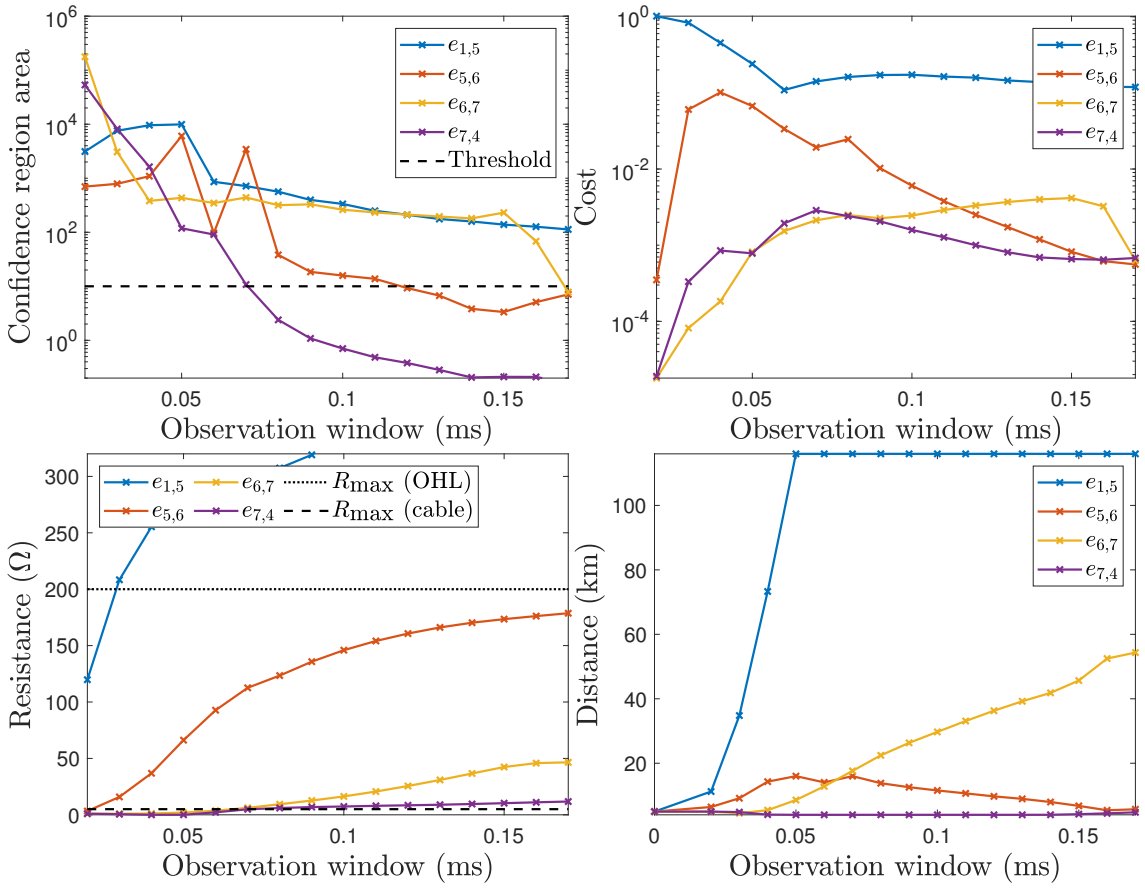


Figure 4.29.: Evolution of the 95% confidence ellipse area (top left), value of the cost function (top right) and estimated fault resistance (bottom left) and distance (bottom right) for each hypothesis. The hypothesizes $e_{6,7}$ is selected after 14 iterations as it is the only one to satisfy both the accuracy and validity tests. By contrast, the estimated fault resistances for the hypothesizes $e_{5,6}$ and $e_{7,4}$ exceed the maximum resistance $R_{\max} = 5\,\Omega$ for the cables.

The evolution of the estimated fault distance and resistance at each iteration is represented in Figure 4.30 and compared with the true fault parameters.

The waveform of the voltage and current for the EMT simulations and parametric model with the estimated fault parameters are compared in Figure 4.31. The model fits relatively well the EMT data despite the error mostly related to the difference between the estimated and actual fault resistance.

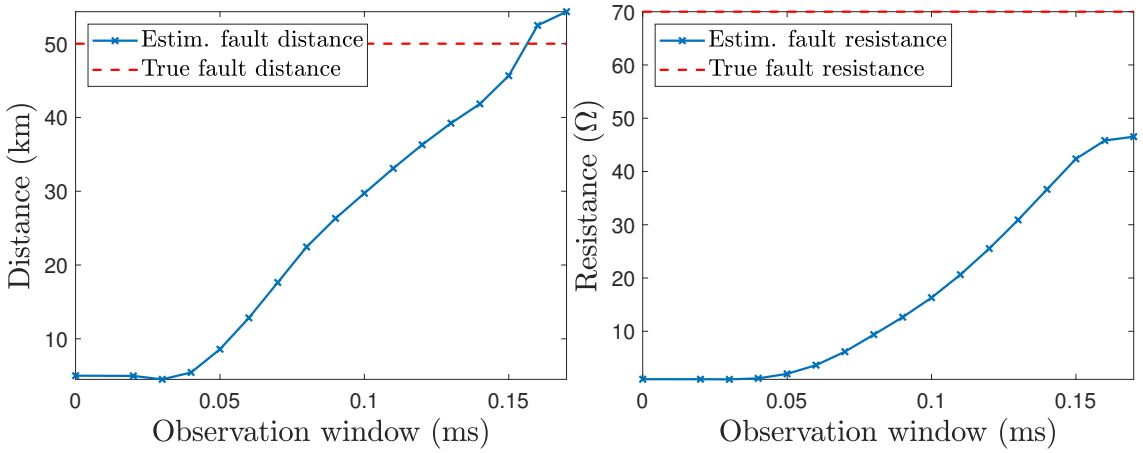


Figure 4.30.: Evolution of the estimated fault distance (left) and resistance (right) after each step of the estimation algorithm.

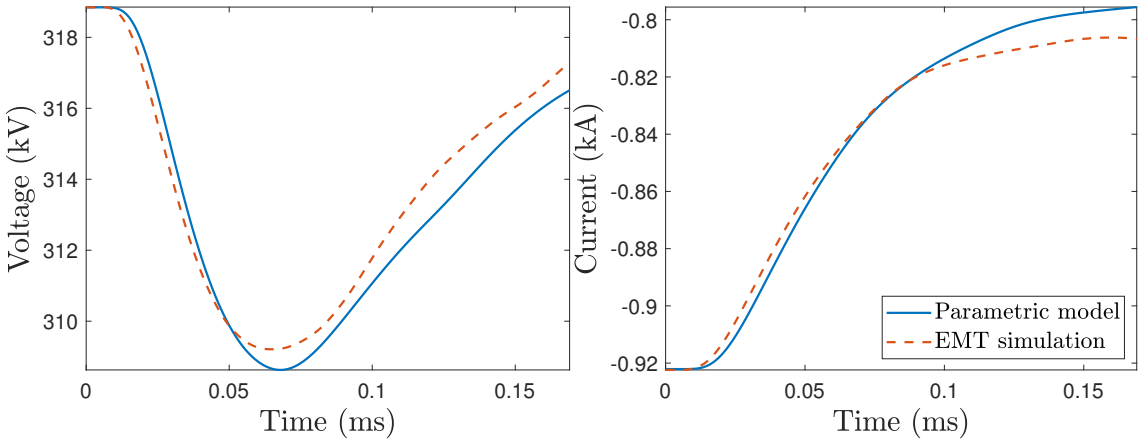


Figure 4.31.: Comparison of the modeled and simulated modal voltage (left) and current (right) at the relay R_{14}. The fault parameters used in the parametric model are those obtained after 16 iterations: $\widehat{R}_\mathrm{f} = 47\,\Omega$ and $\widehat{d}_\mathrm{f} = 54\,\mathrm{km}$.

4.5. Conclusion

In this chapter, the simulations results of the proposed fault identification approach are detailed in the case of bipolar and monopolar overhead lines as well as for hybrid lines. For the case of a grid with OHL, the performance of the algorithm has been evaluated with extensive simulations on a wide range of fault cases. While the algorithm is able to identify the majority of internal faults, we also evidenced some dependability failures for remote faults. A sensitivity analysis regarding the knowledge of the soil resistivity was also proposed, showing an the method is robust to a bad knowledge of the value of the soil resistivity . The next chapter investigates the integration of the fault identification algorithm into a full selective fault clearing strategy.

5. System integration

This chapter addresses the system integration of the proposed fault identification algorithm within a fault clearing strategy. Some of the system requirements especially imposed by the circuit breakers capabilities are detailed in Section 5.1. A collaborative fault identification algorithm is then proposed in Section 5.2 to increase the reliability of the identification. A full selective fault clearing strategy is specified in Section 5.3, including the primary sequence and some backup sequences. The impact of considering more realistic sensors on the fault identification algorithm is detailed in Section 5.4.

5.1. System requirements

We focus in this section on two critical aspects of the fault clearing strategy. The ability of the DC circuit breakers (DCCB) to interrupt the fault current before the current exceeds the breaking capability is first analyzed in Section 5.1.1. The impact of the fault current and the fault neutralization time on the blocking of the converter stations is then investigated in Section 5.1.2.

5.1.1. Opening of the circuit breakers

A fault affecting a line within the grid causes an important and fast rise of the current in the faulty line. This current must be interrupted relatively fast to avoid damages to the sub-stations equipment as well as large disturbances in the remaining DC or connected AC grids. More importantly, the DC circuit breakers must operate before the current exceeds their maximum breaking capability. The use of hybrid circuit breakers [Davidson et al., 2015] is seen as a good compromise between opening time, on-line losses and breaking capabilities. The considered DCCB requirements are specified in Table 5.1, see [PROMOTIoN WP6, 2016]. The neu-

Table 5.1.: Considered hybrid DCCB specifications

Breaking capability	16 kA
Breaker operation time	2 ms

tralization time is defined as the time at which the peak current in the faulty line is

reached [CIGRE JWG A3/B4.34, 2017]. With respect to the fault arrival time at the relay, it is the combination of the relay time required to identify the fault and the operating time of the breaker, *i.e.*,

$$\tau_{\text{nz}} = \tau_{\text{relay}} + \tau_{\text{CB}} \tag{5.1}$$

To evaluate the possibility to disconnect only the faulted line, one has to check that the line current is below the maximum CB capability at the fault neutralization time.

Extensive simulations were performed in the 4-station meshed grid considering a mono-conductor topology, see Figure 4.1. Various faults were simulated on lines L_{13} and L_{14} with fault distance and resistance such that
$d_{\text{f}} \in \{5, 15, 35, 55, 75, 95, 115, 135, 145\}$ km for faults affecting L_{13},
$d_{\text{f}} \in \{5, 15, 20, 25, 45, 50, 80, 85, 100, 140, 155, 185, 220, 240, 260, 270, 280, 290\}$ km for faults affecting L_{14} and $R_{\text{f}} \in \{0, 1, 10, 20\}$ Ω in both cases. The time at which the line current at the extremities of faulty line reaches the maximum breaking capability $i_{\max} = 16$ kA is displayed in Figures 5.1 and 5.2 for faults affecting L_{13} and L_{14}, respectively. For the considered faults, ($R_{\text{f}} = 0$)the maximum breaking current is always reached in more than 3 ms after the fault arrival. For the majority of faults, the maximum current is never reached (yellow area). The maximum current may be reached in 5 ms to 25 ms for low impedance faults occurring as far as 100 km from the line end.

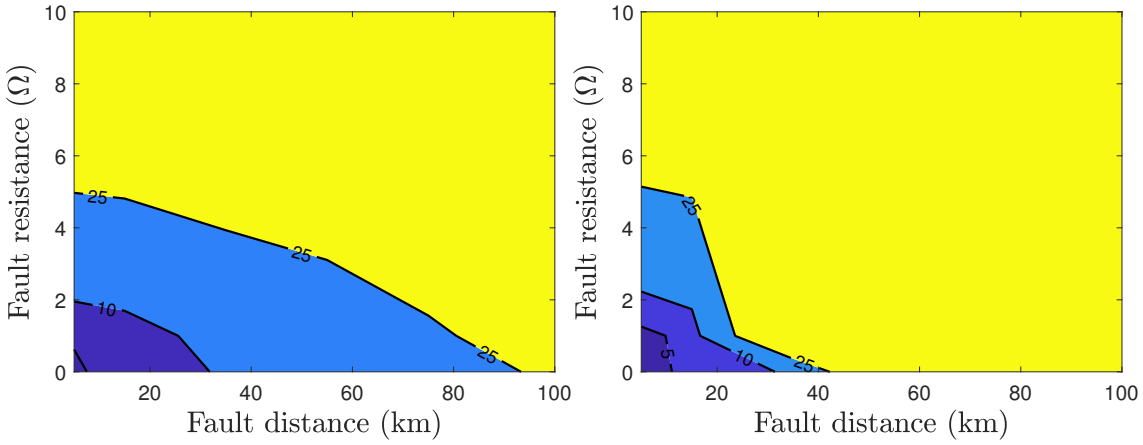


Figure 5.1.: For various faults affecting L_{13}, time at which the line current i_{13} (left) and i_{31} (right) at the two extremities of L_{13} reaches $i_{\max} = 16$ kA. The yellow area corresponds to faults for which the line current never reaches $i_{\max}$. For faults corresponding to the light blue area, $i_{\max} = 16$ kA is reached in less than 25 ms after the fault detection at the relay. The dark blue area denotes faults for which the maximum breaking current is reached in less 10 ms. For the closest fault with $d_{\text{f}} = 5$ km and $R_{\text{f}} = 0\,\Omega$, the maximum current is reached in 3 ms.

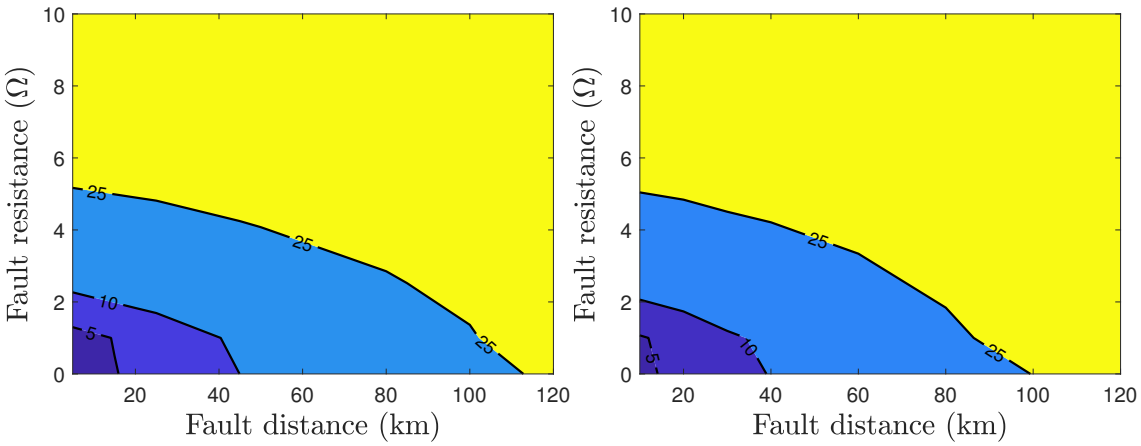


Figure 5.2.: For various faults affecting L_{14}, time at which the line current i_{14} (left) and i_{41} (right) at the two extremities of L_{14} reaches $i_{\max} = 16\,\text{kA}$. The yellow area corresponds to faults for which the line current never reaches $i_{\max}$. For faults corresponding to the light blue area, $i_{\max} = 16\,\text{kA}$ is reached in less than 25 ms after the fault detection at the relay. The dark blue area denotes faults for which the maximum breaking current is reached in less 10 ms. For the closest faults with $R_\text{f} = 0\,\Omega$, the maximum current is reached in 3 ms.

A solid fault occurring at the extremity of a line is further detailed. The inductive behavior of overhead line is negligible for such faults which may impose additional constraints on the DCCB. Consider a fault occurring at $d_\text{f} = 0.2\,\text{km}$ from station 4 (in the first tower span[1]) with $R_\text{f} = 0$. The evolution of the current and voltage at relay R_{41} in the first millisecond after the fault is displayed in Figure 5.3 (red curves). If not inductances are considered, as assumed before, the current reaches $i_{\max} = 16\,\text{kA}$ in 2 ms, which would prevent the opening of the DCCB with the specifications of Table 5.1. One may size the DCCB accordingly, for instance setting the maximum breaking current to 20 kA. An alternative option is to place DC reactors (DCR) at the output of the converter stations, which diminishes the rise of current. In the example of Figure 5.3, DCR of moderate size (50 mH) allows the opening of the DCCB as the maximum breaking current is reached in 4 ms. If DCR should be omitted at the output of the converter stations, the fault may still be cleared by tripping the adjacent DCCB connected to the same busbar. The adjacent lines are then disconnected along with the faulty line, leading to a semi-selective fault clearing strategy, as presented for instance in [Torwelle, 2021].

In case DCR are placed at the converter outputs, the fault identification algorithm can be easily adapted as it only requires to include the DCR value on the equivalent inductance of the MMC station. Though such inductances may be required to clear selectively very close faults or to prevent the blocking of the converter stations, as

[1]Such faults require a simulation time step in EMT software less than 1 μsdue to the small propagation time along a transmission line of length 200 m.

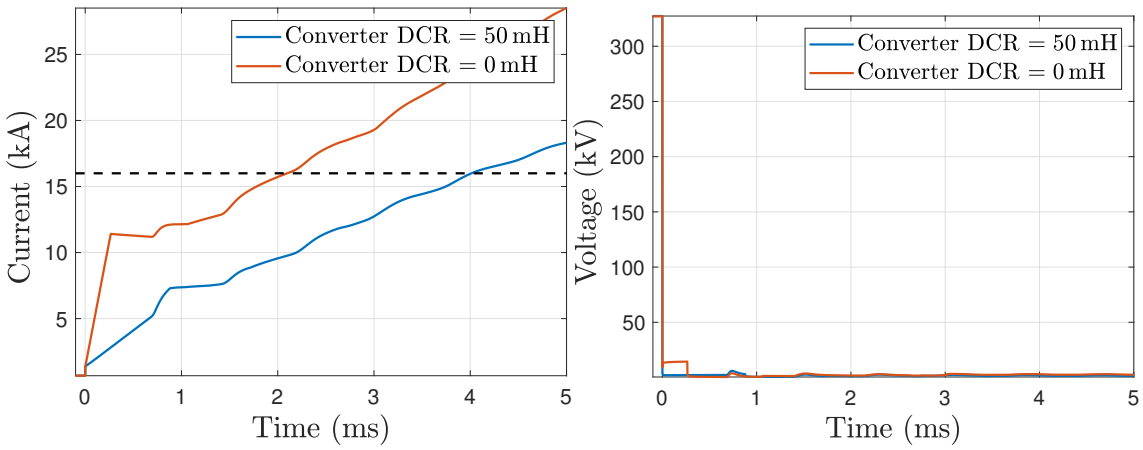


Figure 5.3.: Evolution of the current and voltage at relay R_{41} for a very close solid fault located on line L_{14} at 200 m from station 4. The impact of considering DC reactors at the output of the converter stations is investigated.

detailed in Section 5.1.2, they are omitted in the considered grid. The simulations performed in this section show that for the vast majority of faults, the maximum current breaking capability is reached in at least 3 ms after the fault arrival. Considering the DCCB have an operating time of 2 ms, this imposes that for such faults, the relay time required to identify the fault satisfies $\tau_{\text{relay}} \leq 1$ ms.

5.1.2. Blocking of the MMC stations

During a fault the current flowing through the MMC stations will rise and eventually triggers the power electronic components self-protection. The IGBTs will block if the current or voltage of the arm reaches certain thresholds. An overcurrent protection is considered so that the MMC is blocked if the current exceeds 2 p.u. [Auran, 2017]. This has a negative impact on the healthy lines connected to the blocked stations as it prevents the station to control the power flow on those lines. Once a fault is cleared, the station must then be de-blocked to return to normal operation. The fast opening of the line circuit breakers can prevent the station from blocking. This would diminish the fault impact on the grid and decrease the total duration of the protection strategy [Pang and Wei, 2018].

For the same fault cases as in Section 5.1.1, we determine the time at which the converter stations block due to overcurrent protection. The simulations are performed in open-loop, in particular the effect of the tripping of one or several CB is not simulated. Only the blocking of the converters at the extremities of the faulty line are analyzed as the they are more likely to block just after the fault occurrence. The blocking instants at stations 1 and 3 (with respect to the fault arrival time at those stations) are displayed in Figure 5.4 for the different fault affecting L_{13} and on Figure 5.5 for faults affecting L_{14} and the stations 1 and 4. The blocking of the

MMC stations can be prevented if the line breakers are tripped fast enough. This may not be possible for faults that cause the blocking in less than 3 ms (dark blue), as the DCCB opening itself lasts 2 ms. For many faults occurring all along the line and with a resistance up to 60 Ω ,the station may block after 3 to 25 ms and the blocking may be avoided if the DCCB are tripped before. Though the blocking of station 3 may not be preventable, it is not as damageable as for the other stations as station 3 is only connected to line L_{13} and is thus not available in case this line is faulty.

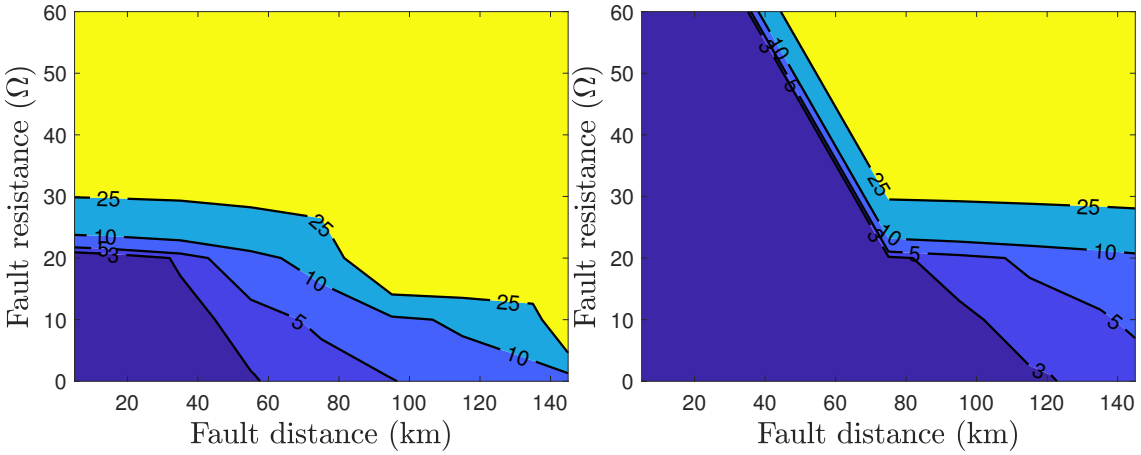


Figure 5.4.: For various faults affecting L_{13}, time at which the MMC at station 1 (left) and 3 (right) block for over-current auto-protection. The yellow area corresponds to faults for which the station may never block, or blocks after 25ms. The dark-purple area denotes the faults for which the MMC blocks in less than 3 ms. The faults for which the station blocks after 3, 5 and 10 ms are depicted in shades of blue.

As evidence in Section 5.1.1, the fast rise of the current in the faulty line for very close solid faults imposes that the line breakers are tripped in less than 3 ms. Furthermore, the fast disconnection of the faulty line for a wide range of fault cases may prevent one or both the converter stations connected to the faulty line to block, which can increase the reliability of the grid [Pang and Wei, 2018].

5.2. Collaborative fault identification

As evidenced in Section 5.1, some faults require that the line breakers are tripped very fast, *e.g.*, in 3 to 5 ms whereas more time may be available for the disconnection associated to other faults. In the latter case, the two relays located at the extremities of a transmission line may collaborate to increase the reliability of the fault identification, as presented in Section 5.2.1. The compliance of this approach with the system requirements is analyzed in Section 5.2.2 through the fault neutralization time.

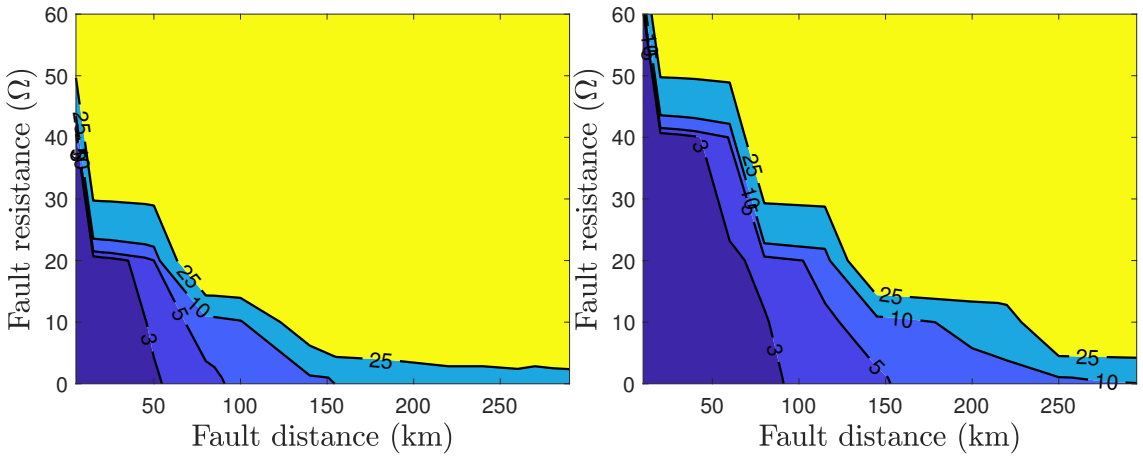


Figure 5.5.: For various faults affecting L_{14}, time at which the MMC at station 1 (left) and 4 (right) block for over-current auto-protection. The yellow area corresponds to faults for which the station may never block, or blocks after 25ms. The dark-purple area denotes the faults for which the MMC blocks in less than 3 ms. The faults for which the station blocks after 3, 5 and 10 ms are depicted in shades of blue.

5.2.1. Distinction between severe and non-severe faults

Based on the extensive simulations carried out in Chapter 4, both in the mono-conductor and multi-conductor configurations, the identification algorithm is able to identify most faults occurring on the monitored line while rejecting external faults, considering a wide range of fault cases. Two situations of possible maloperation of the algorithm have nevertheless be evidenced,

- Faults (F1) occurring on the protected line close to the remote station may fall outside the protected zone and be identified as external. This leads to a *dependability failure*.
- Faults (F2) occurring very close to the remote station but on an adjacent line may be wrongly identified as distant internal faults. This leads to a *security failure*.

Those two situations are depicted in Figure 5.6, where F_1 represents faults that can lead to a dependability and F_2 faults that can cause a security failure, both considering the fault identification process at the local relay. Note that in both cases the estimated fault parameters are rather large, in particular regarding the fault distance which would be close to the total line length.

It is significant that in both cases, the examined fault does not lead to an important rise of the current at the local relay, as the fault occurs near the remote end of the protected line. This leads to distinguish for a given relay, among internal faults, between severe and non-severe faults. In case of *severe faults,* the current flowing through the circuit breakers rise very fast and may exceed the breaker maximum

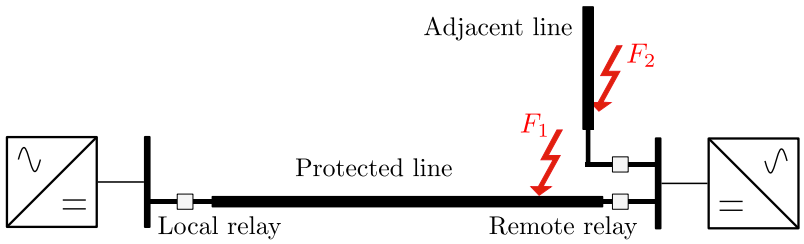


Figure 5.6.: Two evidenced fault cases that can lead to a maloperation of a relay, here denoted as the local relay: F_1 may not be identified and lead to a dependability failure while F_2 may be identified and cause and security failure.

capability. For such faults thus, the breakers have to be tripped fast enough so that they are able to open the line before the fault current reaches a certain value. By contrast, for non-severe faults, the current rise is moderate which gives more time to detect clear the fault. In particular, communication between the two relays that monitor the same line may be employed to improve the robustness of the fault identification.

As evidenced in Section 5.1, the severity of a fault is correlated to the fault parameters, the fault distance d_f and resistance R_f. Low impedance close faults lead to an important rise of the current and may cause the MMC stations to block whereas remote or high impedance faults only lead to a moderate fault current, see Figure 5.7 (left). The discrimination between severe/non-severe faults thus depends on the considered relay as well as the length d_e of the protected line .

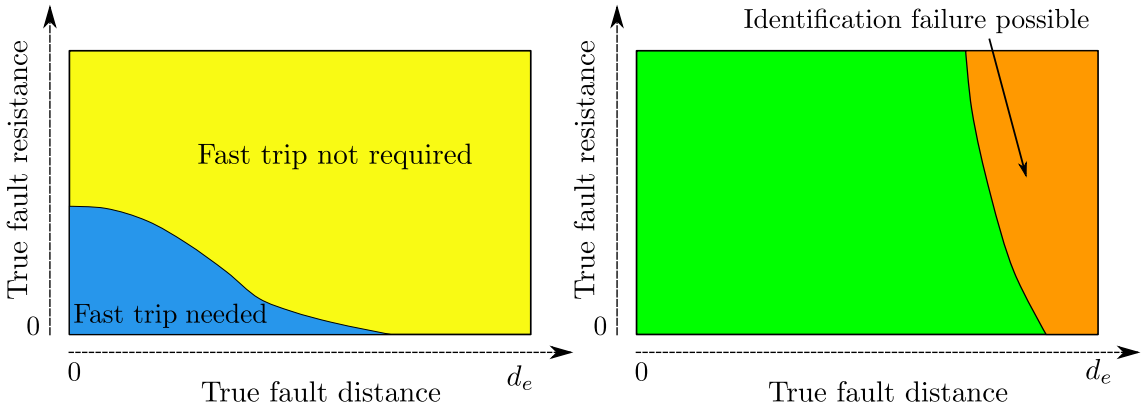


Figure 5.7.: Classification of possible faults based on the fault parameters, regarding the current rise (left) and fault identification algorithm possible failures (right).

This distinction can be compared to the fault parameters for which an identification failure may occur, see Figure 5.7 (right), where a general classification of faults is proposed. Hence the faults for which the algorithm may maloperate do not require a particularly fast action. The precise settings of the boundaries represented in Fig-

ure 5.7 depend on the considered line and hardware, in particular the line breakers, and is not discussed here.

The proposed collaborative algorithm aims to ensure a fast tripping of the breakers when required while diminishing the risk for identification failures by resorting to communication with the remote relay when possible. By contrast, the fault identification algorithm that only employs measurements from the local sensors is referred to as single-ended. The relay can assess the severity of the fault using the estimated fault parameters $\widehat{\mathbf{p}} = \left(\widehat{R}_\mathrm{f}, \widehat{d}_\mathrm{f}\right)$ as represented in Figure 5.8. Based on the type of

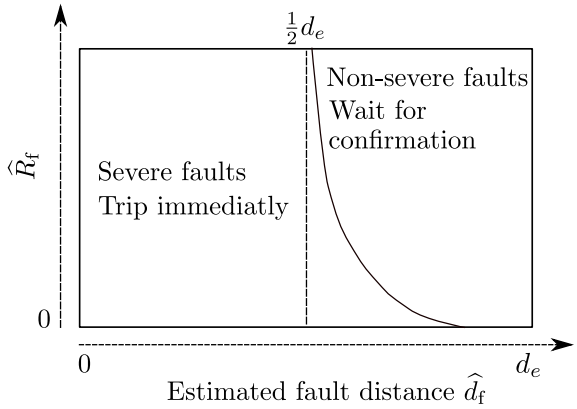


Figure 5.8.: Sending of the tripping orders depending on the estimated fault parameters. The tripping order is sent immediately for severe faults that might lead the current to exceed the DCCB capability while the relay waits for the confirmation in case of non-severe faults.

fault identified by the relay based on the single ended algorithm, different course of action may occur:

- Faults identified as severe by the local relay lead to the immediate trip of the local line breaker.
- Faults identified as non-severe by the local relay must be confirmed by the identification of the remote relay. We propose that the local breaker is tripped only if the remote relay identified a severe fault. This requires that all faults are identified as severe by at least one of the two relays, which can be achieved by ensuring that all faults identified in the first half of the line are considered as severe, see Figure 5.8. Alternatively, the confirmation of faults identified as non-severe by both relays may increase the dependability of the collaborative algorithm at the cost of a reduced security.
- Faults detected but non identified whitin the line protected by the local relay are deemed to be located elsewhere in the grid. Nevertheless, the breaker is tripped if the remote relay identified a severe fault. This *direct intertripping* scheme [Alstom, 2011] increases the dependability and may be justified as the line breaker located at the remote end is already tripped, already stopping

the operation of the line. The precise distinction between severe and non-severe faults must take into account the expected current rise, the expected performances of the identification algorithm as well as the chosen intertripping scheme and is left for further studies. If the identification algorithm at the remote relay does not employ the estimation of the fault parameters, the collaboration with other protection schemes such as directional algorithms is also conceivable.

5.2.2. Fault neutralization time

The neutralization time is defined as the time at which the peak current in the faulty line is reached [CIGRE JWG A3/B4.34, 2017]. With respect to the fault arrival time at the relay, it is the combination of the relay time τ_{relay} required to identify the fault and the operating time of the breaker τ_{CB}, *i.e.*,

$$\tau_{\text{nz}} = \tau_{\text{relay}} + \tau_{\text{CB}} \tag{5.2}$$

In this section it is assumed the identification time at a given relay τ_{ident} is only limited by the length of the required observation window. As evidenced by the extensive simulations performed in Sections 4.2.2 and 4.3.2 one has $\tau_{\text{ident}} < 1\,\text{ms}$ (and for the vast majority of faults the identification time is less than 0.5 ms). For severe faults, the relay time amounts to the time required for the identification performed by the local relay

$$\tau_{\text{relay}} = \tau_{\text{ident}} < 1\,\text{ms}\ , \tag{5.3}$$

Considering is in Table 5.1 the DCCB have a breaking capability of 16 kA in 2 ms, the neutralization time for severe faults is

$$\tau_{\text{nz}}^{\text{severe}} < 3\,\text{ms} \tag{5.4}$$

The required identification time for non-sever faults can be analyzed considering the Figure 5.9. In this situation, the local relay identifies a non-severe fault and waits for the confirmation from the remote relay, which identifies a severe-fault. The neutralization time at the local relay is then

$$\tau_{\text{nz}}^{\text{non-severe}}(d_{\text{f}}) = \underbrace{\tau_{\text{ident}}}_{\text{remote relay ident.}} + \tau_{\text{com}} + \tau_{\text{CB}} + \underbrace{t_{\text{arrival}}^{\text{remote}} - t_{\text{arrival}}^{\text{local}}}_{=\frac{d_e - 2d_{\text{f}}}{c_{\text{w}}}}. \tag{5.5}$$

The communication delay τ_{com} comprises the propagation time (through fiber optics) as well as an additional delay due to communication equipment, here considered to be 500 μs [Johannesson and Norrga, 2019], is

$$\begin{aligned}\tau_{\text{com}} &= \frac{d_e}{2e5} + 500 \times 10^{-6} \\ &\leq 2.3\,\text{ms}\end{aligned}$$

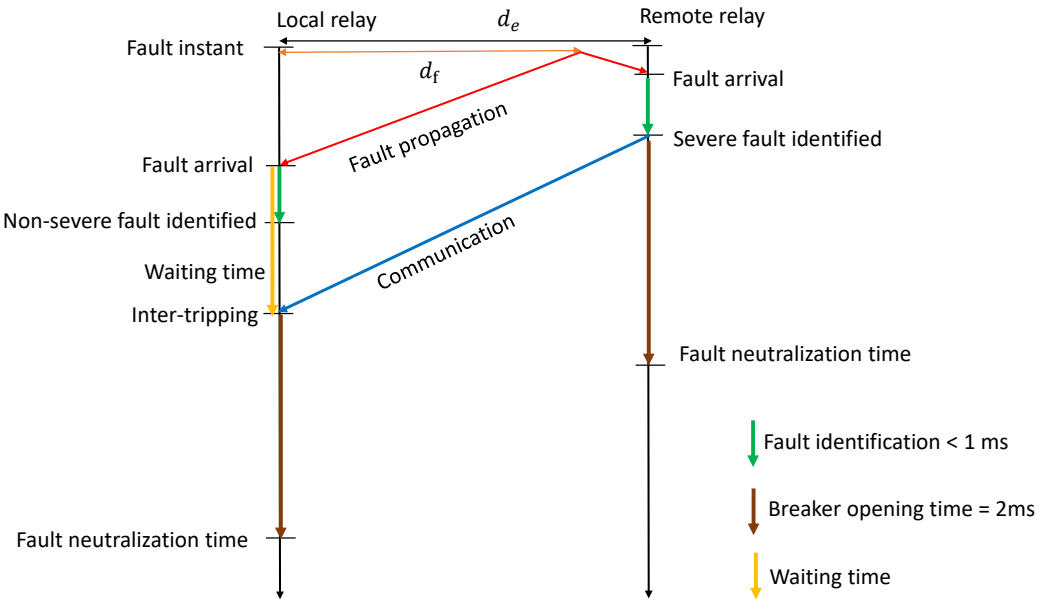


Figure 5.9.: Fault neutralization time for a non-severe fault.

considering the longest line $d_{12} = 350\,\text{km}$ in the grid. The neutralization time for a non-severe fault (5.5) thus depends on the position of the fault along the line. Assuming non-severe faults only lie in the second half on the line, a worst case approach leads to

$$\max_{\frac{d_e}{2} \leq d_f \leq d_e} \tau_{\text{nz}}^{\text{non-severe}}(d_f) = \underbrace{\tau_{\text{ident}}}_{<1\,\text{ms}} + \underbrace{\tau_{\text{com}}}_{<2.3\,\text{ms}} + \underbrace{\tau_{\text{CB}}}_{=2\,\text{ms}} \leq 5.3\,\text{ms} \tag{5.6}$$

The time the local relay may wait after the arrival of the fault for the confirmation of the fault can also be deduced from (5.5)

$$\tau_{\text{wait}}(d_f) = \underbrace{\tau_{\text{ident}}}_{\text{remote relay ident.}} + \tau_{\text{com}} + \underbrace{t_{\text{arrival}}^{\text{remote}} - t_{\text{arrival}}^{\text{local}}}_{=\frac{d_e - 2d_f}{c_w}}. \tag{5.7}$$

which is maximum for $d_f = \frac{d_e}{2}$, thus

$$\max\ \tau_{\text{wait}} = 3.3\,\text{ms}. \tag{5.8}$$

The obtained fault neutralization times of 3 ms and 5.3 ms for severe and non-severe faults respectively is compliant with the requirements evidenced in Section 5.1. It guarantees the line breakers capabilities are not exceeded and it may prevent the blocking of the MMC stations in most cases.

5.3. Full selective fault clearing strategy

This section proposes a full selective fault clearing strategy based on the proposed collaborative fault identification algorithm.

The primary sequence of the fault clearing strategy is summarized in the flowchart of Figure 5.10. The fault is first detected at the relay based on local measures using for instance a threshold on the rate of change of the voltage (ROCOV). The single ended fault identification algorithm is able, based on the value of the estimated fault parameters, to classify the fault into severe, non severe, or non identified, as presented in Section 5.2.1. The collaborative algorithm, based on the results of the single-ended fault identification algorithms at both the local and remote relays, allows to confirm or refute a fault occurred on the protected line. As proposed in Section 5.2, in the case of severe faults, only local data are used which allows the fault clearing to be sufficiently fast.

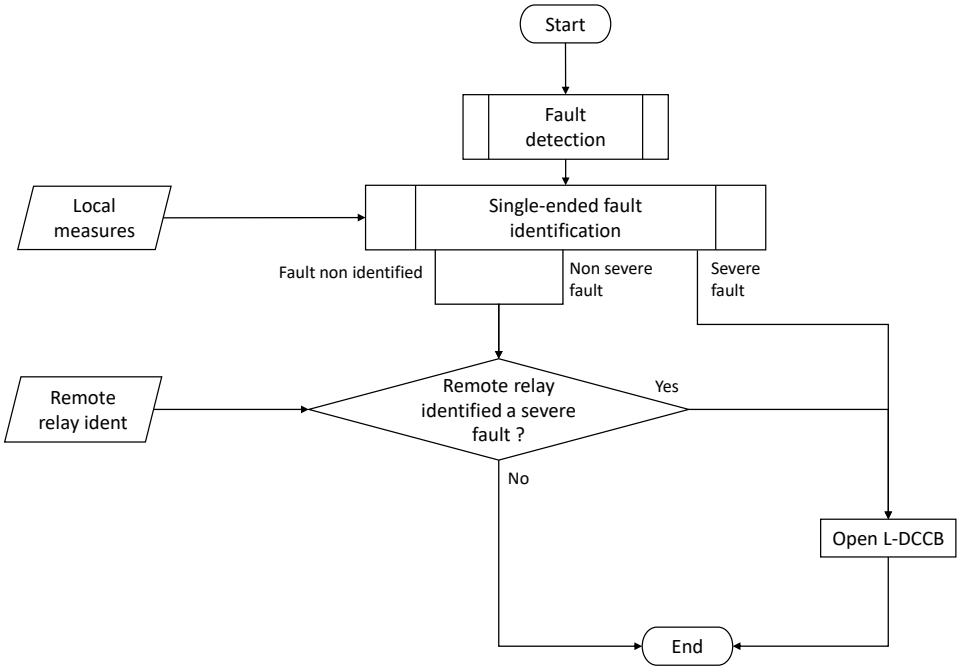


Figure 5.10.: Primary sequence of the full selective fault clearing strategy based on the collaborative fault identification.

In the context of protection strategies, failure modes (FM) refer to the specific ways a protection strategy may fail to operate due to a failure of a component, function or sub-system [Freitas et al., 2019]. Considering the fault clearing strategy of Figure 5.10, two specific FM that may have an important impact on the fault clearing strategy are further investigated:

- Communication failure between the close and remote relays. This may result in the incapacity for the relay to decide whether the fault is internal or not, in the case of non-severe faults. The primary sequence can be adapted so that, for non severe faults, if the communication from the remote relay is not received after some waiting time τ_{wait} evaluated in (5.8), the local relay identifies the fault as internal and trips the line breaker. This conservative approach reduces considerably the likelihood of dependability failures, at the cost of permitting some sympathetic trip for faults occurring on an adjacent line such as F_2, see Figure 5.6. In any case, a backup communication channel should be provided, allowing to significantly reduce he risk of communication failure.
- Line circuit breaker failure to operate. If the breaker controlled by the relay fails to open, the usual action is to trip the adjacent breakers connected to the same bus in order to clear the fault. Due to the important rise of the current as well as the delay required for the detection of the breaker failure, this may not be possible in the case of severe faults. The rise of the current at the output of converters 1 and 4 was evaluated for the same faults affecting line L_{14} as in Section 5.1.1. The time at which current exceeds the maximum breaking current $i_{\max} = 16\,\text{kA}$ is depicted in Figure 5.11. In the worst case, $i_{\max}$ is reached 5 ms after the fault arrival. Considering that the breaker failure to operate can be noticed in 2 ms (corresponding to the CB operating time), the neutralization time of the converter breaker $\tau_{\text{nz}}^{\text{converter}}$ is

$$\begin{aligned}\tau_{\text{nz}}^{\text{converter}} &= \tau_{\text{ident}} + \underbrace{\tau_{\text{CB}}}_{\text{CB failure detection}} + \underbrace{\tau_{\text{CB}}}_{\text{Converter CB opening}} \\ &< 5\,\text{ms}\end{aligned}.$$

  On this particular example, the opening of the adjacent breakers is thus compliant with the rise of current. This may however not be guaranteed for other lines as the rise of current depends for instance on the busbar configuration. One approach can be to choose consequently the specifications of the converter breakers, which may result in an important cost. An alternative approach assumes the breakers have a Fault Current Limiting (FCL) mode which allows to maintain the fault current through the breaker at a reasonable level [Jovcic et al., 2019]. For severe faults detected by a relay, the adjacent breaker at the same bus can be preemptively placed in FCL mode so that, in case of the line breaker failure, they can operate.

5.4. Sensors

The sensors that provide the relay with the current and voltage measurements are located at each extremity of the lines, as well as at the converter stations. This section presents an EMT model that accounts for the limited accuracy of the sensors,

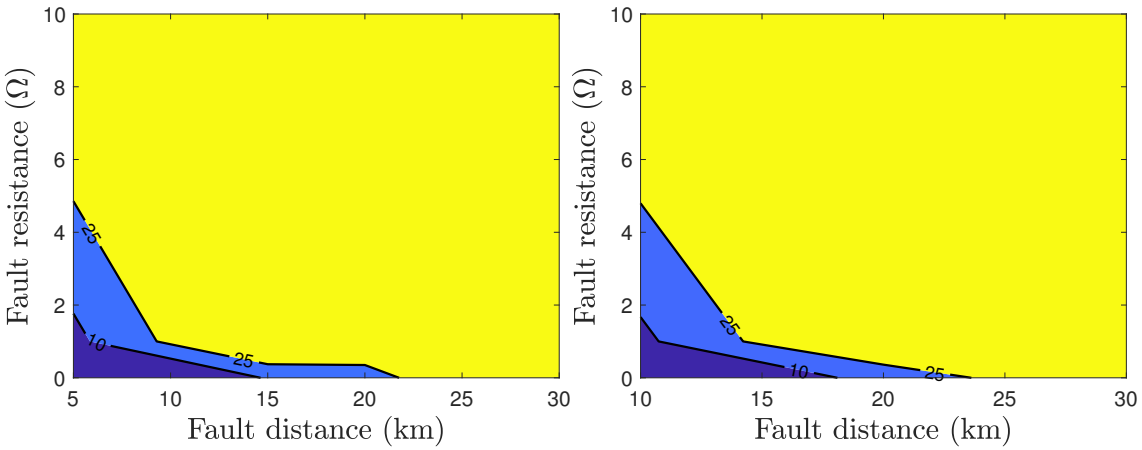


Figure 5.11.: For various faults affecting L_{14}, time at which the current through the converter breakers at station 1 (left) and 4 (right) reaches $i_{\max} = 16\,\text{kA}$. The yellow area corresponds to faults for which the converter breaker current never reaches $i_{\max}$. For faults corresponding to the light blue area, $i_{\max} = 16\,\text{kA}$ is reached in less than 25 ms after the fault detection at the relay. The dark blue area denotes faults for which the maximum breaking current is reached in less 10 ms. For the closest faults with $R_\text{f} = 0\,\Omega$, the maximum current is reached in 5 ms.

see Section 5.4.1. This model is then employed to study the influence of different sensor specifications on the behavior of the (single-ended) fault identification algorithm in Section 5.4.2.

5.4.1. Sensor model

This section introduces a behavioral model [2]for the voltage and current sensors independent from a specific technology. The main elements are reminded here, an more detailed description of the model implemented in the EMT software can be found in Appendix A.

The sensors can be described considering two main components: the transducer and the signal processing parts as presented in Figure 5.12. The transducer part converts the grid quantity $x(t)$ into a substantially proportional analog quantity $y(t)$. The transducer block diagram is presented in Figure 5.13.

The output of the transducer block is thus, before the saturation block,

$$Y(s) = \left[F(s)\exp(-\tau_\text{s} s)\,K_0\left(1+\frac{\Delta K_0}{K_0}\right) + B_0 + n_\text{s}(s)\right] X(s) \tag{5.9}$$

[2]This section uses a model of the sensor implemented in EMT software that results from a collaborative work at the Supergrid Institute. The author wishes to acknowledge for their decisive participation in this work G.Renon, F.Calligari and P.Torwelle.

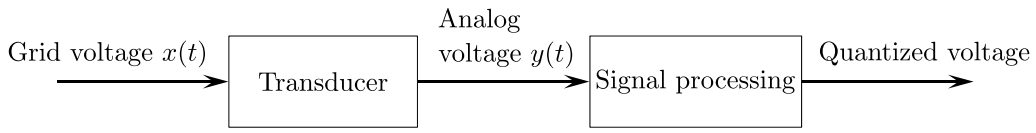


Figure 5.12.: Sensor model considering two main components: a transducer and a signal processing block.

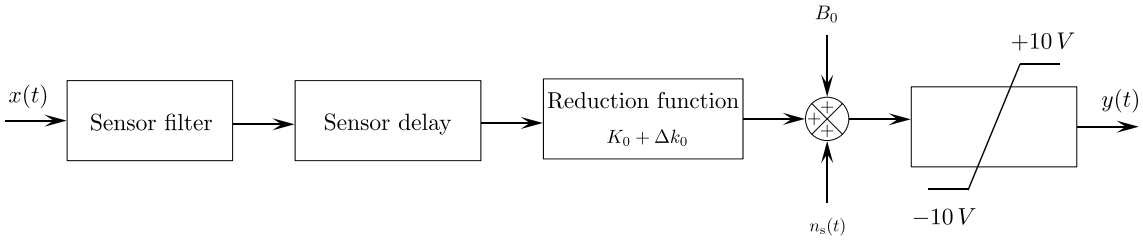


Figure 5.13.: Block diagram of the transducer part of the developed sensor model.

where $F(s)$ is a first order filter that represents the low pass effect of the measurement device with cut-off frequency f_0^{φ}. The total sensor delay $\exp(-\tau_s s)$ is considered constant. The reduction function

$$F_R = K_0 \left(1 + \frac{\Delta K_0}{K_0}\right) \tag{5.10}$$

represents the reduction of the grid quantity to a reduced voltage by a factor

$$K_0 = \frac{|x_{max} - x_{min}|}{|y_{max} - y_{min}|} = \frac{|x_{max} - x_{min}|}{20}.$$

This factor can endure a drift $\frac{\Delta K_0}{K_0}$ typically due to external effects such as temperature or pressure. An additive offset B_0 is considered and may also fluctuate due to external conditions but is here considered as a fixed maximum value. The noise $n_s(t)$ accounts for the internal error of the sensor for instance due electromagnetic perturbations and is represented by a centered white noise. The output error is then, neglecting the sensor filter effects,

$$\begin{aligned} y(t) - K_0 x(t) &= F_R x(t) + B_0 + n_s(t) - K_0 x(t) \\ &= \Delta K_0 x(t) + B_0 + n_s(t) \\ &= x(t)\left(\Delta K_0 + \frac{B_0}{x(t)} + \frac{n_s(t)}{x(t)}\right) \end{aligned}$$

Though the grid quantity $x(t)$ may fluctuate, we use the base value x_{base} to specify the total error with respect to the grid quantity as a percentage

$$\varepsilon(t) = \frac{1}{K_0}\left(\Delta K_0 + \frac{B_0}{x_{base}} + \frac{n_s(t)}{x_{base}}\right). \tag{5.11}$$

The maximum error (considering the peak value for the white noise) is also known as the *accuracy class* of the sensor. The separation of the error into three different terms, namely a drift of the sensitivity factor, a constant offset and a white noise, is advantageous as the distribution of the total error may be different depending on the considered technology. For instance, the materials used for RC-dividers principle are strongly temperature dependent. Thus, the long term variations of their sensitivity factor are important and difficult to reduce. This is the same for optical elements used in optical sensors. On the other hand, it is known that for MEMS technologies [Moagar-Poladian and Moagar-Poladian, 2014] the long term variations of the offset are important due to aging of the mechanical elements. Finally, contrary to optical sensor [Chatrefou et al., 2000], RC-dividers and MEMS sensors operate with electrical components close to the high voltage equipment which leads to high electromagnetic disturbances. Knowing the most influent error sources can guide the choice of the developed principles for new sensors.

In the signal processing part, the analog voltage $y(t)$ is discretized and reconverted into a quantity of the same scale as the input quantity. The signal processing block diagram is presented in Figure 5.14. It is composed of an anti-aliasing filter of bandwidth f^{AAF} and an analog to digital converter with sampling frequency f_{s} and resolution Q. The bandwidth of the sensor is determined by the lowest cut-off frequency, *i.e.*,

$$BW = \min\left\{f^{\mathrm{AAF}}, f_0^{\varphi}\right\}.$$

The detailed sensor model in this section is implemented in EMT software so that

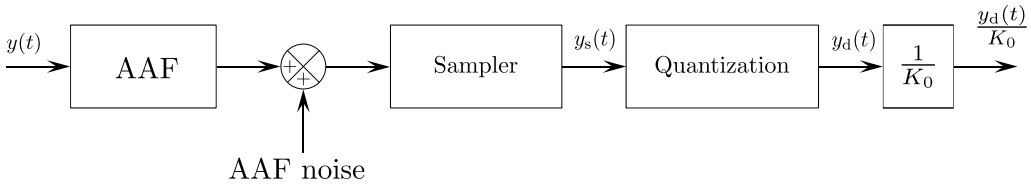


Figure 5.14.: Block diagram of the signal processing part of the sensor developed model.

the influence of the different specifications on the algorithm behavior can be studied, see Section 5.4.2.

5.4.2. Impact of the sensor characteristics

In this section we consider the four station meshed grid from Figure 4.1 with the rigid bipole configuration. Reference values for the sensor specifications have been selected based on existing requirements from [IEC 61869-15, 2018] as well as the the requirements of the algorithm, see Table 5.2. The total error ratio (5.11) is the sum of the sensitivity drift, the maximum offset and the peak of the white noise,

Table 5.2.: Reference specifications for the voltage and current measurement sensors.

Parameter	Voltage sensor	Current sensor
Max/min input value	$[-510, 510]$ kV	$[-30, 30]$ kA
Sampling frequency (kHz)		1000
Bandwidth (kHz)		300
Total error ratio (%)		1
Sensitivity drift (%)		0.4
Maximum offset (%)		0.3
Peak Gaussian noise (%)		0.3
Resolution (bits)		12
Transducer delay (ms)		0.1

all given as a percentage of the base current or voltage. With the considered grid, $\varepsilon = 1\%$ amounts to a maximum error of 3.2 kV and 30 A.

Due to the important simulation time of a grid equipped with the sensor models in EMT software, the effects of varying some of the sensor specifications on the behavior of the fault identification algorithm is only studied on a specific fault case. This particular example is first detailed considering the reference sensor specifications from Table 5.2.

The influence of some of the main sensor characteristics - total error and its components - the sensitivity drift, offset and white noise- are then studied individually. The effect of the sampling frequency and bandwidth is finally detailed. The impact of those different parameters on the dependability of the algorithm as well as on the precision of the estimated fault parameters and on the length of the required measurement window is investigated. When not specified otherwise, the sensor parameters employed are the ones from Table 5.2.

5.4.2.1. Reference case

The rigid bipole configuration is considered for this study. Consider a pole-to-ground fault affecting the line L_{14} at 210 km from station 1 with $R_{\mathrm{f}}^{*} = 60\,\Omega$. The fault identification is triggered at the relays R_{14}, for which $d_{\mathrm{f}}^{*} = 210$ km, and at R_{41} for which $d_{\mathrm{f}}^{*} = 90$ km. Those two situations are referred to, in the following paragraphs, as the *remote* fault and *close* fault, respectively. The behavior of the algorithm at the relays R_{41} is now detailed. The evolution of the estimated fault parameters and the contour plot of the cost function are plotted in Figure 5.15 for the 9 first iterations after which the fault is identified on the line L_{14}. The estimated fault parameters are $\widehat{R}_{\mathrm{f}} = 71\,\Omega$ and $\widehat{d}_{\mathrm{f}} = 68$ km. The evolution of the modeled voltage and current are compared with the EMT data obtained with the sensor model in Figure 5.16. The effect of the reduced bandwidth is observable at the beginning of

the wave, corresponding to the arrival of the aerial mode. The random noise affects more the negative pole quantities as the evolution on the healthy pole is reduced which, by contrast, gives more weight to the error sources.

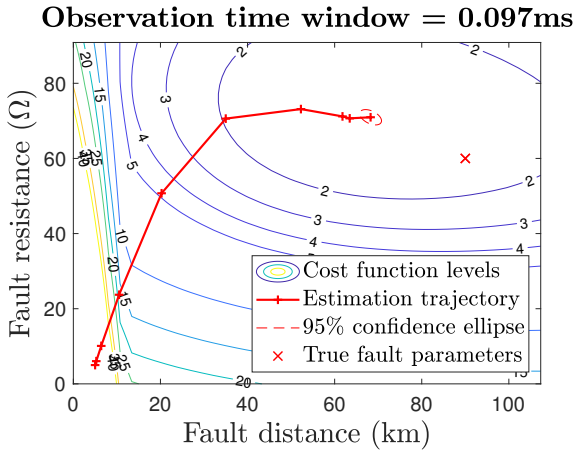


Figure 5.15.: Evolution of the estimated fault parameters and cost function contour when considering sensor model in the reference scenario. The fault is identified after 9 iterations and the estimated fault parameters are $\widehat{R}_{\mathrm{f}} = 71\,\Omega$ and $\widehat{d}_{\mathrm{f}} = 68\,\mathrm{km}$.

5.4.2.2. Impact of the accuracy class

The impact of the accuracy class is first studied as a whole and the significance of the three components is investigated as a second step. Various error ratios are applied such that $\varepsilon = \{0.5\,\%,\ 1\,\%,\ 2\,\%,\ 4\,\%,\ 6\,\%\}$ while its decomposition into the sensitivity drift (0.4ε), maximum offset (0.3ε) and noise (0.3ε) is kept identical. The algorithm successfully identifies the fault as internal in all the cases. The main observed impacts are a decrease of the accuracy of the estimated fault parameters, which is evidenced in Figure 5.17.

The impact of the different sources of error is then studied individually. For each of the three error sources, the two others are set to 0 so that the total error ratio is due of a single error source. The applied total error is then $\varepsilon = \{0.15\,\%,\ 0.3\,\%,\ 0.6\,\%,\ 1.2\,\%,\ 2.4\,\%\}$. Most importantly, the algorithm always identify the fault as an internal fault. The main findings are

- An increase in the white noise extends the observation window required for the identification. It also decreases the accuracy of the estimated fault distance.
- A larger offset only results in an less accurate estimate of the fault resistance, see Figure 5.19 (left). That can be explained given that the offset impacts the amplitude of the current and voltage waves, it can thus be "compensated for" by the estimated fault resistance which also impacts the amplitude of

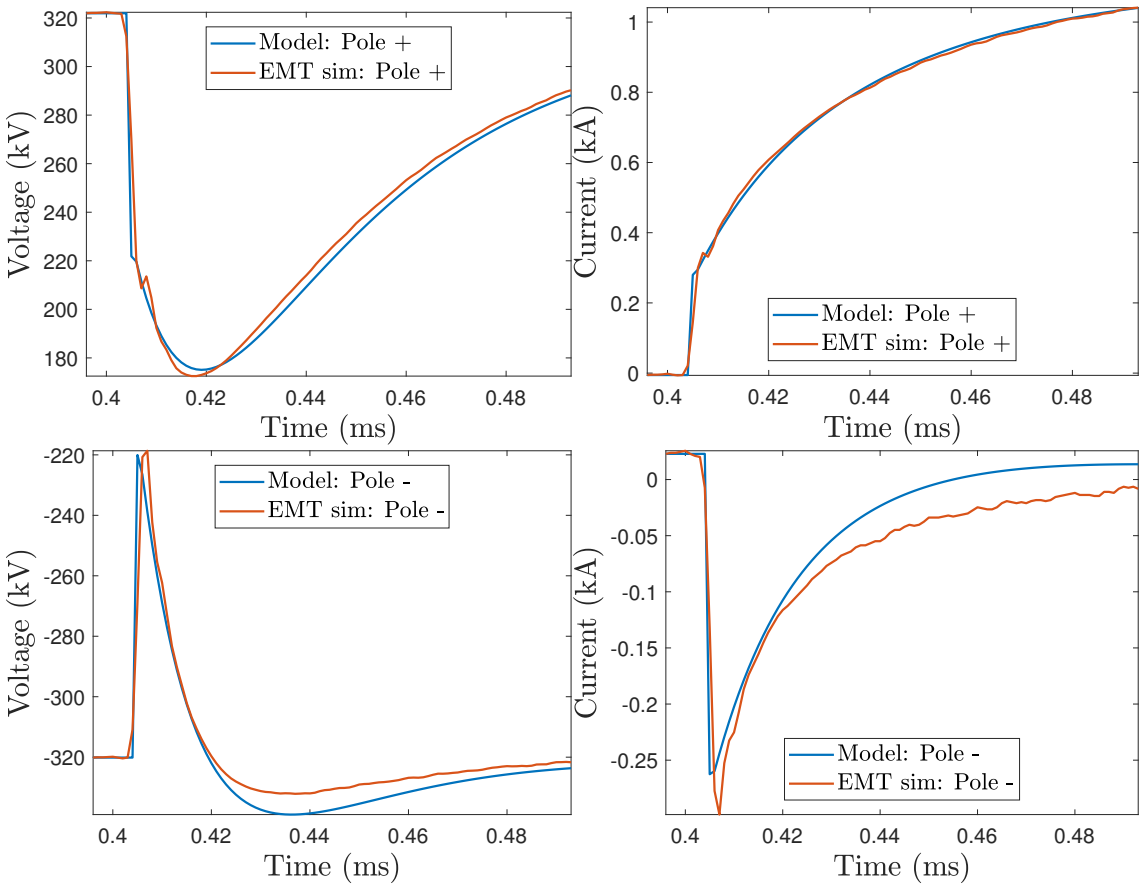


Figure 5.16.: Evolution of the modeled voltage and current compared with the EMT data obtained with the sensor model, in the reference case. The fault parameters employed in the parametric model are the ones obtained after the fault identification $\widehat{R}_{\mathrm{f}} = 71\,\Omega$ and $\widehat{R}_{\mathrm{f}} = 68\,\mathrm{km}$.

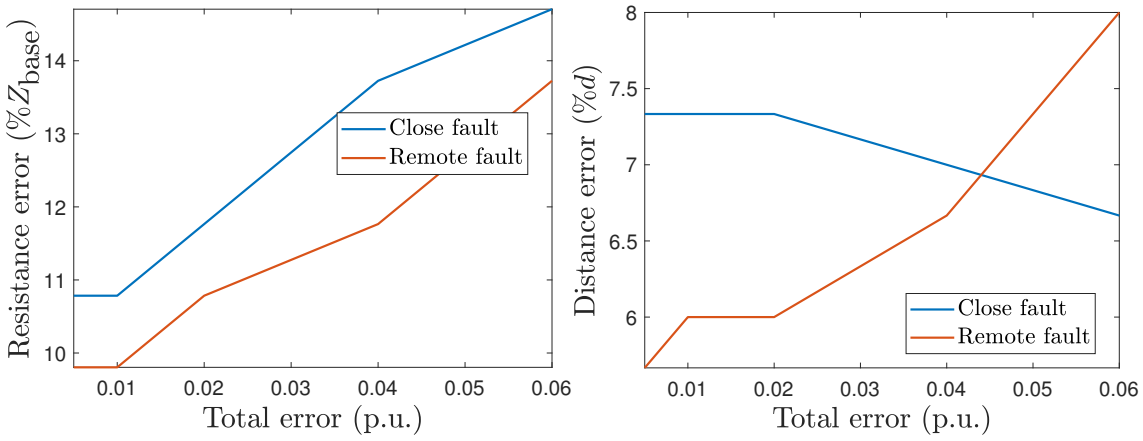


Figure 5.17.: Impact of the total error on the precision of the estimated fault resistance (left) and distance (right).

the modeled waves. On the other hand, the accuracy of the fault distance estimate is not much impacted.

- The sensitivity drift has only minor impact on the behavior of the algorithm besides increasing the value of the cost function for the estimated fault parameters (given in p.u.), see 5.19 (right). This can be explained considering that the voltage model v^{m} is proportional to the measured pre-fault voltage $v_{\mathrm{bf}}^{\mathrm{meas}}$ through the voltage at the fault location (2.32), thus

$$v^{\mathrm{m}} \propto v_{\mathrm{bf}}^{\mathrm{meas}}$$

where $\propto$ represents the proportionality relation and the subscript meas refers to measured quantity prone to measurement errors. Assuming the only source of errors is due to the drift of the sensitivity factor the measured voltage is proportional to the true voltage,

$$v_{\mathrm{bf}}^{\mathrm{meas}} = \Delta K_0 v_{\mathrm{bf}}.$$

The cost function (considering the voltage part only) is then

$$\begin{aligned} v^{\mathrm{m}} - v^{\mathrm{meas}} &\propto \Delta K_0 v_{\mathrm{bf}} - \Delta K_0 v, \\ \underbrace{v^{\mathrm{m}} - v^{\mathrm{meas}}}_{\text{Available cost}} &\propto \Delta K_0 \underbrace{(v_{\mathrm{bf}} - v)}_{\text{Ideal cost}} \end{aligned} \tag{5.12}$$

which shows the sensitivity factor tends to increase the cost by a multiplicative factor. As the cost function also includes the current measurements, the proportional factor in (5.12) may be different for the total cost function. The key assumption here is that the drift ΔK_0 is constant over short time periods (*e.g.* over a millisecond).

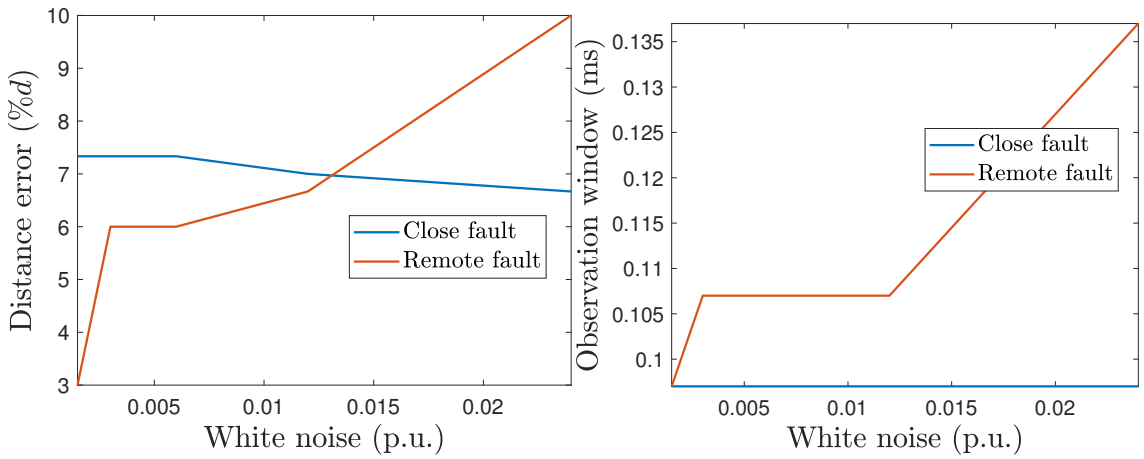


Figure 5.18.: Impact of the white noise on the precision of the estimated fault resistance (left) and distance (right)

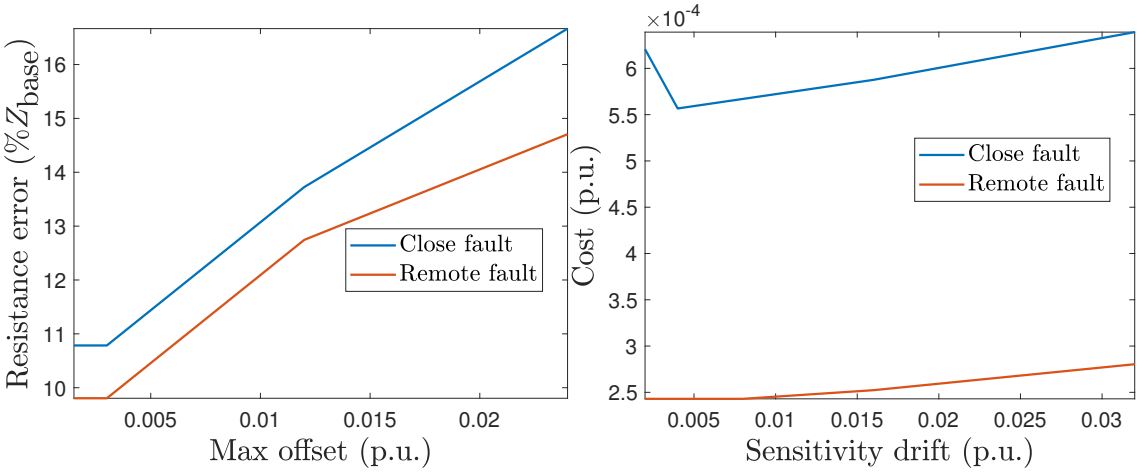


Figure 5.19.: Impact of the maximum offset (left) on the accuracy of the estimate fault resistance and of the sensitivity drift (right) on the value of the cost function reached by the algorithm.

5.4.2.3. Bandwidth and sampling frequency impact

The influence of the sampling frequency and the bandwidth are now investigated. The error terms are set to their reference value, see Table 5.2. The different bandwidth considered for the sensors are $\{40,\ 100,\ 200,\ 300,\ 500\}$ kHz while the sampling frequency of he sensors is kept at $f_s = 1\,\text{MHz}$.

A smaller bandwidth tends to slightly degrade the performances of the algorithm. The length of the required observation window in particular may fluctuate but generally has higher values for smaller bandwidths, see Figure 5.20.

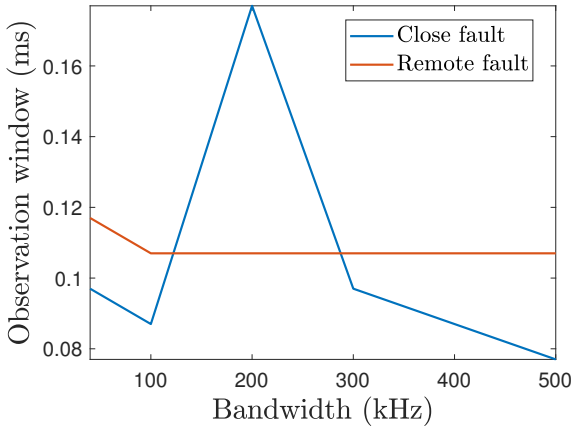


Figure 5.20.: Impact of the bandwidth on the length of the required observation window for the fault identification

The influence of the sampling frequency is finally detailed. The fault is simulated with various sampling time $T_s \in \{1,\ 2,\ 4\ ,\ 6,\ 10\}\ \mu$s (the reference value is $1\,\mu$s).

The bandwidth must be adapted to fulfill Shannon's theorem $BW \leq \frac{f_s}{2}$. The ratio between the sampling frequency and the bandwidth was kept at the same value as in the reference case, *i.e.*, $\frac{f_s}{\mathrm{BW}} = \frac{10}{3}$. The simulation results suggest the sampling time has an important impact on the performances of the fault identification algorithm. In particular, algorithm must be adapted to the various sampling frequencies to obtain acceptable performances. The considered adaptation are

- In the accuracy test, the threshold on the area of the confidence region (3.36) is, all things being equal, proportional the number of sampling points. A slower sampling frequency thus results in larger confidence areas. The threshold (3.37) on the size of the 95% confidence region is thus adjusted proportionally to the sampling frequency variation.
- The number of iterations the optimization algorithm for a given observation window length is kept constant. This leads to adjust κ so that one iteration is performed every $10\,\mu$s.

The different settings are summarized in Table 5.3. Under those adjustments, the

Table 5.3.: Tuning parameters of the identification algorithm adapted to cope with slower sampling frequencies.

Sampling period (μs)	1	2	4	6	10
Threshold c_{95}	20	40	60	50	30
Number of iterations κ	1	2	4	6	10

identification algorithm is able to identify the two considered fault cases for frequencies such that $f_s \geq \frac{1}{6\times10^{-6}} = 167\,\mathrm{kHz}$. For $f_s = 100\,\mathrm{kHz}$ however, the algorithm fails to identify the remote fault case and has a very poor accuracy for the close fault. Furthermore, the performances of the algorithm are, to a large extent, degraded by smaller sampling frequencies. The required observation window may be 2 times larger in the worst case as shown in Figure 5.21 and the precision of the estimated parameters also decreases, see Figure 5.22.

5.4.3. Main observations on the sensor specifications

The influence on the algorithm performances of some of the main specifications of the sensors have been studied using a sensor model implemented in EMT software, see Section 5.4.1. The conclusions drawn have nevertheless a limited scope as only two fault cases have been simulated due to simulation time constraints. The results on those two cases indicate the algorithm is generally robust to less accurate sensors. The precision of the estimated parameters may be degraded due to higher noise or sensitivity drift while the measurement window is scarcely impacted. The sampling frequency is, however, crucial to the algorithm as dependability failures have been observed for $f_s = 100\,\mathrm{kHz}$ and the other performances are significantly degraded.

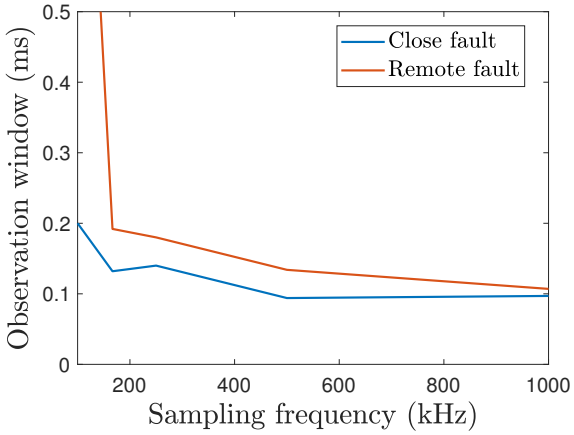


Figure 5.21.: Impact of the sampling frequency on the length of the measurement window required for the fault identification. For $f_s = 10\,\text{kHz}$ the algorithm fails to identify the remote fault.

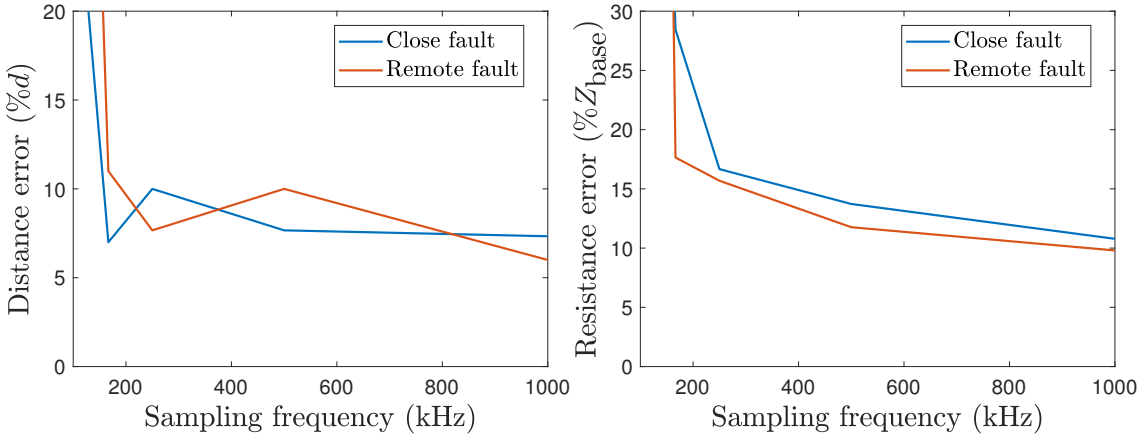


Figure 5.22.: Impact of the sampling frequency on the accuracy of the estimated fault parameters. For $f_s = 10\,\text{kHz}$ the algorithm fails to identify the remote fault.

Those results are summarized in Table 5.4 where recommendations for the specifications are formulated and compared with existing standards [IEC 61869-15, 2018]. The requirement on the sampling frequency appears to be the most stringent one while the algorithm seems to cope with relatively low accuracy classes. The recommended sampling frequencies $f_s \geq 500\,\text{kHz}$ is nevertheless in line with numerous state-of-the-art methods and may for instance be achieved with technologies based on optical sensors, shunt resistance or Rogowski coil [Tzelepis et al., 2020].

Parameter	Recommended range	Existing norm
Bandwidth	$\geq 100\,\text{kHz}$	max: $500\,\text{kHz}$
Sampling frequency	$\geq 500\,\text{kHz}$	max: $100\,\text{kHz}$
Accuracy class	$\leq 4\%$	$[0.1 - 1]\,\%$

Table 5.4.: Comparison of the recommended sensor specifications with some of the existing norm standards. The sampling frequency requirements appear to be the most stringent one.

5.5. Conclusion

The proposed fault identification algorithm can be integrated into a full-selective fault clearing strategy. Such strategies mitigate the impact of faults on the HVDC the healthy parts of the grid can operate continuously. The primary and backup sequences detailed in Section 5.3 allows to handle all faults occurring on a transmission line by having recourse, for non-severe faults, to communication with the other relay protecting the same line. The compliance of the proposed strategy with the rise of the current and the circuit breakers capabilities is investigated in Section 5.1.1. It is shown that, on the considered test grid, the use of DC inductances at the end of each line is not required to clear the fault. The fast clearing of the fault also limits the faults that trigger the blocking of the converter stations, as presented in Section 5.1.2. Finally, the detailed sensor model introduced in Section 5.4 allows to study the robustness of the algorithm to various specifications, in particular regarding the accuracy and sampling frequency of the sensors.

6. Perspectives and conclusion

This chapter gathers different topics of interest which have been the subject of preliminary investigations but would require more in depth studies. It has been noticed in Chapter 4 that the faults occurring very close to the stations may cause the maloperation or degraded performances of the identification algorithm. This issue can be tackled through a multiple hypothesizes approach, as presented in Section 6.1. The assumptions made on the fault model in Chapter 2 are challenged in Section 6.2 where the case of lightning strikes is further considered. Possibilities to improve the execution times of the algorithm are detailed in Section 6.3. A general conclusion is then presented in Section 6.4.

6.1. Very close faults and multiple hypothesizes approach

The extensive simulations performed for both the mono-conductor and multi-conductor configurations in Sections 4.2.2 and 4.3.2 show the proposed approach performs well on a wide range of fault cases. We focus here on the mono-conductor case though the general methodology can be adapted to the multi-conductor case. The behavior of the algorithm for faults occurring close to a station rises a number of issues. Consider for instance the fault depicted in Figure 6.1 affecting line L_{12} and located very close (*e.g.* $d_{\mathrm{f}} \leq 1\,\mathrm{km}$) to station 2.

- For the relay R_{21} the fault is very close and should be identified very fast to ensure the line breakers are tripped as soon as possible. Such faults are characterized by a a large number of reflected waves as well as a reduced impact of the ground resistivity for the current return path. .
- For the relay R_{12} at the other end of the faulty line, the fault is particularly far and the algorithm is likely to identify it as external. Such dependability failures have been observed in Section 4.2.2.
- For the relay R_{42}, the fault may appear as internal due to the very small distance from the station 2. Such security failures have been observed in Section 4.2.2.

In Section 5.2, a collaborative fault identification algorithm was introduced to solve some of the issue raised by the fault depicted in Figure 6.1. The communication

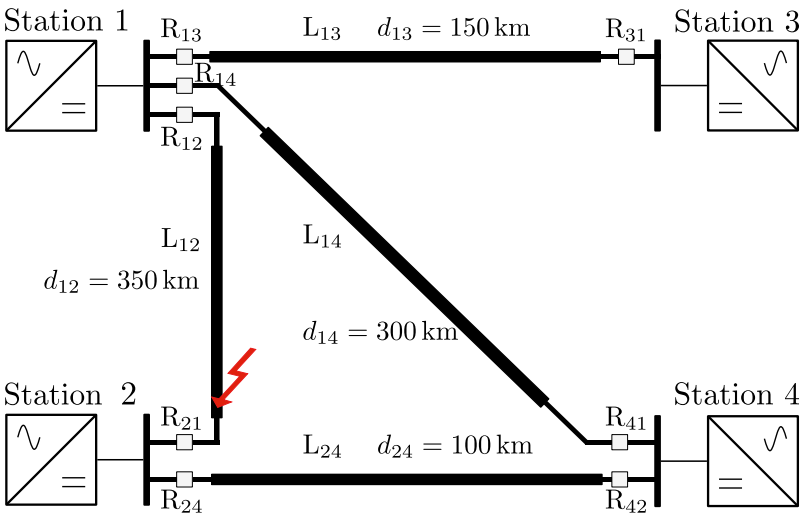


Figure 6.1.: Situation of a fault located very close to the station 2 that may cause a dependability failure at relay R_{12} and a security failure at relay R_{42}.

with the other relay protecting the same line increases the efficiency of the fault identification. Here we address similar situations of faults occurring very close to a station without using the results of the identification from a distant relay. The use of several initial points $\mathbf{p}_0$ for instance may help the algorithm to converge towards the true fault parameters. More generally, the model $\mathbf{y}^{\mathrm{m}}(t, \mathbf{p})$ employed for the fault parameter estimation may be adjusted to the specific cases aforementioned.

Faults that occur very close to a given relay (R_{21} in Figure 6.1) require the modeling of three to four waves but the influence of the ground can be neglected. This leads to consider only the *physical part* within the model developed in Chapter 2.
Faults that occur on an adjacent line but very close to the remote end of the protected line (fault seen from R_{42} in Figure 6.1) can be modeled considering that many waves occur between the fault and the remote station, but for those the ground effect can be neglected. The impact of the soil resistivity mostly comes from the propagation through the whole line (here line L_{24}). The filter corresponding to the whole line length is thus applied. In the considered example, the forward voltage waves that arrive at the station 4 are then

$$v_{\mathrm{f},4}^{\mathrm{m}}(k, \mathbf{p}) = g\left(z^{-1}, d_{24}\right) \otimes v_2^{\mathrm{m}}(k, \mathbf{p}), \tag{6.1}$$

where v_2^{m} is the modeled voltage at the station 2 neglecting the ground effects. In the structure of (6.1) the soil effect is represented but it does not depend on the fault distance as the main distortion comes from the propagation through the line L_{24} of known length d_{24}.

We propose to use four different identification algorithms in parallel at the same relay, as summarized in Table 6.1. Each identification algorithm corresponds to a different suspected fault area location and differ by the initial point $\mathbf{p}_0$. The first

three algorithms are run assuming that the fault is located in the protected line. The last one is initialized with a value of the parameter corresponding to a fault outside of the protected line e, at a distance $d_\mathrm{f} = d_e + 1$, where d_e is the length of the protected line. In the example of Figure 6.1, relay R_{42} considers three hypothesizes of faults affecting line L_{24} and one of a fault affecting the adjacent line L_{12} located close to station 2. For each hypothesis, the model employed by the identification algorithm must remain consistent with the hypothesis. For instance, the algorithm run according to the adjacent line fault hypothesis may not converge to a fault located on the protected line. As the four different estimation algorithms are run in parallel, the fault is identified as soon as one of the estimated fault parameters satisfy both the accuracy and validity tests. When that several fault parameters satisfy the two tests at once, the estimate that for which the cost function is the smallest is selected. This procedure is similar to the one employed in the case of the identification of the faulty segment in hybrid lines, see (3.38).

Table 6.1.: Multiple hypothesizes approach: several fault identification algorithms are launched in parallel, each one is associated to a suspected fault location and parameterized by the initial point $\mathbf{p}_0$. The models employed by some hypothesizes may be adapted according to the suspected fault location. The first three suspected locations correspond to an internal fault whereas the last one corresponds to an external fault.

Suspected location	Initial point $\mathbf{p}_0$	Modeling assumptions
Close fault	$(1\,\Omega, 1\,\mathrm{km})$	Numerous TW with negligible soil effects
Average fault	$(5\,\Omega, 5\,\mathrm{km})$	Standard model
Remote fault	$(50\,\Omega, 0.9 d_e\,\mathrm{km})$	Standard model
Adjacent line fault	$(1\,\Omega, d_e + 1\,\mathrm{km})$	Numerous reflection with remote station, soil main impact due to line e

The proposed approach is illustrated on an example. Consider a fault occurring on line L_{12} at a distance $d_\mathrm{f}^* = 0.2\,\mathrm{km}$ from station 2 (and 349.8 km from station 1) with an impedance of $R_\mathrm{f}^* = 4\,\Omega$, as depicted in Figure 6.1. We successively consider the identification at the relays R_{21}, R_{12}, and R_{42}. Each relays employ the multiple hypothesizes approach described previously.

At the relay R_{21}, for which the fault is located on the protected line at a distance $d_\mathrm{f}^* = 0,2\,\mathrm{km}$, the fault is identified after only four iterations and the initial point corresponding to a very close fault is selected. The confidence region areas and the value of the cost function for the four different hypothesizes are compared in Figure 6.2. The close fault hypothesis has by far the smallest confidence region and cost. In such case, it is imposed that a minimum number of iterations (arbitrarily set to 4) is performed before the algorithm may stop. This ensures the identification is based on a minimum amount of data. This explains why the algorithm does not stop after the first iteration even though the area of the confidence region is smaller

than the threshold. The evolution of the fault parameters as well as the current and voltage of the parametric model and the EMT data are provided in Figure 6.3. It is confirmed the ground effect is negligible as, for instance, the voltage shape resembles the one of a pure step.

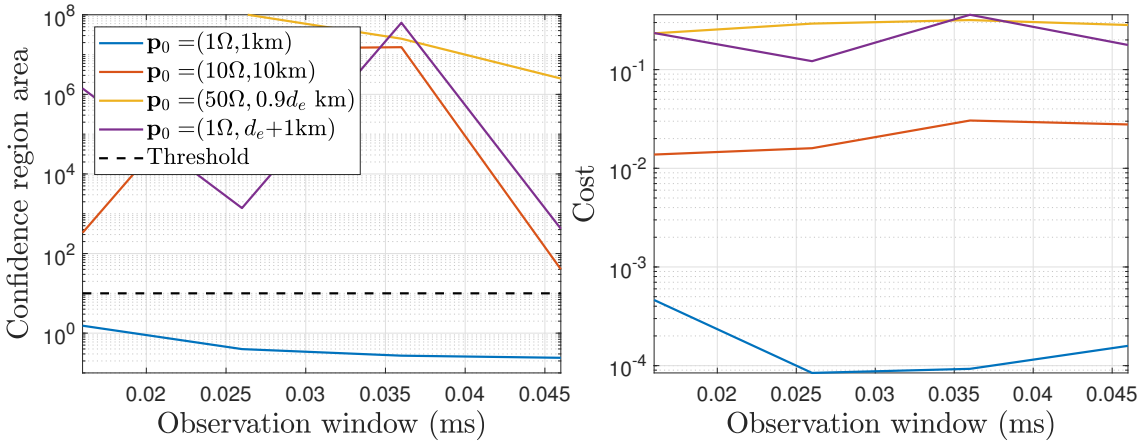


Figure 6.2.: Close fault case: evolution of the confidence region area (left) and cost function (right). The closest initial point allows to converge in only a few steps.

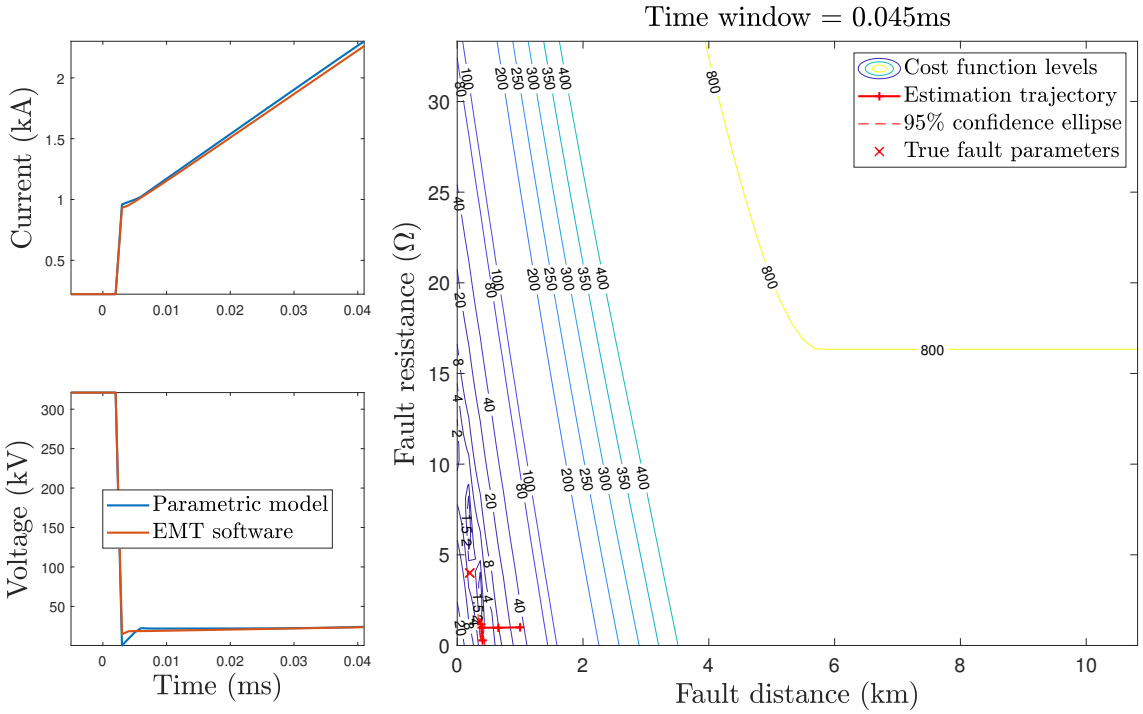


Figure 6.3.: Close fault case: the closest initial point allows to converge in only a few steps. The corresponding model neglects the impact of the soil resistivity but includes up to four traveling waves.

Consider now the algorithm at the relay R_{12}, for which the fault is located on the protected line at a distance $d_{\mathrm{f}}^{*} = 349,8\,\mathrm{km}$. In this case the fault is identified after 17 iterations by the estimation algorithm initialized with $\rho_{0,3}$,which corresponds to a remote fault hypothesis, as seen in Figure 6.4. The trajectory of the estimated fault parameters and the fit between the model and EMT simulation are displayed in Figure 6.4. The estimated parameters first evolve towards a high impedance fault located in the middle of the line before it converges towards the end of the line, closer to the true fault parameters. Despite the moderate accuracy of the estimated fault parameters, the fault is correctly interpreted as an internal fault occurring close to the remote station. By contrast, the identification algorithm run according to the adjacent fault hypothesis leads to a much higher cost function, see Figure 6.4 (right), and is not seleceted.

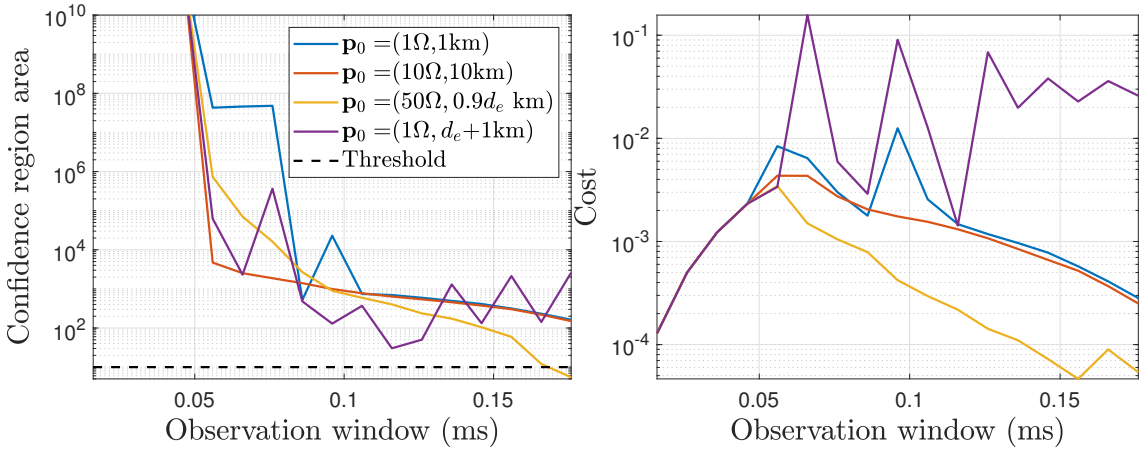


Figure 6.4.: Remote fault case: evolution of the confidence region area (left) and cost function (right). The remote internal fault hypothesis (in yellow) allows to converge in 17 steps.

Consider at last the identification at the relay R_{42} for which the fault is external. Due to the fault being of low impedance and very close to the station 2 it is likely to be misinterpreted as an internal fault. In this situation, the utilization of the adjacent fault model outperforms the results given by the internal fault model, see for instance the cost in Figure 6.6 (right). This allows the algorithm to correctly identify the fault on the line L_{12} after 5 iterations. The corresponding fault parameters trajectory and cost function function contour are provided in Figure 6.7 (right). The evolution of the voltage and current, see Figure 6.7 (left), significantly resembles the one of Figure 6.5 though the first corresponds to an internal fault and the second to an external fault.

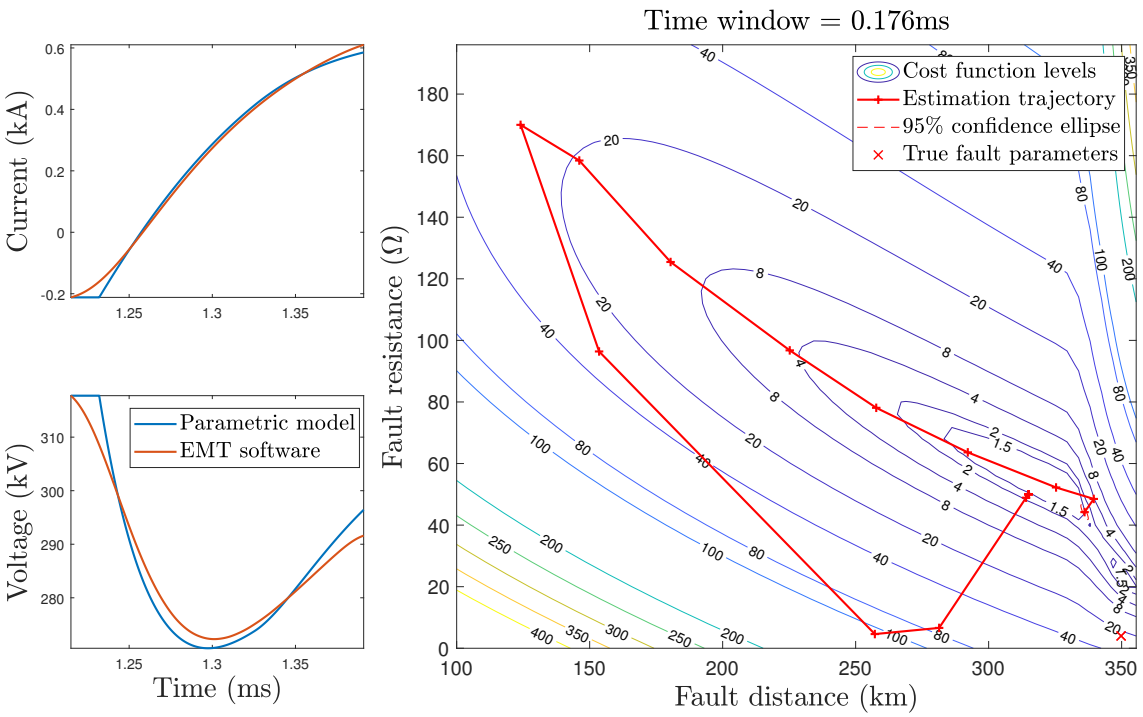


Figure 6.5.: Remote fault case: considering the remote internal fault hypothesis leads to a convergence in 17 steps.

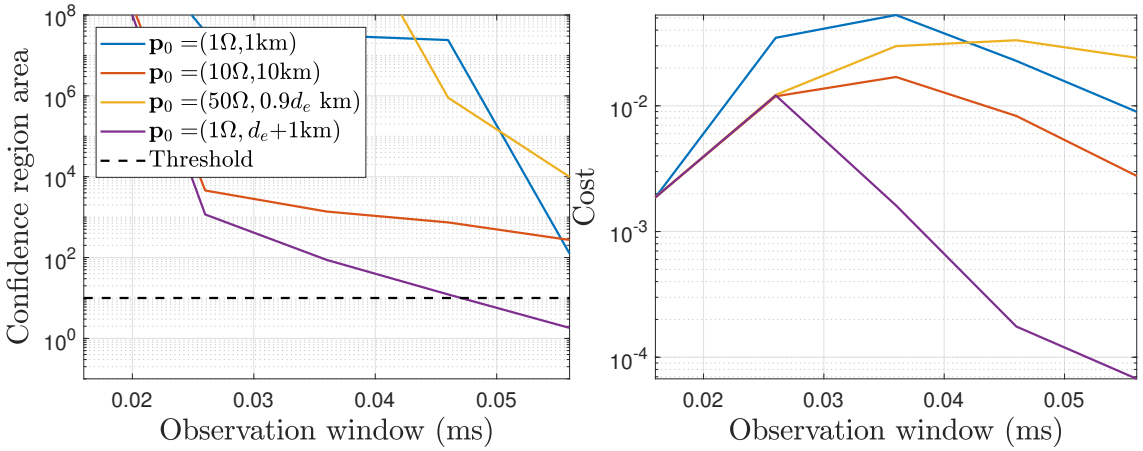


Figure 6.6.: Adjacent fault case: evolution of the confidence region area (left) and cost function (right). The estimated fault parameters obtained according to the adjacent line fault hypothesis (purple) lead to a better fit of the measurements than ones obtained with the internal fault hypothesizes.

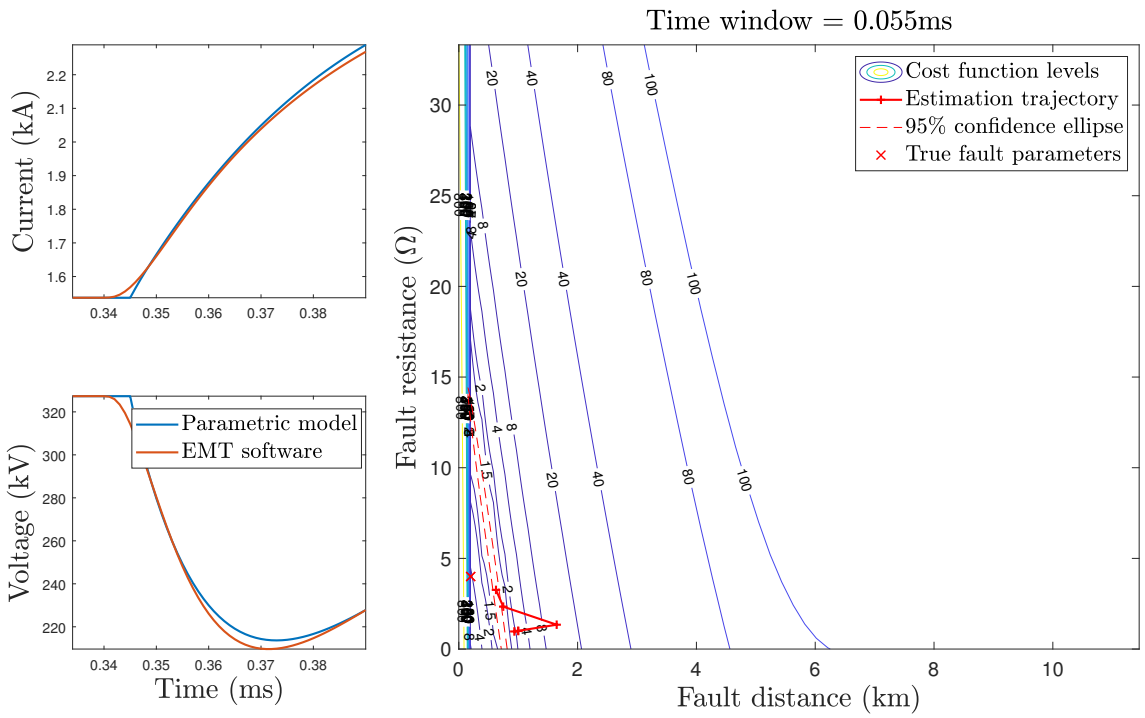


Figure 6.7.: Adjacent fault case: the adjacent line fault hypothesis fits better the received measurements despite the resemblance with the remote internal fault situation from Figure 6.5. The indicated fault distance (right) corresponds to the distance on the adjacent line.

The multiple hypothesizes approach developed in this section allows one to tackle some of the issues encountered, in particular the dependability and security failures occurring for faults located close to a station. It also increases the robustness of the algorithm for the most severe fault occurring at a very small distance. This, however, comes at the cost of an increased complexity as several models must now be considered in parallel and compared in the identification process. The performing of multiple parameter estimations in parallel also increases the computational burden of the algorithm.

6.2. Response to a lightning event

The main causes of faults on overhead lines include pollution, vegetation and lightning strikes [De Almeida et al., 2009]. The pollution and vegetation can be reasonably represented as a switch between the faulted conductor(s) and the ground closing at $t = t_f$. Connected in series with the switch, the fault resistance accounts for the different elements in the current path, such as the arc resistance or the grounding of the tower [Terzija and Dobrijevic, 2007]. Though the fault resistance may vary with time, it can be considered as constant in the first millisecond

[De Andrade Suárez and Sorrentino, 2010]. Modeling the effect of a lightning strike leading to a fault as a closing switch is, however, questionable [Xemard et al., 2021]. This section investigates some of the multiple questions raised by the representation of lightning strikes for fault identification.

Lightning strikes can be classified depending on the point of impact. Most lightning strikes the ground wire or the pylon and are thus called *indirect strikes*. A lightning can nevertheless strike directly a conductor, causing a *direct strike*. For a lightning to cause a fault, the insulator string between the poles and the pylon must flash, which occurs when the insulator voltage goes above its withstand voltage. Due to the insulation constraints being more stringent for HVDC than for HVAC OHL, the probability that a indirect strike causes a fault is reduced. By contrast, direct strikes with only a moderate current can still cause a short-circuit due to the high surge impedance of the pole conductors. Consequently, we focus on direct lightning strikes, which may or may not cause a fault. The bipolar configuration is considered in this section.

The lightning is generally represented by an controlled current source connected to the impact point [You et al., 2010]. The waveform of the lightning current is difficult to characterize random, though usual waveforms for dielectric testing purposes are defined in international standards [CIGRE WG01 SC33, 1991]. Among those, we consider the double exponential waveform

$$i_{\text{light}} = i_0 \left(\exp\left(-\frac{t}{\tau_1}\right) - \exp\left(-\frac{t}{\tau_2}\right) \right) \tag{6.2}$$

where τ_1, τ_2, and i_0 are adjustable parameters. Two current waveform are considered, see Figure 6.8. The current with the highest amplitude causes a fault whereas the other one only induces a disturbance that propagate along the phase conductor.

An accurate simulation of the lightning strike phenomenon in EMT software involves a detailed representation of the stroke pylon as well as the neighboring tower spans [Alipio et al., 2019]. Such simulation, usual for insulator coordination studies, requires small simulation time steps ($< 0.1\,\mu$s) ill-suited for protection studies where multiple faults must be simulated at various location in the grid. We propose instead a simplified approach where the lightning current is connected to the phase conductor and the insulator string is a switch controlled by the voltage at its terminals. The breakdown voltage for the insulator string is set to 2 MV. In case of a fault, we assume that the resistance R_{f} gathers the effect of the pylon, the tower grounding and the arc resistance.

We employ the EMT model of Figure 6.9 to simulate direct lightning strikes that may cause a fault or not, depending on the current amplitude, see Figure 6.8. The obtained waveforms are compared with the one obtained using the ideal switch model, referred to as the "pollution fault". The point of impact is located in line L_{14} at $d_{\text{f}} = 210$ km from station 1. The voltage waveform of the stroke pole (positive)

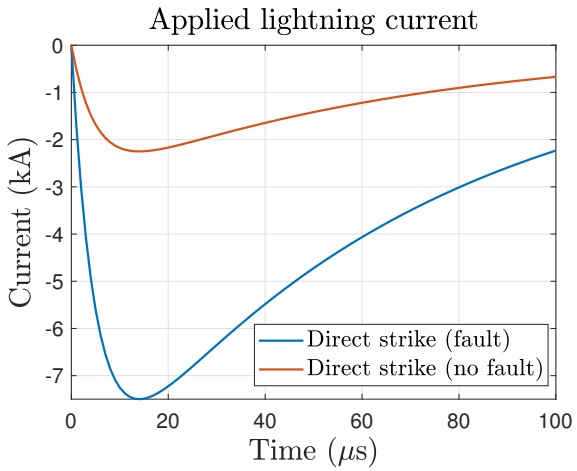


Figure 6.8.: Typical current waveform according to the double exponential model. The current with the highest amplitude causes a fault whereas the other one only induces a disturbance.

and across the insulator at the point of impact can be observed in Figure 6.10. The stroke conductor endures a surge whose amplitude depends on the amplitude of the lightning current injected, see Figure 6.8. Due to the current rise in the conductor, the voltage increases and may exceeds the insulator breakdown voltage, causing a short circuit and the collapse of the voltage (blue curve in Figure 6.10). After the collapse of the voltage, the lightning current mostly flow through the tower (though this may depend on the tower grounding) which limits the variation of the pole voltage. By contrast, if the insulator string never flashes (red curve in Figure 6.10), the lightning current entirely flows through the conductor, causing a large disturbance. The voltage in case of a pollution fault causing a short circuit of the conductor is indicated for the sake of comparison.

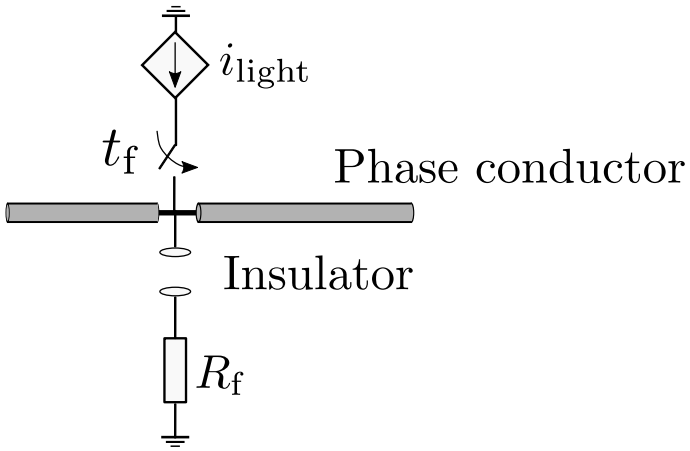


Figure 6.9.: Simplified EMT model of a direct lightning strike as an ideal current source.

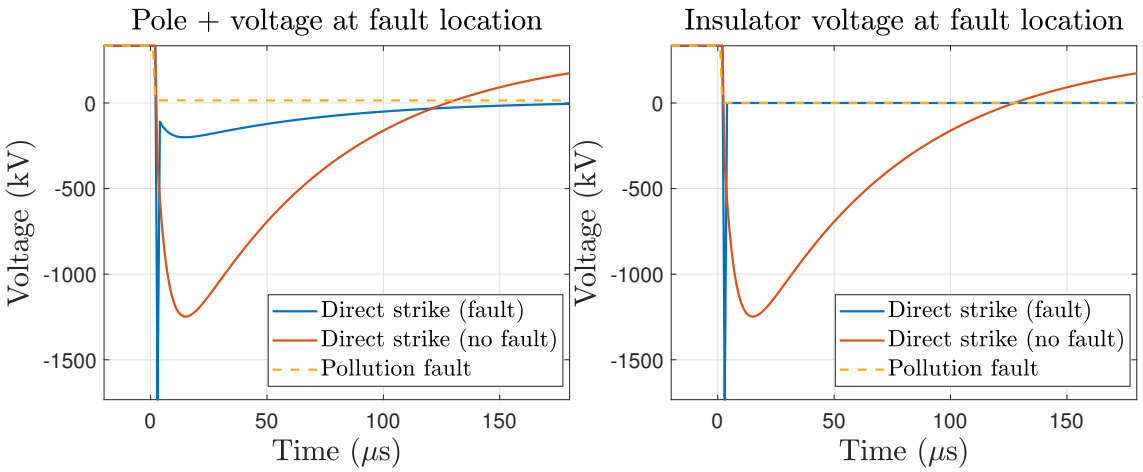


Figure 6.10.: Voltage of the stroke pole (left) and across the insulator string (right) at the point of impact. The waveform for a pollution fault and two direct strikes, causing a fault or a disturbance, are compared.

The voltage observed at the end of the line are plotted in Figure 6.11 on a 2 ms time window. As expected, the TW amplitude due to the lightning disturbance (red) have a greater amplitude than the ones due to a fault (yellow). In the absence of a fault, reflections only occur at the other stations. While the voltage waveforms of a lightning (blue) and a pollution (yellow) induced faults are significantly different for the first two or three TW, they tend to resemble on a longer period. This suggests the ideal representation of faults as a switch is justified, except when considering the very first TW for which the impact of the lightning current itself is important. The

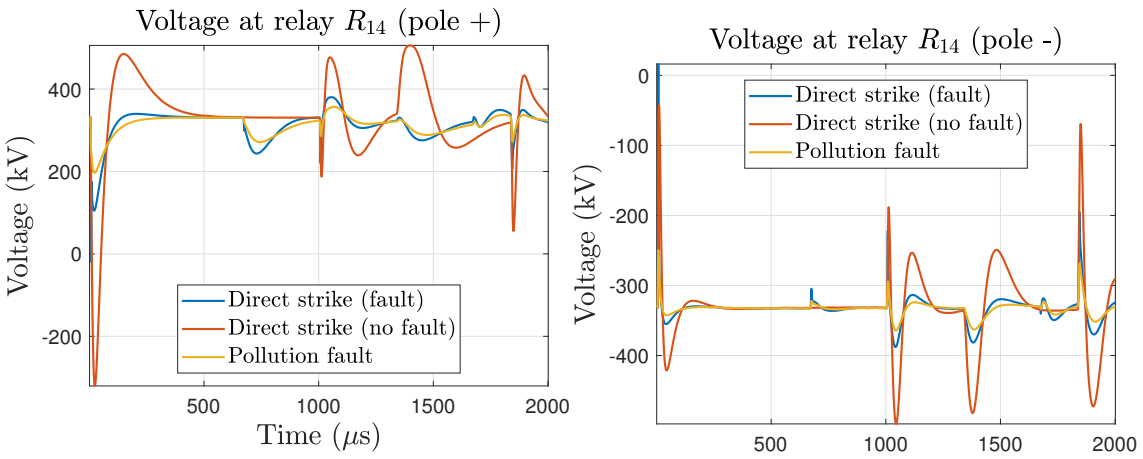


Figure 6.11.: Voltage of the stroke (left) and healthy (right) pole at the extremity of the line. The waveform for a pollution fault and two direct strikes, one causing a fault and the other a disturbance are compared. The waveforms of a lightning (blue) and pollution (yellow) induced fault resemble after the first 2-3 TW.

adaptation of the proposed model-based algorithm to faults caused by a lightning

strike thus requires more studies. It may not be possible to obtain a parametric waveform as in (6.2) that describe the variety of potential lightning strikes. Instead, one can consider the voltage at the fault location as an unknown input that may be estimated, along with the fault distance.

6.3. Complexity and computing times

As presented in Section 4.2.2.3, the execution times of the fault identification algorithm as implemented in Matlab are about 100 times too slow compared to a real time objective. The performances achievable with an optimized C implementation are however difficult to predict using the existing Matlab implementation. In particular, the data processing can be well adapted to parallel computing as, for instance, the different waves - forward and reverse, ground and aerial modes - can be computed separately. Several possibilities may also be considered to improve the performance of the algorithm itself and to limit the computational burden

- In the tuning of the iterative optimization algorithm, see Section 3.3.3, the number of added samples Δn and the number of iterations κ until a new batch of samples is taken into account results in a trade-off between the speed and the accuracy of the algorithm. While the standard tuning $\Delta n = 10$ and $\kappa = 1$ was applied, increasing Δn decreases the number of cost function evaluations required. This however may decrease the accuracy of the estimated fault parameters.

- The sampling time considered for the measurements $f_s = 1\,\text{MHz}$ may be slightly be decreased for instance to 500 kHz as investigated in Section 5.4. This would diminish the number of points on which the the cost function is evaluated.

- As evidenced in Section 4.2.2.3, the computation of more than the first traveling wave greatly increases the execution times. In the few cases where those subsequent waves are used, it may be possible to limit the length of the measurement window so that only the first TW is considered. As suggested by the contour plots of the cost functions, see for instance Figure 4.2, the minimum of the cost function is already well determined using the first 50 μs of measurements.

- The main computations of the algorithm lies in the evaluation of the voltage waveforms obtained through the combined physical and behavioral modeling approach presented in Chapter 2 A pure behavioral approach may lead to more compact formulas, as for instance in proposed the approach in [Zhang et al., 2020]. Such formulas may be tuned off-line based on the existing expressions.

6.4. General conclusion

The protection is one of the technical locks that prevent the emergence of meshed HVDC grids. The fast identification of faults is seen as a key problem and has been the subject of numerous studies. This thesis explores a fast fault identification approach based on the estimation of the fault parameters. The main principle of the approach is reminded in Figure 6.12. A parametric model of the voltage and current transient evolution just after the fault is embedded at each relay. This model depends explicitly on the unknown fault parameters, in particular the fault distance and resistance. When a fault is suspected at a given relay, the fault parameters are estimated using the parametric model and the measurements from the local sensors. Thus the identification algorithm does not use data from distant relays and is thus not limited by the communication delays. Based on the results of the parameter estimation, and in particular using a confidence indicator, a decision logic confirms or refutes that the line protected by the relay is faulty.

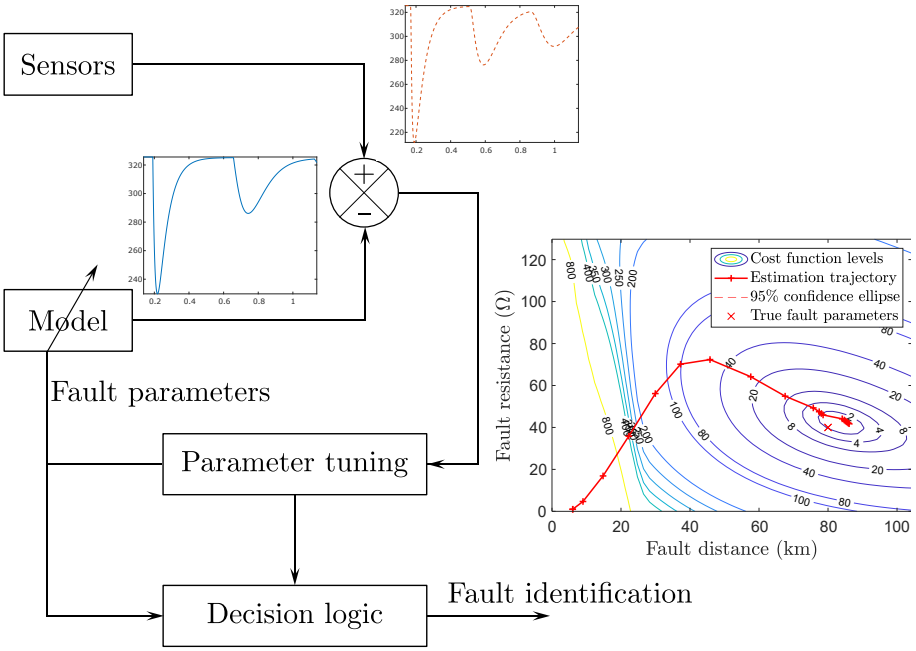


Figure 6.12.: Main principle of the fault identification based on the estimation of the fault parameters.

A general methodology to model DC faults is presented in Chapter 2. Well-known tools such as the traveling waves and modal analysis are first employed to obtain a physical description of the transient behavior of the grid in the first instant after the fault. Explicit formulas can be obtained considering various approximations, and in particular by neglecting the frequency dependent behavior due to the soil resistivity. A behavioral approach is then proposed to account for such effects using low-pass filters. As the distortion and attenuation due to the ground increase with

the fault distance, and the coefficients of the filters representing the ground effects are parameterized with the fault distance. In order to have a model adjustable to any fault distance, an explicit model of the evolution of the ground filters is proposed. A model of the first traveling waves that depend explicitly on the fault parameters is thus obtained. The model is applied to bipolar and monopolar OHL as well as monopolar hybrid lines.

The fault identification algorithm itself is introduced in Chapter 3. The estimation of the fault parameters is performed in the maximum-likelihood sense, leading to a weighted least squares optimization problem. The Levenberg-Marquardt method is then employed to estimate iteratively the fault parameters. An iterative approach is implemented: the estimation starts with very few measurement points and uses more and more data when available. A decision logic is used to confirm or refute that the protected line is faulty. Two tests must be satisfied for the algorithm to identify the fault on the protected line. The validity test compares the value of the estimated fault parameters with minimum and maximum plausible values for the fault distance and resistance. The accuracy test assesses the precision of the estimated parameters by evaluating an approximate confidence region. The area of the obtained ellipse is compared to a pre-determined threshold.

The simulation results of the proposed fault identification approach are presented in Chapter 4. Particular examples and more extensive simulations are both performed. The identification algorithm is able to identify all faults occurring on the protected line except for particularly remote faults for which dependability failures may occur. Security failures are observed in less than 1% of the cases, for faults occurring very close to a station. The length of the required measurement window for identification is particularly small, *i.e.*, always less than 0.8 ms. For most faults, the first traveling wave is sufficient to identify the fault due to the modeling of the ground effects. The accuracy of the estimated fault parameters is acceptable in light of the short amount of data used. A sensitivity analysis is conducted considering an erroneous or non-homogeneous soil resistivity, showing the good robustness of the method.

The system integration of the fault identification algorithm into a full-selective protection strategy is addressed in Chapter 5. A primary protection sequence is proposed to address the dependability and security cases previously observed for faults occurring close to a station. Specifically, the distinction between severe and non-severe faults, based on the estimated fault parameters, allow one to adapt the sequence. Severe faults are likely to induce a very high line current and requires the immediate tripping of the line breakers, whereas for non-severe cases the communication with an other relay can help to correctly identify the fault. Technical specifications regarding the DC breakers as well as the sensors are discussed. An accurate model of the sensors allowed us to study the influence on the performance of the fault identification algorithm of various parameters of the sensors. The impact of the sensor accuracy, bandwidth, and the sampling frequency were thus investigated.

Possible further research topics are presented in Chapter 6. A multiple-hypothesis

approach is proposed on a specific example to tackle the security and dependability failures observed in Chapter 4. The validity of the fault model when considering lightning induced faults is questioned and a framework is proposed to perform an estimation of the input disturbance along with the fault location. More generally, the validation of the proposed model and identification method would greatly benefit from experimental data. Finally, the question of the complexity and the computing times are raised. Despite the relatively good performances obtained compared with existing methods, the gap with real-time is still important. A significant development effort should thus be performed to tackle this issue.

Appendices

A. Test grid and model parameter details

This chapter gather the details on the implemented grid in EMT software EMTP-RV as well as some of the main parameters used in the parametric model. The characteristics of the converter stations are presented in Section A.1. The transmission lines are simulated using the wideband model. The characteristics of the three different considered types of lines - OHL mono-conductor, OHL bipolar and hybrid lines, are detailed in Sections A.2, A.3 and A.4.

A.1. MMC stations

The sub-stations stations are composed of Modular Multilevel Converter (MMC) stations. In the mono-conductor configuration, which corresponds to an asymmetric-monopole, each substation is composed of one MMC. One of the pole of the MMC is connected to the transmission lines through the bus-bar while the other one is directly connected to the ground. In the bipolar case, each sub-station is composed of two MMC connected as a half-bridge. As a metallic return is not considered in the rigid-bipole, the two MMC are grounded through a small resistance ($R_g = 0.5\,\Omega$). More details on the grounding schemes for HVDC grids can be found in [Leterme et al., 2014]. The general topology of an MMC converter is reproduced in Figure A.1, from [Saad et al., 2014].

The MMC sub-stations are simulated in EMT using the model: MMC arm equivalent switching function [Saad et al., 2014] (also known as model 3) with the parameters given in Table A.1. In the bipolar scheme, each converter station is composed of two identical MMC grounded by a small resistor ($0.5\,\Omega$).

The parameters of the RLC equivalent used in the parametric model obtained in (2.31) are indicated in Table A.2. The derivation of the parameters presented in [Leterme and Van Hertem, 2014] are reminded here. The equivalent MMC resistance accounts for all the conduction losses in the IGBT,

$$R_{\mathrm{MMC}} = n_{\mathrm{SM}} \times r_{\mathrm{SM}}.$$

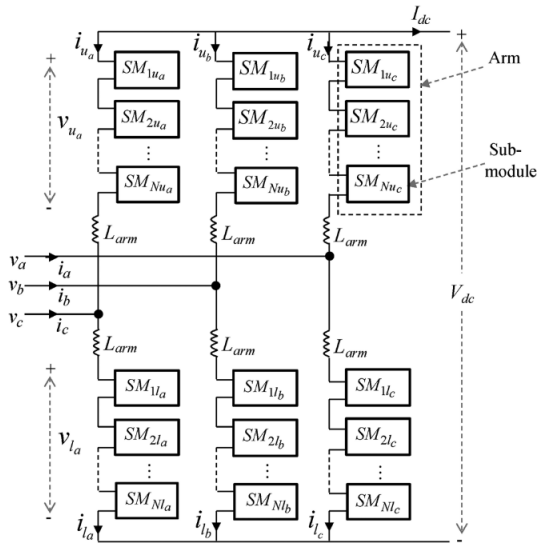


Figure A.1.: Typical MMC topology for a three-phase converter, from [Saad et al., 2014].

Table A.1.: Characteristics of the MMC stations simulated in the EMT software with the model 3: arm switching function.

Rated power S (MVA)	1000
DC rated voltage $V_{\text{base,DC}}$ (kV)	320
Arm inductance $L_{\text{arm,p.u.}}$ (p.u.)	0.15
Transformer reactance (p.u.)	0.18
Capacitor energy in each submodule E_{SM} (kJ/MVA)	40
Conduction losses of each IGBT/diode r_{SM} (Ω)	0.001
Number of sub-modules per arm n_{SM}	400
Grounding impedance for bipole (Ω)	0.5

Table A.2.: Parameters of the equivalent RLC model employed in the reflection and transmission coefficients of the stations.

R_{MMC}	$0.4\,\Omega$
L_{MMC}	8.1 mH
C_{MMC}	$390\,\mu$F

The equivalent inductance is

$$\begin{aligned} L_{\mathrm{MMC}} &= \frac{2}{3} L_{\mathrm{arm}} \\ &= \frac{2}{3} L_{\mathrm{arm,p.u.}} \times \frac{Z_{\mathrm{base,AC}}}{2\pi f_{\mathrm{AC}}} \end{aligned}$$

where the AC frequency is $f_{\mathrm{AC}} = 50\,\mathrm{Hz}$. The equivalent capacitance is

$$\begin{aligned} C_{\mathrm{MMC}} &= 3 C_{\mathrm{arm}} \\ &= 3 \times 2 \frac{E_{\mathrm{arm}}}{V_{\mathrm{base,DC}}^2}, \end{aligned}$$

ans the leg energy E_{arm} is given as

$$E_{\mathrm{arm}} = \frac{1}{6} S \times E_{\mathrm{SM}}.$$

The obtained capacitance value $C_{\mathrm{MMC}} = 390\,\mu\mathrm{F}$ has a negligible influence on the RLC equivalent model for frequencies of interest, typically above. The MMC model can thus be reduced to a RL equivalent.

A.2. Bipolar overhead lines

A bipolar configuration with two poles and without metallic return is considered. Two ground wires are placed above the conductors two ensure the protection against direct lightning strikes. The geometrical configuration of the tower is summarized in Table A.3. The parameters have been deduced from [CIGRE WG B2/B4/C1.17, 2009]. The distributed inductance and capacitance matrices of the fixed frequency model are indicated in Table A.4 for a soil resistivity $\rho_0 = 100\,\Omega\mathrm{m}$.

Table A.3.: Overhead-line characteristics of the rigid bipole configuration for the EMT simulations.

	Poles	Ground wires
DC resistance (Ω/km)	0.024	1.62
Height at tower (m)	37.2	41.7
Height at mid-span (m)	22.2	26.7
Horizontal distance (m)	±4.465	±3.66
Outside diameter (cm)	4.775	0.98

Table A.4.: Distributed inductance and capacitance matrices per unit length at frequency $f_0 = 3.45\,\text{kHz}$ used in the physical part of the model.

Conductor series inductance $[L]_{1,1} = [L]_{2,2}$	$1.45\,\text{mH/km}$
Positive-negative coupling inductance $[L]_{1,2} = [L]_{2,1}$	$0.279\,\text{mH/km}$
Conductor series inductance $[C]_{1,1} = [C]_{2,2}$	$8.29\,\text{nF/km}$
Positive-negative coupling inductance $[C]_{1,2} = [C]_{2,1}$	$-1,23\,\text{nF/km}$

Table A.5.: Overhead-line characteristics of the asymmetric monopole configuration for the EMT simulations.

DC resistance (mΩ/km)	24
Outside diameter (cm)	4.775
Horizontal distance (m)	5
Vertical height at tower (m)	30
Vertical height at mid-span (m)	10
Soil resistivity (Ωm)	100

A.3. Mono-conductor overhead lines

An asymmetric monopole configuration is considered to simulate the case where the transmission lines are composed of a single conductor. The tower geometry is indicated in Table A.5. In this case, no ground wire was included in the EMT model.

The corresponding parameters employed in the model are indicated in Table A.6. In the mono-conductor case, the physical model parameters are obtained from an ideal line where the soil resistivity is neglected, *i.e.*, $\rho = 0\,\Omega\text{m}$. The dependency of the distributed inductance with the frequency is thus removed.

Table A.6.: Distributed parameters per unit length and surge impedance considering $\rho = 0\,\Omega\text{m}$ used in physical part of the mono-conductor model.

Series inductance	$1.45\,\text{mH/km}$
Series capacitance	$7.68\,\text{nF/km}$
Surge impedance	$434\,\Omega$

A.4. Underground cables

For hybrid lines comprising portions of overhead lines and underground cables, the OHL configuration is identical to the one described in Section A.3. The characteristics of the core and screen of the cable are given in Table A.7. The main parameters used in the physical part of the model are detailed in Table A.8 at the frequency $f_0 = 3\,\text{kHz}$.

Table A.7.: Underground cable characteristics for the EMT simulations.

	Core	Screen
Vertical distance (m)		1.33
Outer radius (mm)		63.9
Inside radius (mm)	0	56.9
Outside radius (mm)	32	58.2
Resistivity (nΩm)	17.2	28.3

Table A.8.: Main parameters of the cable used in the physical part of the hybrid line model.

Series inductance	0.129 mH/km
Series capacitance	241 nF/km
Series resistance	83.2 mΩ/km
Surge impedance	23 Ω
Speed	1.8×10^5 km/s

B. Modeling approximations

This section illustrates the approximations proposed in Section 2.1.1 for the physical modeling of multi-conductor transmission lines such as the loss-less and low-loss approaches. The configuration of the grid and the overhead line characteristics are detailed in Appendix A. The variations of the distributed parameters with the frequency is first fully taken into account in Section B.1. The effect of using fixed frequency parameters for the distributed parameters is detailed in Section B.2. In both sections, the obtained quantities are compared with reference values obtained through EMT simulations.

B.1. Evolution of the distributed parameters with the frequency

As seen in Section 2.1.1.2, the transient behavior of a transmission line can be characterized by its surge impedance matrix (in modal domain)

$$[Z_s^c] = [\Gamma]^{-1} [T_V] [Z] [T_V]^T \tag{B.1}$$

and its propagation matrix

$$[\Gamma] = \sqrt{[D_{ZY}]} = \sqrt{[T_V] [Z] [Y] [T_V]^T} \tag{B.2}$$

which can be further decomposed into real and imaginary parts, $[\Gamma] = [\alpha] + j [\beta]$. The propagation along a line of length d is thus modeled as

$$\begin{aligned}
[H] &= \exp(-[\Gamma] d) \\
&= \exp(-[\alpha] d) \exp(-j [\beta] d) \\
&= \underbrace{\exp(-[\alpha] d)}_{=A} \exp\left(-j\omega \frac{[\beta]}{\omega} d\right)
\end{aligned}$$

where $[A] = \exp(-[\alpha] d)$is a real term that represents the attenuation along the line, and $\omega [\beta]^{-1}$ can be interpreted as the propagation speed of the different modes.

The surge impedance $[Z_s^c]$ (in modal domain), the attenuation $[A]$, as well as the propagation speeds can be computed at various frequencies using the full definitions

B.1 and B.2. Low-loss and loss-less approximations have also been introduced in (2.22), (2.23) for $[\Gamma]$ and in (2.24), (2.25) for $[Z_s]$ to be able to get explicit expressions of those quantities. Considering frequency dependent (FD) line parameters, the full model computations and the different approximations are compared with the EMT data in Figures B.1 (attenuation), B.2 (speed), and B.3 (surge impedance). The low-loss approximation shows to be valid as it fits well both the EMT simulations and full analytic model computations. The loss-less approximation show slightly degraded performances for the attenuation of the different modes as well as the surge impedance angle. When using distributed line parameters that vary with the frequency, even the loss-less approximation captures most of the behavior of the line.

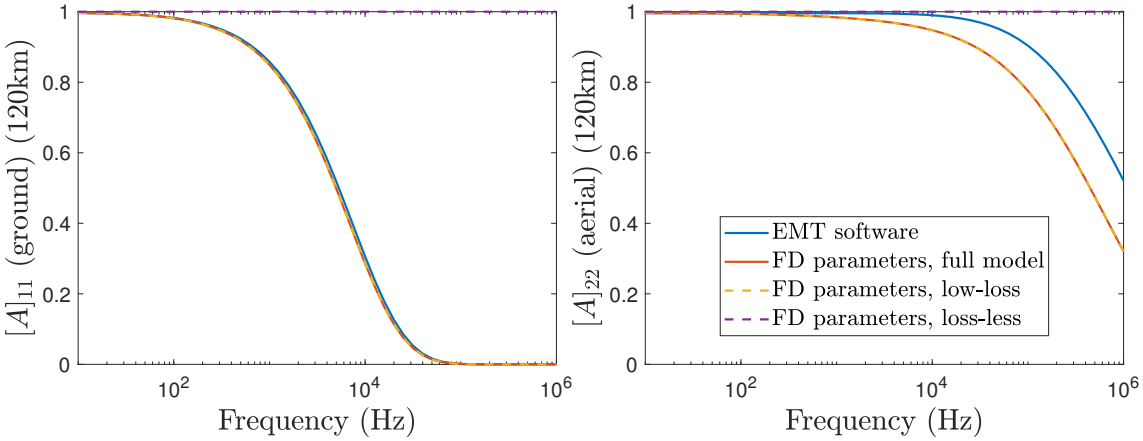


Figure B.1.: Evolution of the attenuation constant of the ground (left) and aerial (right) modes for different approximations with frequency dependent (FD) analytic line parameters compared with the EMT data, for a 120 km long line.

Furthermore, the ground and aerial mode quantities present different behaviors. The ground mode seems more affected by the frequency variations than the aerial mode. In particular, the speed and surge impedance of the aerial mode are almost not affected by the frequency variations. The evolution of the attenuation constant suggests the ground mode is damped for frequencies above 10^3 Hz and the aerial modes for frequencies above 10^5 Hz.

B.2. Distributed parameters at fixed frequency

In order to obtain explicit solutions for the Telegraphers equations, the distributed line parameters have been considered at a fixed frequency. The impact of using constant matrices $[R], [L], [C]$ and $[G]$ on the characteristic quantities of the line is showed in Figures B.4 (attenuation), B.5 (speed), and B.6 (surge impedance). The considered frequency for the line parameters is $f_0 = 1\,\text{kHz}$, within the frequency

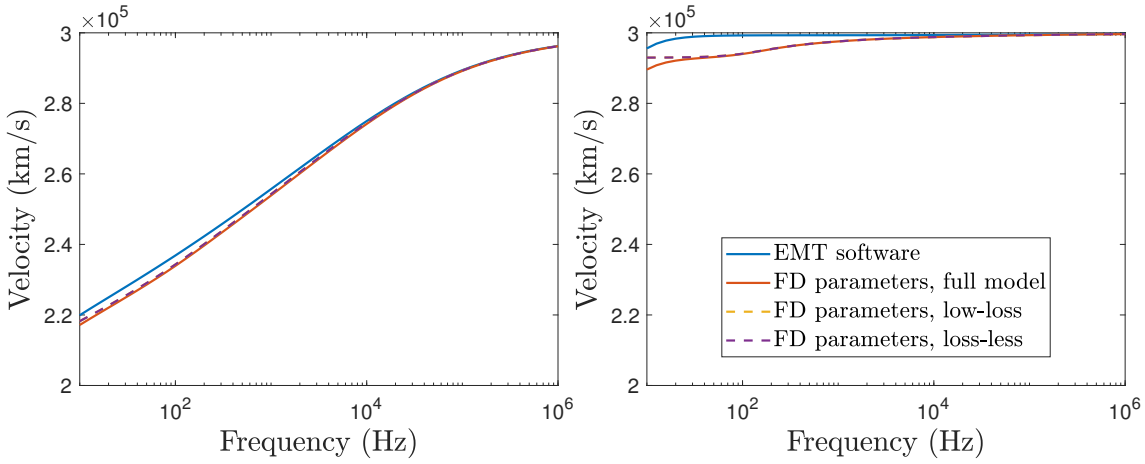


Figure B.2.: Evolution of the propagation speed of the ground (left) and (aerial) modes for different approximations with frequency dependent (FD) line parameters compared with the EMT data.

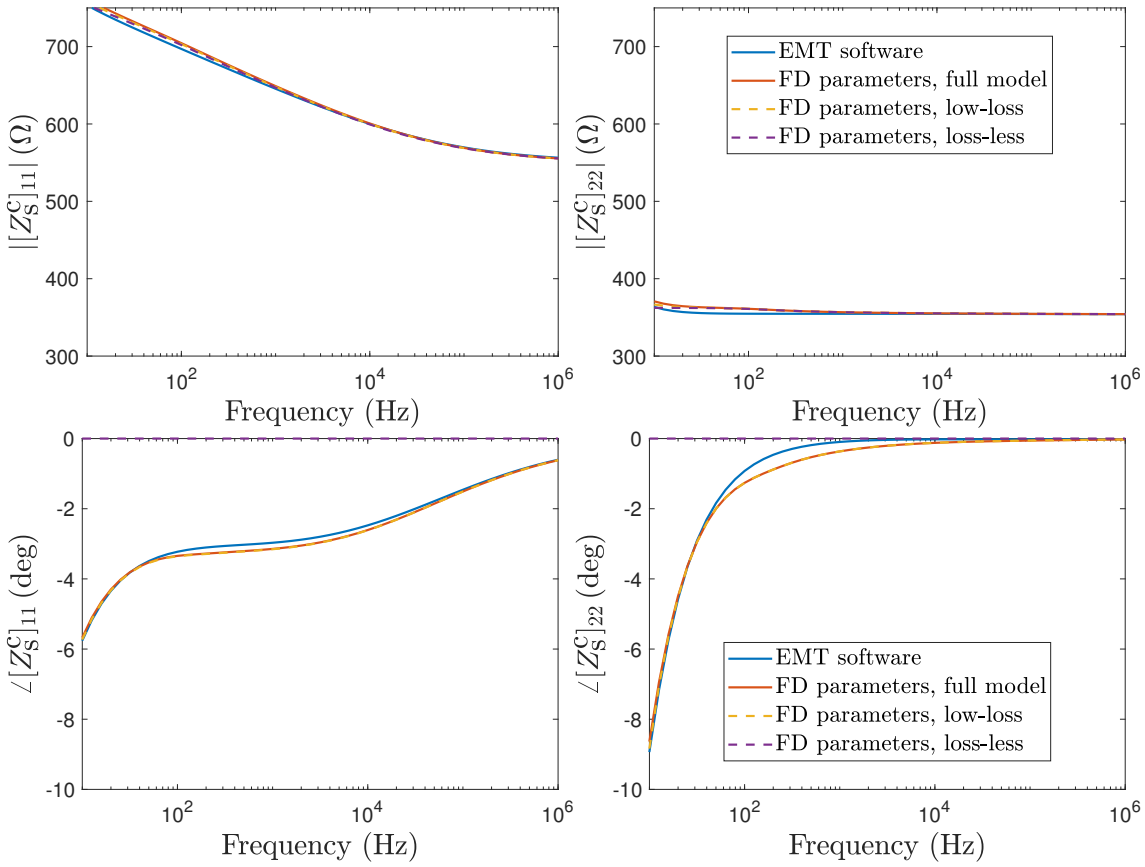


Figure B.3.: Evolution of the modal surge impedance (2.17) module (top) and angle (bottom) of the ground (left) and aerial (right) modes with frequency dependent (FD) line parameters for different approximations compared with the EMT data.

range of interest evidenced in Section B.1. The low-loss and loss-less approximations are compared to the full model computations, all using constant parameters, and the EMT data. It appears that for the ground mode most of the frequency dependent behavior is not captured when using constant line parameters, even without any approximations on the losses. Nevertheless, as the aerial mode is less frequency dependent than the ground mode, the constant parameter model fits better the EMT data for this mode. This justifies that the constant parameters and loss-less approach can be adopted for the aerial mode but is not sufficient to accurately describe the ground mode.

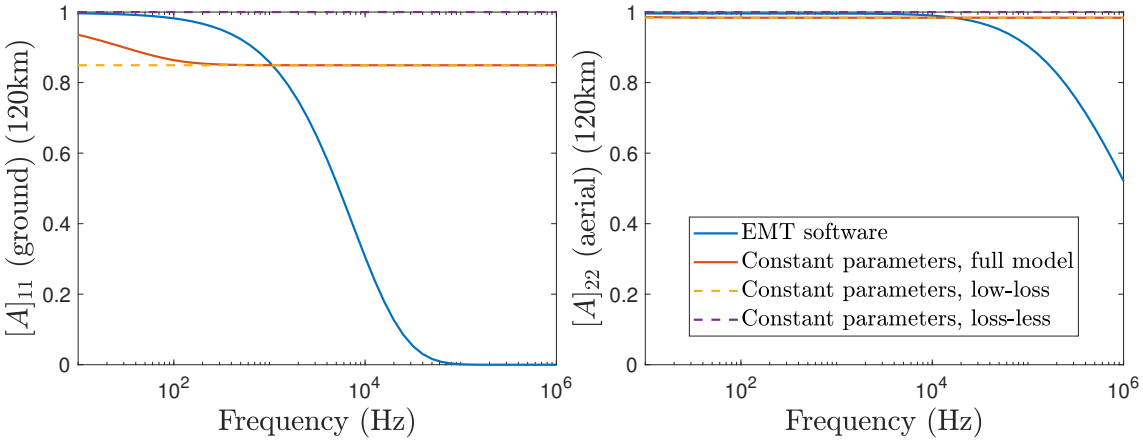


Figure B.4.: Evolution of the attenuation constant of the ground (left) and aerial (right) modes provided by EMT data and using an explicit expressions involving different approximations, for a line of 120 km.

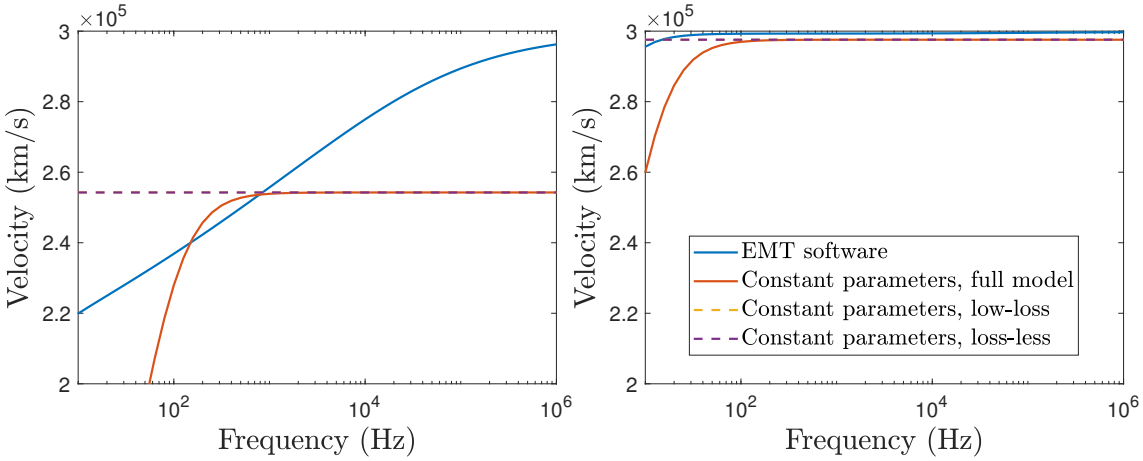


Figure B.5.: Evolution of the propagation speed of the ground (left) and aerial (right) modes provided by EMT data and using an explicit expressions involving different approximations.

The simulations presented in this Section suggest the aerial mode quantities can

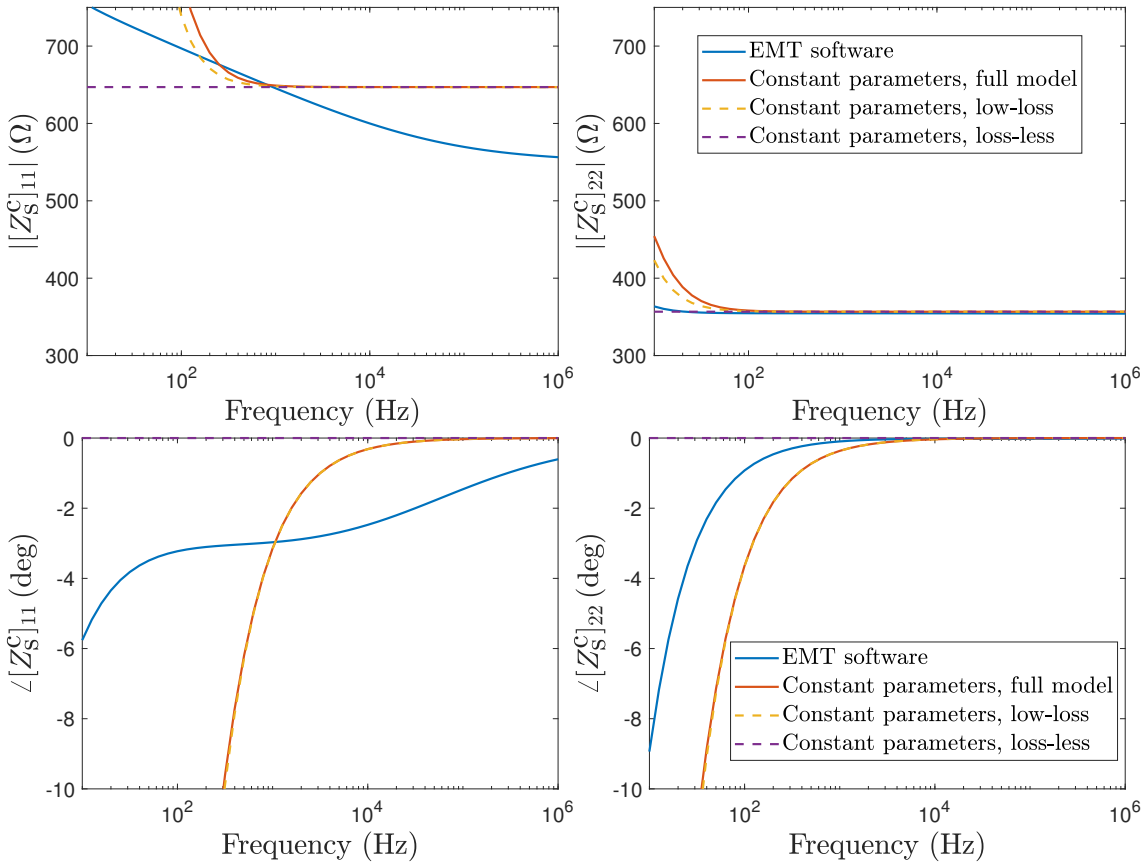


Figure B.6.: Evolution of the surge impedance module (top) and angle (bottom) of the ground (left) and aerial (right) modes provided by EMT data and using an explicit expressions involving different approximations.

reasonably modeled using the distributed parameters at a fixed frequency and low-loss or loss-less approximations. Conversely, the ground mode behavior is more frequency dependent and the constant parameters approach require to be supplemented to obtain an accurate description of the ground mode evolution, as proposed in Section 2.2.

C. Computation of the derivatives

The estimation algorithm presented in Section 3.1.3 reburies the evaluation of the partial derivatives of the cost function with respect to the fault parameters, namely the fault distance and resistance. Though a finite difference approach can be employed, explicit or semi-explicit expressions allow to limit the computational burden of the algorithm. In case the finite difference approach is still required, the increment in the fault distance can be advised based on the temporal resolution of the data, given by the sampling time T_s. A reasonable distance step size is then

$$h_\mathrm{d} = \frac{c}{T_\mathrm{s}} = 0.3\,\mathrm{km}.$$

In this section, closed form expressions for the partial derivatives of the voltage are proposed for the three application cases, namely the nonconductor overhead and hybrid lines and the multi-conductor overhead lines. The partial derivatives of the current can be directly obtained from the voltage using the surge impedance. The partial derivatives may be directly computed from the temporal domain expressions using symbolic computation tools such as the Symbolic Math Toolbox from Matlab. Due to the complexity of the computations however, a detailed by-hand analysis can help reduce the complexity of the obtained formulas.

C.1. Mono-conductor overhead line

For the mono-conductor case, the time-domain voltage expressions for any voltage wave is obtained as the filtering of the physical model by the appropriate ground filter 2.62, *i.e.*,

$$v^{\mathrm{m,g}}(k, \mathbf{p}) = g\left(k, d_\mathrm{f}\right) \otimes v^{\mathrm{m},0}\left(k, \mathbf{p}\right).$$

The derivative of the voltage with respect to the fault resistance thus can be written

$$\frac{\partial v^{\mathrm{m,g}}(k, \mathbf{p})}{\partial R_\mathrm{f}} = g\left(k, d_\mathrm{f}\right) \otimes \frac{\partial v^{\mathrm{m},0}(k, \mathbf{p})}{\partial R_\mathrm{f}}$$

where $\frac{\partial v^{\mathrm{m},0}(k,\mathbf{p})}{\partial R_\mathrm{f}}$ can be obtained using symbolic computation tools.

The derivative with respect to the fault distance involves both the physical and behavioral parts of the model. Considering a first order filter as in (2.63), the combined model can be written more explicitly as

$$v^{\mathrm{m,g}}(k,\mathbf{p}) = b_0\left(d_\mathrm{f}\right) v^{\mathrm{m},0}\left(k,\mathbf{p}\right) + a_1\left(d_\mathrm{f}\right) v^{\mathrm{m,g}}(k-1,\mathbf{p}).$$

Thus the derivative is

$$\begin{aligned}\frac{\partial v^{\mathrm{m,g}}(k,\mathbf{p})}{\partial d_\mathrm{f}} = &\frac{\partial b_0\left(d_\mathrm{f}\right)}{\partial d_\mathrm{f}} v^{\mathrm{m},0}\left(k,\mathbf{p}\right) + b_0\left(d_\mathrm{f}\right)\frac{\partial v^{\mathrm{m},0}\left(k,\mathbf{p}\right)}{\partial d_\mathrm{f}} \\ &+ \frac{\partial a_1\left(d_\mathrm{f}\right)}{\partial d_\mathrm{f}} v^{\mathrm{m},0}\left(k-1,\mathbf{p}\right) + a_1\left(d_\mathrm{f}\right)\frac{\partial v^{\mathrm{m,g}}(k-1,\mathbf{p})}{\partial d_\mathrm{f}}.\end{aligned}$$

As the coefficients b_0 and a_1 are polynomial functions of the fault distance (2.63), their partial derivatives can be directly evaluated. The partial derivative of the physical model can be obtained though symbolic computations.

C.2. Multi-conductor overhead line

The main difficulty in the multi-conductor case arises from the impossibility to commute and gather the different terms within the expressions of the traveling waves. Nevertheless, explicit formulas can be proposed in some cases. We focus on modal quantities only, as phase quantities can be directly obtained from modal ones.

C.2.1. Derivative with respect to the fault resistance

The voltage due to the first traveling wave at station q is (2.35)

$$\mathbf{V}^{\mathrm{m,c}}_{\mathrm{q},1}\left(s,\mathbf{p}\right) = \left[T^{\mathrm{c}}_{\mathrm{q}}\right]\left[H\left(s,d_\mathrm{f}\right)\right]\left[T_V\right]\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]\frac{\exp\left(-t_\mathrm{f}s\right)}{s}\mathbf{v}_{\mathrm{bf}} \tag{C.1}$$

where the reflection matrix at the fault location is

$$\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right] = -\left(2\left[Y_{\mathrm{s},e}\right] + \left[Y_{\mathrm{s,f}}\right]\right)^{-1}\left(\left[Y_{\mathrm{s,f}}\right]\right). \tag{C.2}$$

Assuming now the fault is a pole-to-ground fault affecting the positive pole the expression of (C.2) can be detailed

$$\begin{aligned}\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right] &= \begin{pmatrix} -\frac{\left[Y_\mathrm{s}\right]_{1,1}}{R_\mathrm{f}\left(2\left[Y_\mathrm{s}\right]^2_{1,1}-2\left[Y_\mathrm{s}\right]^2_{1,2}\right)+\left[Y_\mathrm{s}\right]_{1,1}} & 0 \\ \frac{\left[Y_\mathrm{s}\right]_{1,2}}{R_\mathrm{f}\left(2\left[Y_\mathrm{s}\right]^2_{1,1}-2\left[Y_\mathrm{s}\right]^2_{1,2}\right)+\left[Y_\mathrm{s}\right]_{1,1}} & 0 \end{pmatrix} \\ &= \underbrace{\left(R_\mathrm{f}\left(2\left[Y_\mathrm{s}\right]^2_{1,1} - 2\left[Y_\mathrm{s}\right]^2_{1,2}\right) + \left[Y_\mathrm{s}\right]_{1,1}\right)^{-1}}_{=\chi(R_\mathrm{f})}\begin{pmatrix} -\left[Y_\mathrm{s}\right]_{1,1} & 0 \\ \left[Y_\mathrm{s}\right]_{1,2} & 0 \end{pmatrix}\end{aligned}$$

where R_f only occur through a *scalar real* term noted $\chi\left(R_\mathrm{f}\right)$. The derivative of (C.1) is

$$\frac{\partial \mathbf{V}_{\mathrm{q},1}^{\mathrm{m,c}}\left(s,\mathbf{p}\right)}{\partial R_\mathrm{f}} = \left[T_\mathrm{q}^\mathrm{c}\right]\left[H\left(s,d_\mathrm{f}\right)\right]\left[T_V\right]\frac{\partial\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]}{\partial R_\mathrm{f}}\frac{\exp\left(-t_\mathrm{f}s\right)}{s}\mathbf{v}_\mathrm{bf}. \tag{C.3}$$

The derivative of the reflection matrix is

$$\begin{aligned}\frac{\partial\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]}{\partial R_\mathrm{f}} &= \frac{\partial\chi\left(R_\mathrm{f}\right)}{\partial R_\mathrm{f}}\begin{pmatrix}-\left[Y_\mathrm{s}\right]_{1,1} & 0\\ \left[Y_\mathrm{s}\right]_{1,2} & 0\end{pmatrix}\\ &= -\left(2\left[Y_\mathrm{s}\right]_{1,1}^2 - 2\left[Y_\mathrm{s}\right]_{1,2}^2\right)\chi\left(R_\mathrm{f}\right)^2\begin{pmatrix}-\left[Y_\mathrm{s}\right]_{1,1} & 0\\ \left[Y_\mathrm{s}\right]_{1,2} & 0\end{pmatrix}\\ &= -\left(2\left[Y_\mathrm{s}\right]_{1,1}^2 - 2\left[Y_\mathrm{s}\right]_{1,2}^2\right)\chi\left(R_\mathrm{f}\right)\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]\end{aligned}$$

Thus, the voltage derivative (C.3) is

$$\frac{\partial \mathbf{V}_{\mathrm{q},1}^{\mathrm{m,c}}\left(s,\mathbf{p}\right)}{\partial R_\mathrm{f}} = -\left(2\left[Y_\mathrm{s}\right]_{1,1}^2 - 2\left[Y_\mathrm{s}\right]_{1,2}^2\right)\chi\left(R_\mathrm{f}\right)\times \tag{C.4}$$

$$\underbrace{\left[T_\mathrm{q}^\mathrm{c}\right]\left[H\left(s,d_\mathrm{f}\right)\right]\left[T_V\right]\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]\frac{\exp\left(-t_\mathrm{f}s\right)}{s}\mathbf{v}_\mathrm{bf}}_{=\mathbf{V}_{\mathrm{q},1}^{\mathrm{m,c}}(s,\mathbf{p})}. \tag{C.5}$$

$$= \underbrace{-\left(2\left[Y_\mathrm{s}\right]_{1,1}^2 - 2\left[Y_\mathrm{s}\right]_{1,2}^2\right)\chi\left(R_\mathrm{f}\right)}_{\text{scalar real term}}\mathbf{V}_{\mathrm{q},1}^{\mathrm{m,c}}\left(s,\mathbf{p}\right). \tag{C.6}$$

The derivative in temporal domain can then be directly computed, *i.e.*,

$$\frac{\partial \mathbf{v}_{\mathrm{q},1}^{\mathrm{m,c}}\left(t,\mathbf{p}\right)}{\partial R_\mathrm{f}} = 2\left(-\left[Y_\mathrm{s}\right]_{1,1}^2 + \left[Y_\mathrm{s}\right]_{1,2}^2\right)\chi\left(R_\mathrm{f}\right)\mathbf{v}_{\mathrm{q},1}^{\mathrm{m,c}}\left(t,\mathbf{p}\right).$$

Subsequent waves may endure several reflections at the fault location or, in the case of the wave reflected at the remote station, a transmission through the fault. In the latter case it is noteworthy that $\left[T_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right] = I_{n_\mathrm{c}} + \left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]$ so that

$$\frac{\partial\left[T_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]}{\partial R_\mathrm{f}} = \frac{\partial\left[K_{\mathrm{e}\to\mathrm{f}}\left(R_\mathrm{f}\right)\right]}{\partial R_\mathrm{f}}$$

Thus, similar derivations can be applied to any wave and lead to similar results, specifically for a wave that has been reflected and transmitted n times through the fault

$$\frac{\partial \mathbf{v}_{\mathrm{q}}^{\mathrm{m,c}}\left(t,\mathbf{p}\right)}{\partial R_\mathrm{f}} = 2\times n\times\left(-\left[Y_\mathrm{s}\right]_{1,1}^2 + \left[Y_\mathrm{s}\right]_{1,2}^2\right)\chi\left(R_\mathrm{f}\right)\mathbf{v}_{\mathrm{q}}^{\mathrm{m,c}}\left(t,\mathbf{p}\right).$$

C.2.2. Derivative with respect to the fault distance

In the general case, the derivatives with respect to the fault distance must be computed using finite differences. In most of the situations though, only the derivative of the first traveling wave is required, for which an analytic expression has been established using symbolic computation tools.

C.3. Mono-conductor hybrid line

In this section the computation of the partial derivatives of the voltage with respect to the fault distance and fault resistance are established in the case of hybrid lines comprising portions of underground cables and overhead lines. The notations introduced in Section 2.4 are briefly reminded.

A fault is assumed to occur on an edge e_{f} between nodes q_k and q_ℓ within a grid. The two edges connected to the fault node q_{f} are denoted as $e_{\mathrm{f},k}$ and $e_{\mathrm{f},\ell}$ and are of lengths $d_{\mathrm{f},k}$ and $d_{\mathrm{f},\ell}$, respectively. Assuming, without loss of generality, that $k < \ell$, according to the fault distance convention 2.64: $d_{\mathrm{f}} = d_{\mathrm{f},k}$. The computations are detailed for a wave traveling though a path $\pi = (q_{n_1}, \ldots, q_{n_m})$ where $q_{n_i} \in \mathcal{Q}_{\mathrm{f}},\ i = 1, \ldots, m$ and $q_{n_1} = q_{\mathrm{f}}$.

C.3.1. Partial derivative with respect to the fault distance

According to (2.69) and (2.73), the model of the voltage observed at node q_{n_m} resulting from a wave that traveled through the path π can be expressed in the times domain as

$$v_\pi^{\mathrm{m}}(\mathbf{p}, k) = v_{\pi,\mathrm{j}}\left(R_{\mathrm{f}}, \left(k - f_{\mathrm{s}}\tau\left(d_{\mathrm{f}}\right)\right)\right) \otimes h_\pi\left(d_{\mathrm{f}}, k\right),$$

where τ corresponds to the total propagation time through the path π. The delay τ only depends on the fault distance d_{f}.

The derivative with respect to the fault distance d_{f} is then given by (C.7).

$$\begin{aligned}\frac{\partial v_\pi^{\mathrm{m}}(\mathbf{p}, k)}{\partial d_{\mathrm{f}}} &= v_{\pi,\mathrm{j}}\left(R_{\mathrm{f}}, k - f_{\mathrm{s}}\tau\left(d_{\mathrm{f}}\right)\right) \otimes \frac{\partial\left[h_\pi\left(d_{\mathrm{f}}, k\right)\right]}{\partial d_{\mathrm{f}}} \\ &\quad + h_\pi\left(d_{\mathrm{f}}, k\right) \otimes \frac{\partial\left[v_{\pi,\mathrm{j}}\left(R_{\mathrm{f}}, k - f_{\mathrm{s}}\tau\left(d_{\mathrm{f}}\right)\right)\right]}{\partial d_{\mathrm{f}}} \\ &= v_{\pi,\mathrm{j}}\left(R_{\mathrm{f}}, k - f_{\mathrm{s}}\tau\left(d_{\mathrm{f}}\right)\right) \otimes \frac{\partial\left[h_\pi\left(d_{\mathrm{f}}, k\right)\right]}{\partial d_{\mathrm{f}}} \\ &\quad - h_\pi\left(d_{\mathrm{f}}, k\right) \otimes f_{\mathrm{s}} \frac{\partial\tau}{\partial d_{\mathrm{f}}} \frac{\partial\left[v_{\pi,\mathrm{j}}\left(R_{\mathrm{f}}, k - f_{\mathrm{s}}\tau\left(d_{\mathrm{f}}\right)\right)\right]}{\partial\left(k - f_{\mathrm{s}}\tau\left(d_{\mathrm{f}}\right)\right)}.\end{aligned} \tag{C.7}$$

The delay τ due to the propagation along the path π can be expended as

$$\tau\left(d_{\mathrm{f}}\right) = \sum_{\substack{i \\ q_{n_i} \neq q_{\mathrm{f}} \\ q_{n_{i+1}} \neq q_{\mathrm{f}}}} \tau_{e_{n_i,n_{i+1}}} + \sum_{\substack{i \\ e_{n_i,n_{i+1}} = e_{\mathrm{f},k}}} \tau_{e_{n_i,n_{i+1}}} + \sum_{\substack{i \\ e_{n_i,n_{i+1}} = e_{\mathrm{f},\ell}}} \tau_{e_{n_i,n_{i+1}}}, \tag{C.8}$$

where we have isolated the delays due to propagation along the edges $e_{\mathrm{f},k}$ and $e_{\mathrm{f},\ell}$. Introducing $m_{\mathrm{f},k}$ and $m_{\mathrm{f},\ell}$ as the number of times the wave traveled through the two edges connected to the fault $e_{\mathrm{f},k}$ and $e_{\mathrm{f},\ell}$, one gets

$$\tau\left(d_{\mathrm{f}}\right) = \sum_{\substack{i \\ q_{n_i} \neq q_{\mathrm{f}} \\ q_{n_{i+1}} \neq q_{\mathrm{f}}}} \tau_{e_{n_i,n_{i+1}}} + m_{\mathrm{f},k} \frac{d_{\mathrm{f}}}{c_{e_{\mathrm{f}}}} + m_{\mathrm{f},\ell} \frac{d_{e_{\mathrm{f}}} - d_{\mathrm{f}}}{c_{e_{\mathrm{f}}}}$$

The first sum correspond to propagation times along edges not connected to the faulty node. Hence, assuming the propagation speed does not depend on the fault distance

$$\frac{\partial \tau}{\partial d_{\mathrm{f}}} = \frac{m_{\mathrm{f},k} - m_{\mathrm{f},\ell}}{c_{e_{\mathrm{f}}}}. \tag{C.9}$$

Consider now the finite impulse response filter h_π that represents the total distortion along the considered path π. The filter is expressed in the frequency domain in (C.10), where the same decomposition as in (C.8) has been performed.

$$\begin{aligned} H_\pi\left(d_{\mathrm{f}}, \omega\right) &= \prod_{\substack{i=1 \\ q_{n_i} \neq q_{\mathrm{f}} \\ q_{n_{i+1}} \neq q_{\mathrm{f}}}} H_{(q_{n_i}, q_{n_{i+1}})}\left(\omega\right) \times \prod_{\substack{i=1 \\ (q_{n_i}, q_{n_{i+1}}) = e_{\mathrm{f},k}}} H_{(q_{n_i}, q_{n_{i+1}})}\left(d_{\mathrm{f}}, \omega\right) \times \prod_{\substack{i=1 \\ (q_{n_i}, q_{n_{i+1}}) = e_{\mathrm{f},\,\ell}}} H_{(q_{n_i}, q_{n_{i+1}})}\left(d_{\mathrm{f}}, \omega\right) \\ &= \prod_{\substack{i=1 \\ q_{n_i} \neq q_{\mathrm{f}} \\ q_{n_{i+1}}, \neq q_{\mathrm{f}}}} H_{(q_{n_i}, q_{n_{i+1}})}\left(\omega\right) \times H_{e_{\mathrm{f},k}}^{m_{\mathrm{f},k}}\left(d_{\mathrm{f}}, \omega\right) \times H_{e_{\mathrm{f},\ell}}^{m_{\mathrm{f},\ell}}\left(d_{\mathrm{f}}, \omega\right) \end{aligned} \tag{C.10}$$

Taking the derivative with respect to the fault distance, and omitting the dependency in ω, one gets (C.11).

$$\begin{aligned} \frac{\partial H_\pi\left(d_{\mathrm{f}}, \omega\right)}{\partial d_{\mathrm{f}}} &= \prod_{\substack{i=1 \\ q_{n_i} \neq q_{\mathrm{f}} \\ q_{n_{i+1}} \neq q_{\mathrm{f}}}} H_{(q_{n_i}, q_{n_{i+1}})} \left[m_{\mathrm{f},k} H_{e_{\mathrm{f},\ell}}^{m_{\mathrm{f},\ell}} H_{e_{\mathrm{f},k}}^{m_{\mathrm{f},k}-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_{\mathrm{f}}} + m_{\mathrm{f},\ell} H_{e_{\mathrm{f},\ell}}^{m_{\mathrm{f},\ell}-1} H_{e_{\mathrm{f},k}}^{m_{\mathrm{f},k}} \frac{\partial H_{e_{\mathrm{f},\ell}}}{\partial d_{\mathrm{f}}} \right] \\ &= \underbrace{\prod_{\substack{i=1 \\ q_{n_i} \neq q_{\mathrm{f}} \\ q_{n_{i+1}} \neq q_{\mathrm{f}}}} H_{(q_{n_i}, q_{n_{i+1}})} H_{e_{\mathrm{f},k}}^{m_{\mathrm{f},k}} H_{e_{\mathrm{f},k}}^{m_{\mathrm{f},\ell}}}_{H_\pi} \left[m_{e_{\mathrm{f},k}} H_{e_{\mathrm{f},k}}^{-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_{\mathrm{f}}} + m_{e_{\mathrm{f},\ell}} H_{e_{\mathrm{f},\ell}}^{-1} \frac{\partial H_{e_{\mathrm{f},\ell}}}{\partial d_{\mathrm{f}}} \right] \end{aligned} \tag{C.11}$$

Moreover, we assume that a wave that travels successively through $e_{\mathrm{f},k}$ and $e_{\mathrm{f},\ell}$ is prone to the same distortion as a wave that travels through e_{f}, *i.e.*,

$$H_{e_{\mathrm{f},k}}\left(d_{\mathrm{f}}, \omega\right) H_{e_{\mathrm{f},\ell}}\left(d_{\mathrm{f}}, \omega\right) = H_{e_{\mathrm{f}}}\left(\omega\right)$$

Where H_{e_f} does not depend on the fault distance. Hence,

$$\frac{\partial H_{e_{\mathrm{f},\ell}}(d_\mathrm{f},\omega)}{\partial d_\mathrm{f}} = -\frac{\partial H_{e_{\mathrm{f},k}}(d_\mathrm{f},\omega)}{\partial d_\mathrm{f}} \frac{H_{e_\mathrm{f}}(\omega)}{H_{e_{\mathrm{f},k}}(d_\mathrm{f},\omega)^2}$$

The expression of $\frac{\partial H_\pi(d_\mathrm{f},\omega)}{\partial d_\mathrm{f}}$ can thus be further simplified

$$\begin{aligned}
\frac{\partial H_\pi(d_\mathrm{f},\omega)}{\partial d_\mathrm{f}} &= H_\pi \left[m_{e_{\mathrm{f},k}} H_{e_{\mathrm{f},k}}^{-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} + m_{e_{\mathrm{f},\ell}} H_{e_{\mathrm{f},\ell}}^{-1} \frac{\partial H_{e_{\mathrm{f},\ell}}}{\partial d_\mathrm{f}} \right] \\
&= H_\pi \left[m_{e_{\mathrm{f},k}} H_{e_{\mathrm{f},k}}^{-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} - \frac{m_{e_{\mathrm{f},\ell}} H_{e_\mathrm{f}}}{H_{e_{\mathrm{f},\ell}} H_{e_{\mathrm{f},k}}^2} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} \right] \\
&= H_\pi \left[m_{e_{\mathrm{f},k}} H_{e_{\mathrm{f},k}}^{-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} - m_{e_{\mathrm{f},\ell}} H_{e_{\mathrm{f},k}}^{-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} \right] \\
&= H_\pi H_{e_{\mathrm{f},k}}^{-1} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} \left[m_{e_{\mathrm{f},k}} - m_{e_{\mathrm{f},\ell}} \right].
\end{aligned}$$

Taking the inverse Fourier transform

$$\frac{\partial h_\pi(d_\mathrm{f},k)}{\partial d_\mathrm{f}} = \left[m_{e_{\mathrm{f},k}} - m_{e_{\mathrm{f},\ell}} \right] \mathcal{F}^{-1} \left\{ \frac{H_\pi(d_\mathrm{f},\omega)}{H_{e_{\mathrm{f},k}}} \frac{\partial H_{e_{\mathrm{f},k}}}{\partial d_\mathrm{f}} \right\} \tag{C.12}$$

The derivative of the impulse response $H_{e_{\mathrm{f},k}} = \mathcal{F}\left(h_{e_{\mathrm{f},k}}\right)$ can be obtained from (2.71)

$$h_e(k) = f_\mathrm{s} \left(u_{d,\rho}(d_\mathrm{f}, k+1) - u_{d,\rho}(d_\mathrm{f}, k) \right)$$

and the linear interpolation (2.70),

$$u_{d,\rho}(d_\mathrm{f},k) = \frac{u_{d_2,\rho}(k) - u_{d_1,\rho}(k)}{(d_2 - d_1)} (d_\mathrm{f} - d_1) + u_{d_1,\rho}(k)$$

where $d_1 < d_\mathrm{f} < d_2$. Hence,

$$\frac{\partial u_{d,\rho}(k)}{\partial d_\mathrm{f}} = \frac{u_{d_2,\rho}(k) - u_{d_1,\rho}(k)}{(d_2 - d_1)}.$$

The last derivative to compute in (C.7) is

$$\frac{\partial v_{\pi,\mathrm{j}}(R_\mathrm{f}, k - f_\mathrm{s}\tau(d_\mathrm{f}))}{\partial (k - f_\mathrm{s}\tau(d_\mathrm{f}))} = \left. \frac{\partial v_{\pi,\mathrm{j}}(R_\mathrm{f}, k')}{\partial k'} \right|_{k' = k - f_\mathrm{s}\tau},$$

which may be approximated by the finite difference

$$\frac{\partial v_{\pi,\mathrm{j}}\left(R_\mathrm{f},k'\right)}{\partial k'} \simeq \left(v_{\pi,\mathrm{j}}\left(R_\mathrm{f},k'+1\right)-v_{\pi,\mathrm{j}}\left(R_\mathrm{f},k'\right)\right). \tag{C.13}$$

The final voltage derivative (C.14) is obtained combining (C.9), (C.12) and (C.13). The evaluation of (C.9) is simple as it only requires counting how many times the wave traveled through the edges connected to the fault. Similarly, (C.13) involves a discrete differentiation. The computation of (C.12) is more demanding but still relatively efficient as h_π is already available from the computations of the voltage. This approach thus leads to a direct evaluation of the derivative with respect to the fault distance more effective than a finite difference approach.

$$\frac{\partial v_\pi^\mathrm{m}\left(\mathbf{p},k\right)}{\partial d_\mathrm{f}} = v_{\pi,\mathrm{j}}\left(R_\mathrm{f},k-f_\mathrm{s}\tau\left(d_\mathrm{f}\right)\right) \otimes \frac{\partial\left[h_\pi\left(d_\mathrm{f},k\right)\right]}{\partial d_\mathrm{f}} \tag{C.14}$$

$$- f_\mathrm{s}\frac{\partial\tau}{\partial d_\mathrm{f}} h_\pi\left(d_\mathrm{f},k\right) \otimes \frac{\partial\left[v_{\pi,\mathrm{j}}\left(R_\mathrm{f},k-f_\mathrm{s}\tau\left(d_\mathrm{f}\right)\right)\right]}{\partial\left(k-f_\mathrm{s}\tau\left(d_\mathrm{f}\right)\right)} \tag{C.15}$$

C.3.2. Partial derivative with respect to the fault resistance

The parametric model (2.73) depends on the fault resistance R_f only through the interactions at the fault location, appearing in the reflection and transmission coefficients. Furthermore, in the loss–less transmission line model, the surge impedance is a real number. Since the fault impedance is considered as purely resistive, the reflection and transmission coefficients at the fault location are thus also real numbers.

The part of the model that computes the reflection and transmission at the different interfaces $V_{\pi,\mathrm{j}}\left(R_\mathrm{f},s\right)$ is

$$V_{\pi,\mathrm{j}}(R_\mathrm{f},s) = \prod_{i=1}^{n} J_{e_{n_{i-1},n_i}\to q_{n_i}}(s,R_\mathrm{f})\frac{\exp(-t_\mathrm{f}s)}{s}V_\mathrm{bf}$$

where the first term accounts for the initial surge at the fault location $J_{e_{n_0,n_1}\to q_{n_1}} = K_{q_{n_1},,q_{n_2},\leftrightarrow q_\mathrm{f}}$. One can separate the reflections and transmissions at the fault location, which involves the fault resistance, and the interactions at the other junctions

$$V_{\pi,\mathrm{j}}(R_\mathrm{f},s) = e^{-st_\mathrm{f}}V_\mathrm{bf}K_{q_{n_1},,q_{n_2},\leftrightarrow q_\mathrm{f}}\left(R_\mathrm{f}\right)\prod_{\substack{i\\ q^i\neq q_\mathrm{f}}} J_{e_{n_{i-1},n_i}\to q_{n_i}}$$

$$\times \prod_{\substack{i=1\\ q_{n_i}=q_\mathrm{f}\\ q_{n_{i+1}}=q_{n_{i-1}}}} \underbrace{J_{e_{n_{i-1},n_i}\to q_{n_i}}}_{=K_{q_{n_1},,q_{n_2},\leftrightarrow q_\mathrm{f}}} \times \prod_{\substack{i=1\\ q_{n_i}=q_\mathrm{f}\\ q_{n_{i+1}}\neq q_{n_{i-1}}}} \underbrace{J_{e_{n_{i-1},n_i}\to q_{n_i}}}_{=T_{q_{n_1},,q_{n_2},\to q_\mathrm{f}}}.$$

The reflection and transmission coefficients at the fault location are noted K_f and T_f to lighten notations. Consider the numbers $m_{\mathrm{K_f}}$ and $m_{\mathrm{T_f}}$ of the reflections and transmissions at the fault location. Hence,

$$V_{\pi,\mathrm{j}}(R_\mathrm{f},s) = e^{-st_\mathrm{f}} V_\mathrm{bf} K_\mathrm{f}^{1+m_{K_\mathrm{f}}}(R_\mathrm{f})\, T_\mathrm{f}^{m_{T_\mathrm{f}}}(R_\mathrm{f}) \prod_{\substack{i \\ q_{n_i}\neq q_\mathrm{f}}} J_{e_{n_{i-1},n_i}\to q_{n_i}}.$$

The derivative with respect to the fault resistance is detailed in (C.16)

$$\begin{aligned}\frac{\partial V_{\pi,\mathrm{j}}(R_\mathrm{f},s)}{\partial R_\mathrm{f}} &= \exp(-st_\mathrm{f}) V_\mathrm{bf} \prod_{\substack{i \\ q^i\neq q_\mathrm{f}}} J_{e_{n_{i-1},n_i}\to q_{n_i}} \\ &\times \left[(1+m_{K_\mathrm{f}})\, K_\mathrm{f}^{m_{K_\mathrm{f}}} \frac{\partial K_{e_{\mathrm{f}\to f}}}{\partial R_\mathrm{f}} \times T_\mathrm{f}^{m_{T_\mathrm{f}}} + m_{T_\mathrm{f}} T_\mathrm{f}^{m_{K_\mathrm{f}}-1} \frac{\partial T_{e_{\mathrm{f}\to f}}}{\partial R_\mathrm{f}} \times K_\mathrm{f}^{1+m_{T_\mathrm{f}}}\right] \\ &= \underbrace{\exp(-st_\mathrm{f}) V_\mathrm{bf} * \prod_{\substack{i \\ q^i\neq q_\mathrm{f}}} J_{e_{n_{i-1},n_i}\to q_{n_i}} \times K_\mathrm{f}^{1+m_{K_\mathrm{f}}} \times T_\mathrm{f}^{m_{T_\mathrm{f}}}}_{=V_{\pi,\mathrm{j}}(R_\mathrm{f},s)} \\ &\times \left[(1+m_{K_\mathrm{f}})\, K_\mathrm{f}^{-1} \frac{\partial K_\mathrm{f}}{\partial R_\mathrm{f}} + m_{T_\mathrm{f}} T_\mathrm{f}^{-1} \frac{\partial T_\mathrm{f}}{\partial R_\mathrm{f}}\right].\end{aligned} \tag{C.16}$$

Hence, one gets

$$\frac{\partial V_{\pi,\mathrm{j}}(R_\mathrm{f},s)}{\partial R_\mathrm{f}} = V_{\pi,\mathrm{j}}(R_\mathrm{f},s) \left[(1+m_{K_\mathrm{f}})\, K_\mathrm{f}^{-1} \frac{\partial K_\mathrm{f}}{\partial R_\mathrm{f}} + m_{T_\mathrm{f}} T_\mathrm{f}^{-1} \frac{\partial T_\mathrm{f}}{\partial R_\mathrm{f}}\right]$$

The reflection and transmission at the fault location are

$$\begin{aligned}K_\mathrm{f} &= \frac{R_\mathrm{f}}{0.5 Z_{\mathrm{s},e_\mathrm{f}} + R_\mathrm{f}} \\ T_\mathrm{f} &= \frac{-Z_{\mathrm{s},e_\mathrm{f}}}{Z_{\mathrm{s},e_\mathrm{f}} + 2R_\mathrm{f}}\end{aligned}$$

whose derivatives are

$$\frac{\partial T_\mathrm{f}}{\partial R_\mathrm{f}} = \frac{0.5 Z_{e_\mathrm{f}}}{(0.5 Z_{e_\mathrm{f}} + R_\mathrm{f})^2} = \frac{K_\mathrm{f} T_\mathrm{f}}{R_\mathrm{f}}$$

$$\frac{\partial K_\mathrm{f}}{\partial R_\mathrm{f}} = \frac{\partial (T_\mathrm{f} - 1)}{\partial R_\mathrm{f}} = \frac{\partial T_\mathrm{f}}{\partial R_\mathrm{f}}$$

Hence the previous expressions can be even further simplified

$$\frac{\partial V_{\pi,\mathrm{j}}(R_\mathrm{f},s)}{\partial R_\mathrm{f}} = V_{\pi,\mathrm{j}}(R_\mathrm{f},s) \frac{[(1+m_{K_\mathrm{f}})T_\mathrm{f} + m_{T_\mathrm{f}} K_\mathrm{f}]}{R_\mathrm{f}}.$$

The voltage derivative with respect to the fault resistance hence amounts to a multiplication by a real coefficient of the voltage expression. Thus the same coefficient can be applied in time domain, *i.e.*,

$$\frac{\partial v_{\pi}^{\mathrm{m}}\left(\mathbf{p}, t_k\right)}{\partial R_{\mathrm{f}}} = v_{\pi}^{\mathrm{m}}\left(\mathbf{p}, t_k\right) \frac{\left[(1 + m_{K_{\mathrm{f}}})T_{\mathrm{f}} + m_{T_{\mathrm{f}}}K_{\mathrm{f}}\right]}{R_{\mathrm{f}}}. \tag{C.17}$$

D. Influence of the initial point

The Levenberg-Marquardt algorithm introduced in Section 3.1.3 to compute the minimum of the cost function requires to provide an initial guess for the value of the fault parameters $\mathbf{p_0} = (R_{f,0}, d_{f,0})$. The impact of different initial guess on the performances of the fault identification algorithm is further investigated in this section. For different fault cases stemmed from the extensive simulations performed in the bipolar configuration, see Section 4.2.2, 50 identification algorithms are launched with a random initial fault resistance and distance. The impact of the initial fault parameters on the ability of the algorithm to identify internal faults, or dependability, as well as on the size of the required measurement window is evidenced.

Consider first two close faults occurring on line L_{24} of length $d_{24} = 100\,\text{km}$ at $d_f^* = 1\,\text{km}$ from relay R_{24} with $R_f^* = 0\,\Omega$ and $R_f^* = 160\,\Omega$. The result of the algorithm for the different initial points are represented for the two fault cases in Figure D.1. The initial points that lead to a correct identification are indicated in green and those for which a dependability failure occur are in plotted in orange. The size of the circles is proportional to the length of measurement window required for the fault identification. Significantly, for the fault occurring at $d_f^* = 1\,\text{km}$ with $R_f^* = 0\,\Omega$, the algorithms that starts with a small fault distance (e.g. $d_{f,0} \leq 15\,\text{km}$) correctly identify the fault whereas larger initial fault distances may lead to dependability failures. A small initial distance also leads to shorter observation windows. By contrast, for the fault occurring at $d_f^* = 1\,\text{km}$ with $R_f^* = 160\,\Omega$ all the initial points but one lead to a correct identification.

Two average faults occurring on line L_{14} are now considered. The result of the identification algorithm at relay R_{12} are indicated on Figure D.2 for a fault occurring at $d_f^* = 70\,\text{km}$ with $R_f^* = 160\,\Omega$ (left) and for a fault at $d_f^* = 210\,\text{km}$ with $R_f^* = 80\,\Omega$ (right). In both cases, the identification is successful for all initial points and the length of the observation window does not variate significantly depending on the initial point.

Finally, the particular case of faults occurring at the remote end of the protected line is investigated. The results of the identification algorithm at relay R_{24} are indicated on Figure D.3 for a fault occurring at $d_f^* = 99\,\text{km}$ with $R_f^* = 0\,\Omega$ (left) and for a fault at $d_f^* = 99\,\text{km}$ with $R_f^* = 160\,\Omega$ (right). As evidenced in Section 4.2.2 such cases are particularly challenging for the fault identification process. For the high impedance fault, see Figure D.3 (right) , none of the different initial point leads to the correct identification of the fault. For the solid fault case (left), initial points with a large $R_{f,0}$ and small $d_{f,0}$ seem to increase the risk of dependability failure

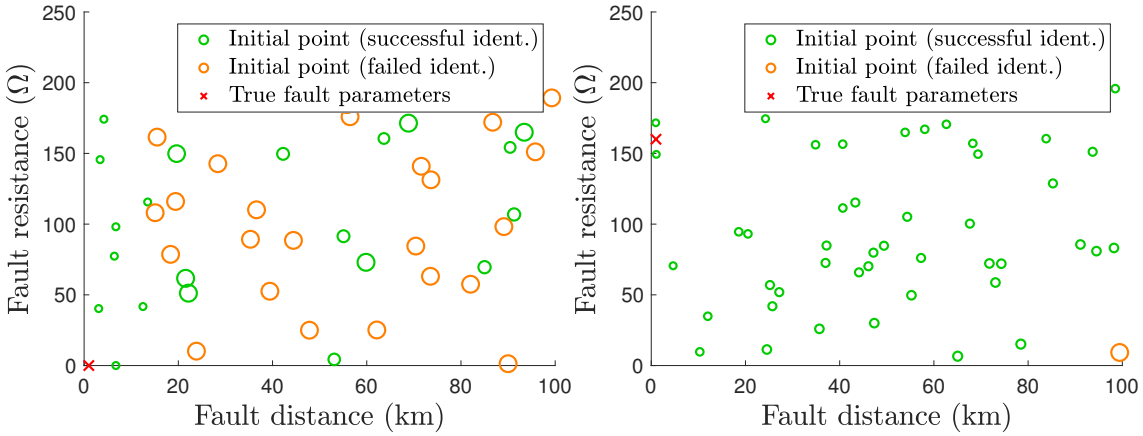


Figure D.1.: Result of the fault identification algorithm for different initial points for low (left) and high (right) impedance faults that occur at $d_{\mathrm{f}}^{*} = 1\,\mathrm{km}$ from relay R_{24}.

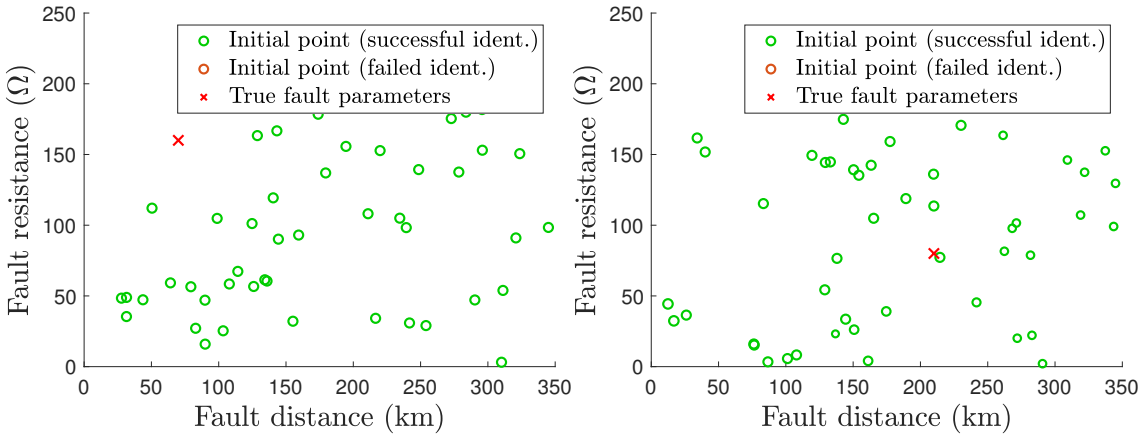


Figure D.2.: Result of the fault identification algorithm for two average faults on line L_{14} of length $d_{24} = 350\,\mathrm{km}$.

whereas the initial point with a small $R_{\mathrm{f},0}$ and $d_{\mathrm{f},0}$ requires a smaller observation window than the other initial points.

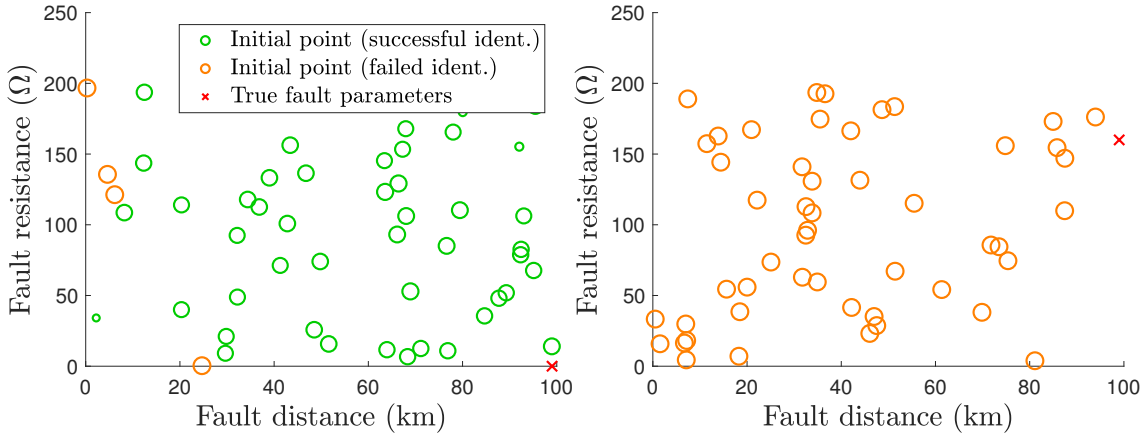


Figure D.3.: Result of the fault identification algorithm for two average faults on line L_{24} of length $d_{24} = 100\,\mathrm{km}$.

The examination of different cases representative of close, average, and remote, as well as low and high impedance faults showed that for the majority of the fault cases, the result of the identification algorithm is insensitive to the choice of the initial point. Nevertheless, for the very close solid faults, which lead to the most severe consquences, an initialization in the first 10-15 km is required to avoid dependability failures. The modeling of very close faults involve at least a second traveling wave due to the fast reflection from the fault location. It may be harder for the model to include this second TW if the initial fault distance is too large. Finally, a small initial resistance may increase the chance of correct identification for faults occurring at the remote end of the line. Those observations justify the adopted choice of initial point $(R_{\mathrm{f},0},\, d_{\mathrm{f},0}) = (5\,\Omega, 5\,\mathrm{km})$ presented in Section 3.3.3. This choice has been further validated with the extensive simulations performed in Section 4.2.2.

Bibliography

[Abedrabbo et al., 2017] Abedrabbo, M., Wang, M., Tielens, P., Dejene, F. Z., Leterme, W., Beerten, J., and Van Hertem, D. (2017). Impact of DC grid contingencies on AC system stability. In *Proc. 13th IET International Conference on AC and DC Power Transmission (ACDC)*, pages 1–7, Manchester. IET.

[Abu-Elanien et al., 2016] Abu-Elanien, A. E., Elserougi, A. A., Abdel-Khalik, A. S., Massoud, A. M., and Ahmed, S. (2016). A differential protection technique for multi-terminal HVDC. *Electric Power Systems Research*, 130:78–88.

[Afa and Anaele, 2010] Afa, J. and Anaele, C. (2010). Seasonal Variation of Soil Resistivity and Soil Temperature in Bayelsa State. *American Journal of Engineering and Applied Sciences*, 3(4):704–709.

[Ali Al Hage et al., 2016] Ali Al Hage, A., Piepenbreier, B., Buchstaller, D., and Engel, M. (2016). Multiple model based fault localization for HVDC transmission systems: Robustness and real-world performance. In *Proc. IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pages 1–5, Dallas.

[Alipio et al., 2019] Alipio, R., De Conti, A., Miranda, A., and Correia De Barros, M. T. (2019). Lightning Overvoltages Including Frequency-Dependent Soil Parameters in the Transmission Line Model. In *Proc. International Conference on Power Systems Transients (IPST)*, Perpignan.

[Alstom, 2011] Alstom (2011). *Network Protection & Automation Guide*. Alstom Grid.

[Ashouri et al., 2018] Ashouri, M., Bak, C. L., and Faria Da Silva, F. (2018). A review of the protection algorithms for multi-terminal VCD-HVDC grids. In *Proc. IEEE International Conference on Industrial Technology (ICIT)*, pages 1673–1678, Lyon.

[Auran, 2017] Auran, G. (2017). *Full selective protection strategy for multi-terminal cable HVDC grids based on HB-MMC converters*. Phd thesis, Université Grenoble Alpes.

[Auran et al., 2017] Auran, G., Descloux, J., Nguefeu, S., and Raison, B. (2017). Non-unit full selective protection algorithm for MTDC grids. In *Proc. IEEE Power & Energy Society General Meeting (PES)*, pages 1–5, Chicago.

[Azad et al., 2015] Azad, S. P., Leterme, W., and Van Hertem, D. (2015). A DC grid primary protection algorithm based on current measurements. In *Proc. 17th*

European Conference on Power Electronics and Applications, EPE-ECCE Europe, pages 1–10, Geneva.

[Chang et al., 2017] Chang, B., Cwikowski, O., Barnes, M., Shuttleworth, R., Beddard, A., and Coventry, P. (2017). Review of different fault detection methods and their impact on pre-emptive VSC-HVDC dc protection performance. *High Voltage*, 2(4):211–219.

[Chatrefou et al., 2000] Chatrefou, D., Pristchepa, M., Uhde, D., and Alstom, T. (2000). Application of Optical Sensors for Measurement of High Frequency Overvoltages in Power Transformers. In *Proc. IEEE Power Engineering Society Winter Meeting*, Singapore.

[CIGRE JWG A3/B4.34, 2017] CIGRE JWG A3/B4.34 (2017). Technical requirements and specifications of state-of-the-art HVDC switching equipment. Technical report, CIGRE.

[CIGRE WG B2/B4/C1.17, 2009] CIGRE WG B2/B4/C1.17 (2009). Impacts of HVDC lines in the economics of HVDC projects. Technical report, CIGRE.

[CIGRE WG01 SC33, 1991] CIGRE WG01 SC33 (1991). Guide to procedures for estimating the lightning performance of transmission lines. Technical Report October.

[Coelho et al., 2015] Coelho, V. L., Piantini, A., Almaguer, H. A., Coelho, R. A., Boaventura, W. D. C., and Paulino, J. O. S. (2015). The influence of seasonal soil moisture on the behavior of soil resistivity and power distribution grounding systems. *Electric Power Systems Research*, 118:76–82.

[Davidson et al., 2015] Davidson, C. C., Whitehouse, R. S., Barker, C. D., Dupraz, J.-P., and Grieshaber, W. (2015). A new ultra-fast HVDC Circuit breaker for meshed DC networks. In *Proc. 11th IET International Conference on AC and DC Power Transmission*, Birmingham.

[De Almeida et al., 2009] De Almeida, S. A., Pestana, R., and Barbosa, F. P. (2009). The main causes of incidents in the portuguese transmission system - Their characterization and how they can be used for risk assessment. In *Proc. 6th International Conference on the European Energy Market, EEM*, Leuven.

[De Andrade Suárez and Sorrentino, 2010] De Andrade Suárez, J. and Sorrentino, E. (2010). Typical expected values of the fault resistance in power systems. In *IEEE/PES Transmission and Distribution Conference and Exposition: Latin America (T&D-LA)*, pages 602–609, Sao Paulo.

[De Kerf et al., 2011] De Kerf, K., Srivastava, K., Reza, M., Bekaert, D., Cole, S., Van Hertem, D., and Belmans, R. (2011). Wavelet-based protection strategy for DC faults in multi-terminal VSC HVDC systems. *IET Generation, Transmission & Distribution*.

[ENTSO-E, 2020] ENTSO-E (2020). TYNDP 2020 Scenario Report. Technical report.

[Fairley, 2013] Fairley, P. (2013). Germany jump-starts the supergrid. *IEEE Spectrum*, 50(5):36–41.

[Farshad and Sadeh, 2013] Farshad, M. and Sadeh, J. (2013). A novel fault-location method for HVDC transmission lines based on similarity measure of voltage signals. *IEEE Transactions on Power Delivery*, 28(4):2483–2490.

[Fawcett, 2006] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.

[Franklin, 2009] Franklin, G. A. (2009). Using modal analysis to estimate received signal levels for a power-line carrier channel on a 500-kV transmission line. *IEEE Transactions on Power Delivery*, 24(4):2446–2454.

[Freitas et al., 2019] Freitas, G. D., Bertinato, A., Raison, B., Niel, E., and Despouys, O. (2019). FMEA of a non-selective fault-clearing strategy for HVDC grids. In *Proc. 15th IET International Conference on AC and DC Power Transmission*, Conventry.

[Guo et al., 2015] Guo, J., Wang, G., Liang, Y., and Zeng, D. (2015). Global-Sensitivity-Based Theoretical Analysis and Fast Prediction of Traveling Waves with Respect to Fault Resistance on HVDC Transmission Lines. *IEEE Transactions on Power Delivery*, 30(4):2007–2016.

[Guzman-Casillas et al., 2018] Guzman-Casillas, A., Kasztenny, B., Tong, Y., and Mynam, M. V. (2018). Accurate single-end fault locating using traveling-wave reflection information. In *Proc. 14th International Conference on Developments in Power System Protection (DPSP)*, pages 1–6, Belfast.

[Henry et al., 2014] Henry, S., Panciatici, P., and Parisot, A. (2014). Going green: Transmission grids as enablers of the transition to a low-carbon European economy. *IEEE Power and Energy Magazine*, 12(2):26–35.

[IEC 61869-15, 2018] IEC 61869-15 (2018). Instrument transformers Part 15: Additional requirements for voltage transformers for DC applications. Technical report, IEC 61869-6:2016 Edition 1.0.

[IEEE Substations Committee, 2012] IEEE Substations Committee (2012). IEEE Guide for Measuring Earth Resistivity, Ground Impedance, and Earth Surface Potentials of a Grounding System. Technical report.

[Isermann, 1984] Isermann, R. (1984). Process fault detection based on modeling and estimation methods-A survey. *Automatica*, 20(4):387–404.

[Johannesson and Norrga, 2018] Johannesson, N. and Norrga, S. (2018). Estimation of travelling wave arrival time in longitudinal differential protections for multi-terminal HVDC systems. In *Proc. 14th International Conference on Developments in Power System Protection (DPSP)*, pages 1007–1011, Belfast.

[Johannesson and Norrga, 2019] Johannesson, N. and Norrga, S. (2019). Contribution of travelling wave propagation time to the speed of optical link protections

in multi-terminal high-voltage DC systems. *IET Generation, Transmission and Distribution*, 13(14):3078–3085.

[Johannesson et al., 2016] Johannesson, N., Norrga, S., and Wikström, C. (2016). Selective Wave-Front Based Protection Algorithm for MTDC Systems. In *Proc. 13th International Conference on Development in Power System Protection (DPSP)*, pages 1–6, Edinburgh.

[Jovcic et al., 2019] Jovcic, D., Tang, G., and Pang, H. (2019). Adopting Circuit Breakers for High-Voltage dc Networks: Appropriating the Vast Advantages of dc Transmission Grids. *IEEE Power and Energy Magazine*, 17(3):82–93.

[Kay, 1993] Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall PTR.

[Kimbark, 1970] Kimbark, E. W. (1970). Transient Overvoltages Caused by Monopolar Ground Fault on Bipolar DC Line: Theory and Simulation. *IEEE Transactions on Power Apparatus and Systems*, PAS-89(4):584–592.

[Kontos et al., 2014] Kontos, E., Rodrigues, S., Teixeira Pinto, R., and Bauer, P. (2014). Optimization of limiting reactors design for DC fault protection of multi-terminal HVDC networks. In *Proc. IEEE Energy Conversion Congress and Exposition, ECCE*, pages 5347–5354.

[Le Blond et al., 2016] Le Blond, S., Bertho, R., Coury, D. V., and Vieira, J. C. (2016). Design of protection schemes for multi-terminal HVDC systems. *Renewable and Sustainable Energy Reviews*, 56:965–974.

[Leterme et al., 2016] Leterme, W., Beerten, J., and Van Hertem, D. (2016). Non-unit protection of HVDC grids with inductive DC cable termination. *IEEE Transactions on Power Delivery*, 31(2):820–828.

[Leterme et al., 2014] Leterme, W., Tielens, P., De Boeck, S., and Van Hertem, D. (2014). Overview of grounding and configuration options for meshed HVDC grids. *IEEE Transactions on Power Delivery*, 29(6):2467–2475.

[Leterme and Van Hertem, 2014] Leterme, W. and Van Hertem, D. (2014). Reduced Modular Multilevel Converter Model to Evaluate Fault Transients in DC Grids. In *Proc. 12th IET International Conference on Developments in Power System Protection (DPSP)*, Copenhagen.

[Levenberg, 1944] Levenberg, K. (1944). a Method for the Solution of Certain Non-Linear Problems in Least Squares. *Quarterly of Applied Mathematics*, 2(2):164–168.

[Lewis et al., 2016] Lewis, P. T., Grainger, B. M., Al Hassan, H. A., Barchowsky, A., and Reed, G. F. (2016). Fault Section Identification Protection Algorithm for Modular Multilevel Converter-Based High Voltage DC with a Hybrid Transmission Corridor. *IEEE Transactions on Industrial Electronics*, 63(9):5652–5662.

[Li et al., 2017] Li, R., Xu, L., and Yao, L. (2017). DC fault detection and location in meshed multiterminal HVDC systems based on DC reactor voltage change rate. *IEEE Transactions on Power Delivery*, 32(3):1516–1526.

[Ljung, 1987] Ljung, L. (1987). *System Identification - Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.

[Macilwain, 2010] Macilwain, C. (2010). Supergrid. *Nature*, 468:624–625.

[Mahseredjian et al., 2007] Mahseredjian, J., Dennetière, S., Dubé, L., Khodabakhchian, B., and Gérin-Lajoie, L. (2007). On a new approach for the simulation of transients in power systems. *Electric Power Systems Research*, 77(11):1514–1520.

[Mahseredjian et al., 1993] Mahseredjian, J., Lefebvre, S., and Dai Do, X. (1993). A New Method for Time-Domain Modelling of Nonlinear Circuits in Large Linear Networks. In *Proc. of 11th Power Systems Computation Conference (PSCC)*, pages 915–922, Avignon.

[Marquardt, 1963] Marquardt, D. W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441.

[Marten et al., 2015] Marten, A.-K., Troitzsch, C., and Westermann, D. (2015). Non-telecommunication based DC line fault detection methodology for meshed HVDC grids. In *Proc. AC and DC Power Transmission (ACDC)*, Birmingham.

[Meghwani et al., 2018] Meghwani, A., Srivastava, S. C., and Chakrabarti, S. (2018). Local measurement-based technique for estimating fault location in multi-source DC microgrids. *IET Generation, Transmission and Distribution*, 12(13):3305–3313.

[Moagar-Poladian and Moagar-Poladian, 2014] Moagar-Poladian, V. and Moagar-Poladian, G. (2014). MEMS sensors for high voltage lines. In *Proc. European Solid-State Device Research Conference*, pages 358–361, Bucharest.

[Morched et al., 1999] Morched, A., Gustavsen, B., and Tartibi, M. (1999). A Universal Model for Accurate Calculation of Electromagnetic Transients on Overhead Lines and Underground Cables. *IEEE Transactions on Power Delivery*, 14(3):1032–1038.

[Pang and Wei, 2018] Pang, H. and Wei, X. (2018). Research on Key Technology and Equipment for Zhangbei 500kV DC Grid. In *Proc. International Power Electronics Conference (IPEC-ECCE Asia)*, pages 2343–2351, Niigata. IEEJ Industry Application Society.

[Pesch et al., 2014] Pesch, T., Allelein, H., and Hake, J. (2014). Impacts of the transformation of the German energy system on the transmission grid. *The European Physical Journal*, 2575:2561–2575.

[Pierri et al., 2017] Pierri, E., Binder, O., Hemdan, N. G., and Kurrat, M. (2017). Challenges and opportunities for a European HVDC grid. *Renewable and Sustainable Energy Reviews*, 70:427–456.

[Pipelzadeh et al., 2015] Pipelzadeh, Y., Chaudhuri, B., and Green, T. C. (2015). Modelling and Dynamic Operation of the Zhoushan DC Grid: Worlds First Five-Terminal VSC-HVDC Project. In *Proc. International High Voltage Direct Current*, pages 87–95, Seoul.

[Pirooz Azad and Van Hertem, 2017] Pirooz Azad, S. and Van Hertem, D. (2017). A Fast Local Bus Current-Based Primary Relaying Algorithm for HVDC Grids. *IEEE Transactions on Power Delivery*, 32(1):193–202.

[PROMOTIoN WP4, 2018] PROMOTIoN WP4 (2018). Report on the broad comparison of protection philosophies for the identified grid topologies. Technical report.

[PROMOTIoN WP6, 2016] PROMOTIoN WP6 (2016). Develop system level model for hybrid DC CB. Technical report.

[Razzaghi et al., 2013] Razzaghi, R., Lugrin, G., Manesh, H., Romero, C., Paolone, M., and Rachidi, F. (2013). An efficient method based on the electromagnetic time reversal to locate faults in power networks. *IEEE Transactions on Power Delivery*, 28(3):1663–1673.

[Razzaghi et al., 2014] Razzaghi, R., Paolone, M., Rachidi, F., Descloux, J., Raison, B., and Retière, N. (2014). Fault location in multi-terminal HVDC networks based on Electromagnetic Time Reversal with limited time reversal window. In *Proc. Power Systems Computation Conference, PSCC*, pages 1–7, Wroclaw.

[Reed et al., 2019] Reed, L., Granger Morgan, M., Vaishnav, P., and Armanios, D. E. (2019). Converting existing transmission corridors to HVDC is an overlooked option for increasing transmission capacity. *Proceedings of the National Academy of Sciences of the United States of America*, 116(28):13879–13884.

[Saad et al., 2014] Saad, H., Dennetiere, S., Mahseredjian, J., Delarue, P., Guillaud, X., Peralta, J., and Nguefeu, S. (2014). Modular multilevel converter models for electromagnetic transients. *IEEE Transactions on Power Delivery*, 29(3):1481–1489.

[Savitzky and Golay, 1964] Savitzky, A. and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639.

[Schaber et al., 2012] Schaber, K., Steinke, F., Pascal, M., and Hamacher, T. (2012). Parametric study of variable renewable energy integration in Europe : Advantages and costs of transmission grid extensions. *Energy Policy*, 42:498–508.

[Schlachtberger et al., 2017] Schlachtberger, D. P., Brown, T., Schramm, S., and Greiner, M. (2017). The Benefits of Cooperation in a Highly Renewable European Electricity Network. *Energy*, 134:469–481.

[Shinoda et al., 2019] Shinoda, K., Benchaib, A., Dai, J., and Guillaud, X. (2019). Virtual Capacitor Control for Stability Improvement of HVDC System Comprising DC Reactors. In *Proc. 15th IET International Conference on AC and DC Power Transmission (ACDC)*, pages 1–6, Belfast.

[Suonan et al., 2010] Suonan, J., Gao, S., Song, G., Jiao, Z., and Kang, X. (2010). A novel fault-location method for HVDC transmission lines. *IEEE Transactions on Power Delivery*, 25(2):1203–1209.

[Suonan et al., 2013] Suonan, J., Zhang, J., Jiao, Z., Yang, L., and Song, G. (2013). Distance Protection for HVDC Transmission Lines Considering Frequency-Dependent Parameters. *IEEE Transactions on Power Delivery*, 28(2):723–732.

[Tang et al., 2018] Tang, L., Dong, X., Shi, S., and Wang, B. (2018). Analysis of the characteristics of fault-induced travelling waves in MMC-HVDC grid. In *Proc. 14th International Conference on Developments in Power System Protection (DPSP)*, volume 2018, pages 1349–1353.

[Terzija and Dobrijevic, 2007] Terzija, V. V. and Dobrijevic, D. M. (2007). Short circuit studies in transmission networks using improved fault model. In *Proc. IEEE POWER TECH*, pages 1752–1757, Lausanne.

[Tong et al., 2019] Tong, N., Lin, X., Li, Y., Hu, Z., Jin, N., Wei, F., and Li, Z. (2019). Local Measurement-Based Ultra-High-Speed Main Protection for Long Distance VSC-MTDC. *IEEE Transactions on Power Delivery*, 34(1):353–364.

[Torwelle, 2021] Torwelle, P. (2021). *Development of a HVDC grid protection strategy based on hybrid OHL and cable lines*. PhD thesis, Université Grenoble Alpes.

[Torwelle et al., 2020] Torwelle, P., Bertinato, A., Raison, B., Le, T. D., and Petit, M. (2020). Compliance of a Cable-Based Protection Strategy for Ohl Grids : Analysis and Remediation. In IET, editor, *Proc. 15th IET International Conference on Developments in Power System Protection (DPSP)*, pages 1–6, Liverpool.

[Transtrum and Sethna, 2012] Transtrum, M. K. and Sethna, J. P. (2012). Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization. *arXiv preprint arXiv:1201.5885*.

[Tzelepis et al., 2020] Tzelepis, D., Psaras, V., Tsotsopoulou, E., Mirsaeidi, S., Dysko, A., Hong, Q., Dong, X., Blair, S. M., Nikolaidis, V. C., Papaspiliotopoulos, V., Fusiek, G., Burt, G. M., Niewczas, P., and D. Booth, C. (2020). Voltage and Current Measuring Technologies for High Voltage Direct Current Supergrids: A Technology Review Identifying the Options for Protection, Fault Location and Automation Applications. *IEEE Access*, 8(November):203398–203428.

[Van Hertem and Ghandhari, 2010] Van Hertem, D. and Ghandhari, M. (2010). Multi-terminal VSC HVDC for the European supergrid: Obstacles. *Renewable and Sustainable Energy Reviews*, 14(9):3156–3163.

[Van Hertem et al., 2016] Van Hertem, D., Gomis-Bellmunt, O., and Liang, J.

(2016). *HVDC Grids For Offshore and Supergrid of the Future*. Wiley-IEEE Press.

[Walter, 2014] Walter, E. (2014). *Numerical Methods and Optimization: A Consumer Guide*. Springer, Cham.

[Walter and Pronzato, 1997] Walter, E. and Pronzato, L. (1997). *Identification of parametric models from experimental data*. Springer-Verlag London.

[Wasserrab and Balzer, 2015] Wasserrab, A. and Balzer, G. (2015). The significance of frequency-dependent overhead lines for the calculation of HVDC line short-circuit currents. *Electrical Engineering*, 97(3):213–223.

[Wedepohl, 1963] Wedepohl, L. (1963). Application of matrix methods to the solution of travelling-wave phenomena in polyphase systems. In *Proceedings of the Institution of Electrical Engineers*, volume 110, page 2200.

[Wedepohl and Nguyen, 1996] Wedepohl, L. M. and Nguyen, H. V. (1996). Frequency-dependent transformation matrices for untransposed transmission lines using newton-raphson method. *IEEE Transactions on Power Systems*, 11(3):1538–1546.

[Xemard et al., 2021] Xemard, A., Sellin, E., Tarafi, R., Bertinato, A., and Verrax, P. (2021). Lightning overvoltages on a dc transmission line, calculated based on measured bipolar lightning strokes. *Electric Power Systems Research*, 197:107331.

[Xi et al., 2016] Xi, C., Chen, Q., and Wang, L. (2016). A single-terminal traveling wave fault location method for VSC-HVDC transmission lines based on S-transform. In *Proc. IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pages 1008–1012, Xi'an.

[You et al., 2010] You, M., Zhang, B. H., Cheng, L. Y., Bo, Z. Q., and Klimek, A. (2010). Lightning model for HVDC transmission lines. In *Proc. 10th IET International Conference on Developments in Power System Protection (DPSP)*, volume 2010, pages 0–4, Manchester.

[Zarazua de Rubens and Noel, 2019] Zarazua de Rubens, G. and Noel, L. (2019). The non-technical barriers to large scale electricity networks: Analysing the case for the US and EU supergrids. *Energy Policy*, 135(October):111018.

[Zhang et al., 2020] Zhang, C., Song, G., Wang, T., Wu, L., and Yang, L. (2020). Non-unit Traveling Wave Protection of HVDC Grids Using Levenberg-Marquart Optimal Approximation. *IEEE Transactions on Power Delivery*, PP(c):1–1.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Algorithme d'identification des défauts pour la protection des réseaux maillés HVDC

Mots clés: réseau HVDC, protection, identification des paramètres

Résumé: La protection des réseaux maillés HVDC se heurte à différents verrous parmi lesquels l'identification fiable et rapide des défauts affectant les lignes du réseau. Cette thèse propose un concept d'algorithme d'identification basé sur l'estimation des paramètres du défaut. Un modèle analytique décrivant l'évolution de la tension et du courant dans la ligne en défaut est développé. Le modèle obtenu dépend explicitement des paramètres du défaut (e.g. la distance et la résistance du défaut). Lorsqu'un défaut est détecté, le relais protégeant la ligne considérée estime les paramètres du défaut, a priori inconnus, à partir des mesures reçues des capteurs les plus proches. Le résultat de l'estimation des paramètres permet de confirmer ou non que la ligne considérée est en défaut. En particulier, un indicateur de confiance évalue la précision des paramètres estimés. La méthode proposée est testée sur un réseau 4 terminaux basé sur des lignes aériennes, implémenté dans un logiciel de simulation des transitoires électromagnétiques. La présence d'inductances de lignes n'est pas requise pour identifier fiablement les défauts. L'algorithme est capable d'identifier la plupart des défauts en utilisant une fenêtre de mesure réduite (moins de 0.5 ms). L'algorithme d'identification peut être intégré dans une stratégie de protection sélective, où chaque élément du réseau est protégé individuellement. La compatibilité de l'approche avec les aspects systèmes tels que le dimensionnement des disjoncteurs ou les spécifications des capteurs est aussi vérifiée.

Title: Transient-based fault identification algorithm for the protection of meshed HVDC grids

Keywords: HVDC grid, protection, parameter estimation

Abstract: The protection of meshed HVDC grids comprises many locks, among which the reliable and fast identification of faults affecting the lines of the grid is seen as particularly challenging. The present thesis investigates the concept of a single-ended fault identification algorithm based on the estimation of the fault parameters. A closed-form model of the transient evolution of the voltage and current in a faulty line is first developed. The obtained model depends explicitly on the fault parameters (e.g. the fault distance and resistance) and is thus adaptable to various fault cases. When a fault is suspected at a relay protecting one of the transmission lines, the proposed algorithm estimates the unknown fault parameters based on the received measurements and the developed model. The result of the parameter estimation process allows the relay to confirm or reject that the protected faulty line is faulty or not. In particular, a confidence indicator is employed to evaluate the accuracy of the estimated fault parameters. The proposed method is tested on a 4 station meshed grid based on overhead lines, implemented in Electro-Magnetic Transient software. Contrary to most existing approaches, the presence of DC reactors at the extremity of each lines is not required for the algorithm to distinguish between faults occurring on different transmission lines. The proposed approach is able to identify most of the faults while using a short observation window (less than 0.5 ms long). The fault identification algorithm is thus a good candidate to be used in a selective fault clearing strategy, where each elements of the grid is protected individually. The compliance of the approach with system requirements such as the DC circuit breakers and the sensor specifications is also investigated.

Université Paris-Saclay
3 rue Joliot Curie, Bâtiment Bréguet, 91190 Gif-sur-Yvette