Robust Faulted Line-section Location for Distribution Networks Based on Normalized Quantile Hausdorff Distance

Guangxiao Zhang, *Member*, *IEEE*, Gaoxi Xiao, *Senior Member*, *IEEE*, Xinghua Liu, *Senior Member*, *IEEE*, Yan Xu, *Senior Member*, *IEEE*, and Peng Wang, *Fellow*, *IEEE*

***Abstract*—This paper proposes a robust faulted line-section location method based on the normalized quantile Hausdorff distance (NQHD) algorithm for detecting single-phase-to-ground faults in distribution networks. The faulted line section is determined according to the characteristic differences between the zero-sequence currents on the faulted and healthy line sections. Specifically, the zero-sequence currents at both ends of a healthy line section are highly similar to each other, while such is generally not the case on a faulted line section. The NQHD algorithm can disregard extremes or outliers while also providing a normalized scaling in different scenarios. Thus, it can be applied to calculate the robust waveform similarity of zero-sequence current waveforms at both ends of different line sections for identifying reliably the faulted line section even under the interference of outliers. The results demonstrate the good performance of the proposed method in detecting single-phase-to-ground faults under different fault conditions. Comparative tests with the existing methods confirm the advantageous robustness of the proposed method against the impacts of outliers and noises.**

***Index Terms*—Faulted line-section location, single-phase-to-ground fault, distribution network, normalized quantile Hausdorff distance, outlier.**

I. Introduction

NEUTRAL ineffectively grounded networks are commonly employed in medium-voltage (MV) distribution networks [1], [2]. A key advantage is that following a single-phase-to-ground fault, the phase-to-phase voltage remains symmetrical, typically allowing the system to operate for an additional 1-2 hours. However, single-phase-to-ground faults constitute over 70% of all faults [3], and the normal phase voltage escalates to be nearly the line voltage during faults. Operating an MV distribution network for an extended period during a single-phase-to-ground fault may not only endanger the safety of personnel and equipment, but also potentially lead to more severe phase-to-phase faults as a result of insulation breakdown [4]. For secure operation and the implementation of automated fault management systems in distribution networks, it is crucial to detect faulted line section accurately and reliably.

To address the challenge of locating line sections with single-phase-to-ground faults in distribution networks, several faulted line-section location methods have been proposed. These methods are generally classified into three categories: active injection-based, signal processing-based, and learning-based. Reference [5] proposes an active injection-based method for locating the faulted line section using signal injection driven by the soft open point; however, the signal injection method for fault detection has limitations, including system interference, waveform/frequency selection, complexity, and limited accuracy and sensitivity of sensor/device. With the development of artificial intelligence (AI) technology, there has been increasing attention to learning-based methods. In [6], the eigenvalues of the time-series signals are used to train an improved *K*-means clustering model for locating the faulted line section. A method for faulted line-section location in resonant grounding distribution networks utilizing waveform concatenation and 1-dimensional convolutional neural network (1-D CNN) is proposed in [7]. In [8], a faulted line-section location method based on an autoencoder and a backpropagation neural network is proposed for detecting single-phase-to-ground faults in distribution networks by extracting features from transient zero-sequence current and voltage. Nevertheless, it should be noted that the above learning-based methods may face certain limitations such as a considerable need for data quantity and dimensionality.

The advantage of signal processing-based methods lies in their use of existing measurement signals for fault identification, thereby avoiding the need for additional signal injection equipment as in active injection-based methods, and reducing the data burden associated with learning-based meth-

Manuscript received: December 5, 2024; revised: March 25, 2025; accepted: June 7, 2025. Date of CrossCheck: June 7, 2025. Date of online publication: July 3, 2025.

This work was supported by the Future Resilient Systems (FRS-II) Project at the Singapore-ETH Centre (SEC), which was funded by the National Research Foundation of Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) program.

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

G. Zhang is with the Institute of Catastrophe Risk Management, Nanyang Technological University, Singapore, and he is also with Future Resilient Systems, Singapore-ETH Centre, Singapore (e-mail: guangxiao.zhang@ntu.edu.sg).

G. Xiao (corresponding author), Y. Xu, and P. Wang are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail: egxxiao@ntu.edu.sg; xuyan@ntu.edu.sg; epwang@ntu.edu.sg).

X. Liu is with the School of Electrical Engineering, Xi'an University of Technology, Xi'an 710048, China (e-mail: liuxh@xaut.edu.cn).

DOI: 10.35833/MPCE.2024.001299

ods. Reference [9] proposes a feeder terminal unit (FTU)-based method for identifying single-phase high-impedance faults (HIFs) in resonant grounding systems using transient zero-sequence admittance. This method remains effective regardless of fault inception angles, arcing faults, or HIFs. The degree of distortion in equivalent admittances is utilized in [10] to locate line section in distribution networks under single-phase-to-ground fault. In [11], the leakage energy, derived from both fault phase voltage and current, is utilized as the criterion for faulted line-section location. However, these methods require additional sensors to collect voltage and current signals, which may increase the cost and complexity of the faulted line-section location system. Reference [12] develops an MV faulted line-section location method that relies solely on unsynchronized low-voltage (LV) measurements in distribution networks. A new faulted line-section location method for HIF is proposed in [13], utilizing the significant differences in the declining periodic components of zero-sequence current, which are obtained through applying the extended Prony method at upstream and downstream of the fault point. Reference [14] formulates a linear model for faulted line-section location, utilizing the distribution characteristics of transient zero-sequence current direction, which proves capable of pinpointing the faulted line section with accuracy and efficiency. In [15], a faulted line-section location method based on the Hausdorff distance (HD) is proposed by comparing the similarity of transient zero-sequence currents at both ends of line sections to identify the faulted line section. In [16], a dynamic time warping (DTW)-based method for locating faulted line sections is proposed. This method utilizes waveform similarity calculations for identifying single-phase-to-ground faults in distribution networks and offers the advantage of not requiring strict time synchronization.

The key limitation of these methods is that their reliability can be significantly degraded by outliers caused by communication failures, measurement interference, or malicious cyber activities. As outliers sent to a faulted line-section location system may lead to incorrect decision-making, some research works have concentrated on developing robust techniques to ensure the cyber security of locating faulted line sections. The methods presented in [17]-[19] can address data incompleteness and uncertainties by utilizing multi-source binary decision information. Since [17]-[19] do not provide a method for obtaining binary decision information through electrical signals, [20] bridges this gap by detailing this step. Subsequently, [20] employs the mixed-integer linear programming (MILP) to combine multi-source binary decision information, enhancing the robustness of faulted cable-section location when the decision information contains anomalous data. To effectively address the outliers in electrical signals, a modified Hausdorff distance (MHD)-based faulted line-section location method is proposed in [21]. By using the average value instead of the maximum, the MHD algorithm resists minor outlier interference but remains sensitive to extreme values. Furthermore, some robust fault detection methods designed for transmission networks can be adapted for use in distribution networks. For instance, a learning-based framework is utilized in [22]-[24] to detect outliers during the fault detection process, thereby guaranteeing reliable fault detection results. Nonetheless, the learning-based methods necessitate a substantial volume of data for training the outlier detection model. To ensure the safety of fault detection in scenarios where outliers cannot be entirely removed from the current waveforms, [25] proposes a robust fault detection method based on Kendall's tau coefficient (KTC) to minimize the impact of undetected outliers. The KTC-based method relies on rank order rather than metric values, thus keeping outliers at a manageable level. However, KTC-based method can produce inaccurate results in the presence of strong noise interference because the ordering of data points can be altered under such conditions.

In this paper, a robust faulted line-section location method based on normalized quantile Hausdorff distance (NQHD) algorithm is proposed for detecting single-phase-to-ground faults in distribution networks. The distinction between faulted and healthy line sections is achieved by evaluating the NQHD value between the zero-sequence currents at both ends of the line sections. The effectiveness of the proposed method is demonstrated under different conditions including fault distances, fault resistances, and fault inception angles. Its robustness is further validated in the presence of HIF. The main contributions are summarized as follows.

1) The proposed method enhances the reliability of single-phase-to-ground fault detection in distribution networks by robustly assessing the similarity of sampled zero-sequence currents. Compared with existing faulted line-section location methods and traditional current differential techniques, the proposed method demonstrates greater robustness, reliably detecting faults even when the raw signals contain multiple outliers. Furthermore, it eliminates the need for additional outlier detection mechanisms and, in contrast to multi-source data fusion methods, requires fewer measurement devices and a less complex decision-making process, making it a more efficient and practical solution.

2) By utilizing quantiles to disregard outliers and leveraging metric values rather than rank order to minimize variations caused by noise interference, the NQHD algorithm focuses on the overall data distribution. This enables the proposed method to maintain its performance, offering an advantage over commonly used similarity metrics in fault detection, which may suffer from misjudgments and performance degradation in presence of outliers and noises.

3) The proposed method leverages the diagonal of a robust bounding box for two sets of points as a natural scaling factor to normalize distances, ensuring consistency and comparability across different datasets and scenarios. By using percentiles instead of the absolute minimum and maximum values, the proposed method offers resistance to outliers and enhances its robustness. This not only makes the normalized distance more intuitive and consistent across different scales, but also simplifies the threshold setting, particularly in the presence of extreme values.

The remainder of this paper is structured as follows. Section II analyzes the fault characteristics of the zero-sequence currents under single-phase-to-ground faults in both faulted

and healthy line sections. Section III presents the faulted line-section location method based on the NQHD algorithm. Section IV describes the system structure of the proposed method. Section V presents case studies and discussion to validate the effectiveness of the proposed method. This paper concludes in Section VI.

II. Fault Characteristics of Zero-sequence Currents

This section focuses on identifying the distinguishing features of zero-sequence current waveforms between faulted and healthy line sections in a distribution network during single-phase-to-ground faults. Figure 1 illustrates the equivalent zero-sequence circuit during a single-phase-to-ground fault that occurs at point *f* within the line section *MN* of feeder L_1 in a typical distribution network with an arc suppression coil [21]. It shows four detection devices located at points *M*, *N*, *P*, and *Q* along the feeder L_1, with i_M, i_N, i_P, and i_Q denoting the corresponding zero-sequence currents measured by each detection device.

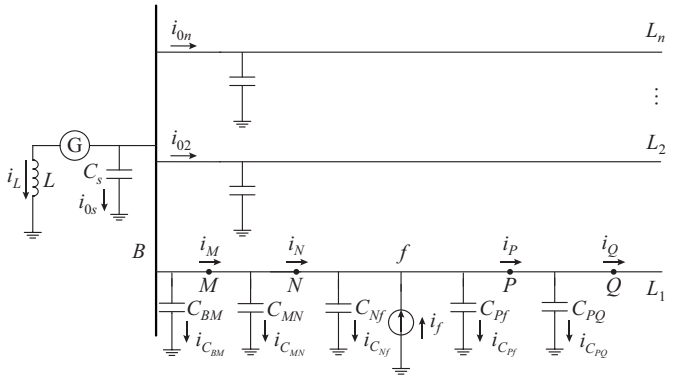

Fig. 1. Equivalent zero-sequence circuit during a single-phase-to-ground fault.

A. Line Section with an External Fault

The zero-sequence current i_M observed at detection point *M*, which is located upstream of the fault point *f* on the faulted feeder, can be expressed as:

$$i_M = i_{C_{BM}} + i_\Sigma + i_L \tag{1}$$

where $i_{C_{BM}}$ is the grounding capacitance current of equivalent capacitances C_{BM} on the line section *BM*, which connects the busbar node *B* to the detection point *M*; i_Σ is the sum of the grounding capacitive currents of all healthy feeders and transformer, given by $i_\Sigma = i_{0s} + i_{02} + ... + i_{0n}$, where i_{0s} is the capacitive current through the transformer-side capacitance C_s, and i_{02}-i_{0n} represent the capacitive currents from the healthy feeders L_2-L_n; and i_L is the inductance current generated by the arc suppression coil *L*.

Correspondingly, the zero-sequence current i_N detected at detection point *N* can be expressed as:

$$i_N = i_{C_{BM}} + i_{C_{MN}} + i_\Sigma + i_L \tag{2}$$

where $i_{C_{MN}}$ is the grounding capacitance current of equivalent capacitances C_{MN} on the line section *MN*.

By comparing (1) and (2), we can obtain the differences in zero-sequence current between the two ends of the healthy line section, which is upstream of the fault point:

$$i_M - i_N = i_{C_{MN}} \tag{3}$$

Similarly, the differences in zero-sequence currents between detection points *P* and *Q* for the healthy line section *PQ* downstream of the fault point can be obtained as:

$$i_P - i_Q = i_{C_{PQ}} \tag{4}$$

where $i_{C_{PQ}}$ is the grounding capacitance current of equivalent capacitances C_{PQ} on the line section *PQ*.

Equations (3) and (4) indicate that the differences in zero-sequence currents detected at both ends of the healthy line section are mainly influenced by the grounding capacitance currents of the equivalent capacitance between them. Given that the length between the two ends of the line section is relatively short, the grounding capacitance current between adjacent detection points can be ignored. Consequently, for healthy line sections, the zero-sequence currents observed at two adjacent detection points exhibit similar waveforms, i.e., $i_M \approx i_N$ and $i_P \approx i_Q$.

B. Line Section with an Internal Fault

In the case of a faulted line section, the difference in zero-sequence current between detection points *N* and *P* can be described by:

$$i_N - i_P = i_{C_{Nf}} + i_{C_{Pf}} - i_f \tag{5}$$

where $i_{C_{Nf}}$ and $i_{C_{Pf}}$ are the grounding capacitance currents of the equivalent capacitances C_{Nf} and C_{Pf}, respectively; and i_f is the fault current that flows through the fault point.

According to (5), the currents associated with grounding capacitance are significantly smaller than the fault current present in the faulted line section, resulting in a notable difference in magnitude between i_N and i_P. Hence, by comparing the similarity differences of zero-sequence current waveforms at both ends of the line section, we can effectively identify the faulted line sections during single-phase-to-ground faults in the distribution network.

C. Considerations in Utilizing Fault Characteristics amid Interfering Factors

During single-phase-to-ground faults, one of the distinguishing signal characteristics at different detection points is the substantial increase in the zero-sequence current. Under normal operation with balanced three-phase currents, this zero-sequence current is almost negligible. However, when a fault occurs, this zero-sequence current exhibits a rapid increase. Single-phase-to-ground faults can be categorized into three types: Phase A to ground, Phase B to ground, and Phase C to ground. Given that the zero-sequence current represents the collective sum of the currents of Phases A, B, and C, analyzing it allows for the effective identification of these single-phase-to-ground faults. Based on the analysis in the two preceding subsections, a notable feature emerges that differentiates faulted line sections from healthy ones: the zero-sequence current waveforms at both ends of a healthy line section are highly similar, whereas those of a faulted line section display significant differences.

Although similarity metrics such as HD, cosine similarity

(CS), or Pearson correlation coefficient (PCC) can be used to locate faulted line sections, challenges such as data losses and outliers in the cyber system would lead to erroneous fault detection results. To address these challenges, a common solution involves deploying additional outlier detection mechanisms or implementing the fusion of data from multiple sources. However, these solutions necessitate extensive equipment and sophisticated decision-making algorithms, often incurring significant costs. To tackle the problem, this paper aims to develop a robust similarity metrics to locate faulted line section even in the presence of outlier interferences.

III. Faulted Line-section Location Method Based on NQHD Algorithm

This section will start with a brief introduction to the quantile Hausdorff distance (QHD) algorithm. Subsequently, an NQHD algorithm will be provided to offer a normalized scale. Finally, the implementation of the proposed method for locating the faulted line section will be presented.

A. Introduction of QHD

The HD algorithm is a widely used measure of similarity or dissimilarity between two sets in different applications. To compute the HD value between two vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, which contain M items each denoted as x_1, x_2, ..., x_i, ..., x_M and y_1, y_2, ..., y_j, ..., y_M, respectively, we have [15]:

$$H(\boldsymbol{X},\boldsymbol{Y})=\max\{h(\boldsymbol{X},\boldsymbol{Y}),h(\boldsymbol{Y},\boldsymbol{X})\} \tag{6}$$

$$h(\boldsymbol{X},\boldsymbol{Y})=\max_{x_i\in\boldsymbol{X}}\left\{\min_{y_j\in\boldsymbol{Y}}\left\|x_i-y_j\right\|\right\} \tag{7}$$

$$h(\boldsymbol{Y},\boldsymbol{X})=\max_{y_j\in\boldsymbol{Y}}\left\{\min_{x_i\in\boldsymbol{X}}\left\|x_i-y_j\right\|\right\} \tag{8}$$

where $\|\cdot\|$ represents the Euclidean distance, defined on the points of $\boldsymbol{X}$ and $\boldsymbol{Y}$. Equation (7) signifies that the minimal distance is initially determined for each point in $\boldsymbol{X}$ relative to $\boldsymbol{Y}$. Subsequently, the maximum value from this set of minimum distances is selected as the value of $h(\boldsymbol{X},\boldsymbol{Y})$. The metric $h(\boldsymbol{Y},\boldsymbol{X})$ in (8) can be calculated in a similar manner. Finally, $H(\boldsymbol{X},\boldsymbol{Y})$ in (6) can be derived by choosing the larger value between $h(\boldsymbol{X},\boldsymbol{Y})$ and $h(\boldsymbol{Y},\boldsymbol{X})$.

The HD algorithm measures the difference between two sets by calculating the maximum distance between points in the sets. This method is particularly effective in fault signal analysis for its ability to emphasize discrepancies based on amplitude, contrasting with CS or PCC which focuses on direction aspects. By concentrating on the largest minimum distances, it offers better robustness against noise and a more prominent measurement of dissimilarities compared with average-based methods like Euclidean distance or rank-order metrics such as the KTC method. However, it is highly sensitive to outliers or extreme values, which can lead to inaccurate distance estimation. In contrast, the QHD algorithm is more robust to the interference of outliers compared with the HD algorithm due to its utilization of quantiles. By using quantiles, QHD can disregard extreme or outlier values and focus more on the overall distribution of data points. This enhances the robustness of distance estimation, resulting in a more reliable measure, particularly when dealing with outliers. The QHD metric between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined as [26]:

$$H_q(\boldsymbol{X},\boldsymbol{Y})=\max\{h_q(\boldsymbol{X},\boldsymbol{Y}),h_q(\boldsymbol{Y},\boldsymbol{X})\} \tag{9}$$

$$h_q(\boldsymbol{X},\boldsymbol{Y})=\underset{x_i\in\boldsymbol{X}}{quantile^q}\left\{\min_{y_j\in\boldsymbol{Y}}\left\|x_i-y_j\right\|\right\} \tag{10}$$

$$h_q(\boldsymbol{Y},\boldsymbol{X})=\underset{y_j\in\boldsymbol{Y}}{quantile^q}\left\{\min_{x_i\in\boldsymbol{X}}\left\|x_i-y_j\right\|\right\} \tag{11}$$

where $h_q(\boldsymbol{X},\boldsymbol{Y})$ indicates that the shortest distance is first computed for each point in $\boldsymbol{X}$ with respect to $\boldsymbol{Y}$; and the function $quantile^q\{\cdot\}$ denotes the q^{th} quantile of the given set. Following this, the q^{th} quantile from these minimum distances is selected as the value of $h_q(\boldsymbol{X},\boldsymbol{Y})$. Here, q represents the quantile value, which ranges between 0 and 1. Similarly, the metric $h_q(\boldsymbol{Y},\boldsymbol{X})$, as specified in (11), can be computed. Finally, the value of $H_q(\boldsymbol{X},\boldsymbol{Y})$, as specified in (9), is obtained by choosing the larger value between $h_q(\boldsymbol{X},\boldsymbol{Y})$ and $h_q(\boldsymbol{Y},\boldsymbol{X})$.

The QHD algorithm is used to assess the similarity between zero-sequence currents of healthy and faulted line sections. The fault characteristics discussed in Section II suggest that the zero-sequence currents in $\boldsymbol{X}$ and $\boldsymbol{Y}$ exhibit high similarity for healthy line sections, as reflected by a relatively small QHD value. However, for the faulted line section, the zero-sequence currents in $\boldsymbol{X}$ and $\boldsymbol{Y}$ exhibit significant dissimilarity, with the largest QHD value. Based on the analysis above, it can be concluded that the QHD values between $\boldsymbol{X}$ and $\boldsymbol{Y}$ can indicate the status of the line section as either faulted or healthy during a single-phase-to-ground fault. As a result, the faulted line section can be found according to (12).

$$H_q(\boldsymbol{X},\boldsymbol{Y})=\max_{s_n\in S}\{H_{q(s_n)}(\boldsymbol{X},\boldsymbol{Y})\} \tag{12}$$

where $H_{q(s_n)}(\boldsymbol{X},\boldsymbol{Y})$ is the QHD value of the n^{th} line section in the set of line sections $S=\{s_1,s_2,...,s_n,...,s_N\}$, which contains N items. In short, (12) indicates that the line section with the maximum QHD value is identified as the faulted line section.

B. NQHD Algorithm

The faulted line-section location method based on QHD algorithm may fail to provide reliable results if it solely relies on selecting the maximum QHD value to distinguish between line sections. Without establishing a set threshold, there is a heightened risk of either overreactions or continuous interventions, as there is no stable benchmark to guide the decision-making.

Normalization is a valuable technique, especially when comparisons across different datasets or scenarios are desired. Min-max normalization is a simple linear scaling technique with the primary objective of rescaling data values to lie within a specified range, which is often [0, 1]. However, a significant limitation of this method is its high sensitivity to outliers. Since the min-max normalization directly utilizes the minimum and maximum values of the data for scaling, a single outlier can drastically affect the scale of normalization. Thus, when outliers are present in the dataset, direct ap-

plication of the min-max normalization might not be ideal. In such scenarios, a more robust method of data scaling or transformation is needed.

To facilitate the threshold setting, the diagonal of a robust bounding box for two sets of points, $\boldsymbol{X}$ and $\boldsymbol{Y}$, is utilized for normalization. This aims to provide a natural scaling factor, allowing the resulting normalized distance to be compared across different datasets and scenarios, even in the presence of outliers. To enhance robustness, the bounding box is defined using percentiles rather than the absolute minimum and maximum values, reducing sensitivity to extreme values.

$$D = |U - L| \tag{13}$$

$$L = quantile^{\alpha}\{\boldsymbol{X} \cup \boldsymbol{Y}\} \tag{14}$$

$$U = quantile^{\beta}\{\boldsymbol{X} \cup \boldsymbol{Y}\} \tag{15}$$

where L is the lower bound determined by the α^{th} percentile of the combined set of $\boldsymbol{X}$ and $\boldsymbol{Y}$; U is the upper bound derived from the β^{th} percentile of the same combined set; α and β are the variables representing the percentiles, with $0 \leq \alpha < \beta \leq 100$; and D is the absolute difference between U and L, representing the overall spread of the data. This metric remains robust to outliers due to the flexible selection of the percentiles α and β. In this paper, α is specifically chosen as the 10^{th} percentile and β as the 90^{th} percentile, because, with 50 samples (10 kHz sampling rate in this paper) used for similarity comparison, removing 10% of data from each tail allows the method to suppress up to 5 outliers. Although a higher sampling frequency allows more outliers to be tolerated, in practical applications, the actual number of outliers may be limited. Moreover, higher sampling rates increase data transmission, leading to greater communication overhead and computational load, especially in extended networks. Therefore, a balance between outlier suppression and data preservation is achieved, allowing consecutive outliers to be removed without significantly reducing the amount of useful information.

To emphasize the differences between two sets of points and enhance computational efficiency by eliminating square root operations, the squared Euclidean distance is chosen over the standard Euclidean distance in QHD calculations. Consequently, (10) and (11) can be reformulated, with the original terms $h_q(\boldsymbol{X}, \boldsymbol{Y})$ and $h_q(\boldsymbol{Y}, \boldsymbol{X})$ being replaced by $h_q^2(\boldsymbol{X}, \boldsymbol{Y})$ and $h_q^2(\boldsymbol{Y}, \boldsymbol{X})$, as:

$$h_q^2(\boldsymbol{X}, \boldsymbol{Y}) = \underset{x_i \in \boldsymbol{X}}{quantile^{q}} \left\{ \min_{y_j \in \boldsymbol{Y}} \left\| x_i - y_j \right\|^2 \right\} \tag{16}$$

$$h_q^2(\boldsymbol{Y}, \boldsymbol{X}) = \underset{y_j \in \boldsymbol{Y}}{quantile^{q}} \left\{ \min_{x_i \in \boldsymbol{X}} \left\| x_i - y_j \right\|^2 \right\} \tag{17}$$

where $\|\cdot\|^2$ represents the squared Euclidean distance.

By integrating the QHD algorithm based on the squared Euclidean distance with the diagonal of a robust bounding box for two sets of points, an NQHD algorithm is presented as:

$$HD(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\max\{h_q^2(\boldsymbol{X}, \boldsymbol{Y}), h_q^2(\boldsymbol{Y}, \boldsymbol{X})\}}{D} \tag{18}$$

where $HD(\boldsymbol{X}, \boldsymbol{Y})$ denotes an NQHD value computed using the squared Euclidean distance.

C. Faulted Line-section Location Method Based on NQHD Algorithm

Based on the analysis above, it can be concluded that the NQHD values between $\boldsymbol{X}$ and $\boldsymbol{Y}$ can indicate the status of the line section as either faulted or healthy during a single-phase-to-ground fault. As a result, fault detection criteria can be established by using the noticeable differences in the NQHD values to distinguish between faulted and healthy line sections.

Under stable network conditions, the threshold can be empirically determined, while a dynamic threshold is proposed to ensure robustness in different operation scenarios. Specifically, if the following condition is met, the line section will be identified as faulted:

$$HD(\boldsymbol{X}, \boldsymbol{Y}) \geq \frac{1}{2} \sum_i HD_i \tag{19}$$

where $HD(\boldsymbol{X}, \boldsymbol{Y})$ is the NQHD value of the current line section; and $\sum_i HD_i$ is the total sum of NQHD values across all line sections. A line section is identified as faulted if its NQHD value exceeds half of the total sum of NQHD values of all line sections. Setting half of the total sum of NQHD values of all line sections as the threshold is based on the fault contribution ratio principle. In the presence of a fault, the NQHD value of the faulted line section is significantly higher than that of non-faulted line sections, contributing disproportionately to the total NQHD distribution. The proposed method ensures that the threshold dynamically adjusts to different network conditions without relying on predefined parameters. By setting the threshold at half of the total sum of NQHD values, the method effectively captures the dominance of the faulted line section while maintaining adaptability to different feeder configurations and operating conditions. Note that this method requires a centralized system for processing, as it involves aggregating and comparing NQHD values across multiple sections. Such a method can be effectively implemented within the centralized decision-making system described in Section IV.

The proposed method locates faulted and healthy line sections by measuring the similarity between two datasets with fault characteristics based on the NQHD algorithm. The flowchart of the proposed method is shown in Supplementary Material A Fig. SA1, and the step-by-step implementation details are outlined below.

Step 1: detect whether a fault occurs. The zero-sequence voltage is continuously monitored for its instantaneous value. Subsequently, the following start-up criterion is applied to detect the presence of a fault condition [9].

$$\Delta u_0(t) \geq k u_n \tag{20}$$

where $\Delta u_0(t)$ is an increment of zero-sequence voltage u_0 at instant t; u_n is the rated voltage; and the coefficient k, which has an empirical value of 0.15, is used to sensitively distinguish between normal and faulted conditions in different scenarios while maintaining an appropriate speed of action. If the condition is met for three consecutive samples beginning at instant t, the algorithm proceeds to the following step.

Step 2: acquire the measurement data from both sides of each line section in distribution networks. Once a fault condition is detected, the zero-sequence currents are collected. As a result, $\boldsymbol{X}$ and $\boldsymbol{Y}$, each containing M samples for a quarter-cycle data window, are obtained at both ends of each line section.

Step 3: calculate the NQHD values for each line section. This is achieved by computing the NQHD values between $\boldsymbol{X}$ and $\boldsymbol{Y}$ using (18). The similarity metrics $HD(\boldsymbol{X}, \boldsymbol{Y})$ of each line section s_n $(n = 1, 2, ..., N)$ can then be utilized as the fault indicator.

Step 4: locate the faulted line section. The NQHD value is evaluated to determine whether a line section satisfies the criterion described in (19). If the criterion is satisfied, the line section is identified as the faulted line section. Conversely, if the criterion is not satisfied, the line section is considered a healthy one.

IV. System Structure of Proposed Method

The overall system structure of the proposed method is depicted in Fig. 2. The proposed method can be implemented into an Internet of Things (IoT) infrastructure, which would offer advantages such as enhanced real-time monitoring, the potential for advanced data analytics and management, remote access and control capabilities, and improved user's interaction [27].

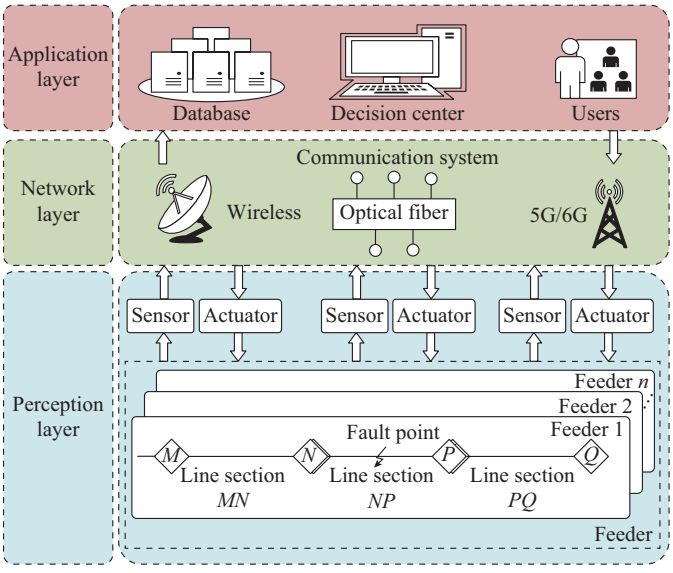

Fig. 2. Overall system structure of proposed method.

As illustrated in Fig. 2, the IoT infrastructure of the proposed method, which is designed to identify faulted line-section locations, consists of three layers: the perception layer, the network layer, and the application layer. At the lowest level of this platform lies the perception layer, equipped with intelligent sensors and actuators. These sensors are tasked with sampling and processing measurements within the distribution network. The raw or pre-processed data are then transmitted to the network layer. Meanwhile, the actuators receive decision-making information and convert it into tangible actions. In the middle, the network layer connects geographically dispersed sensor and actuator nodes from the perception layer, offering functionalities for communication, storage, and data aggregation. Positioned at the top of the hierarchy, the application layer delves into in-depth data analysis and supports a variety of sophisticated applications within the IoT-based faulted line-section location system. The proposed method remains computationally efficient even in large systems with many nodes and lines, because it processes each line section independently and in parallel, although it follows a centralized method. The system simultaneously collects data from all line sections, and the faulted line-section locations for each line section are performed in parallel, ensuring that an increase in the number of lines and nodes does not significantly impact overall computational efficiency. This structure allows the algorithm to scale efficiently with the system size while maintaining real-time performance.

The faulted line-section location system identifies faults using measurements obtained from sensors in the distribution network. When a single-phase-to-ground fault occurs, intelligent sensors in the perception layer capture three-phase currents and extract their zero-sequence components. These data are relayed to the application layer by the network layer utilizing cutting-edge communication technologies such as optical fiber, 5G/6G, and other wireless networks. Once the application layer acquires the data, it employs advanced analytical methods to identify the location of the faulted line section. This fault information is then documented and sent remotely to the client, optimizing the user experience and ensuring efficient human-computer interaction. To circumvent overloading the centralized decision-making center with excessive data transmission, the faulted feeder is identified using the methods outlined in [28], [29], which involve comparing zero-sequence currents from the head-end of each feeder. This narrows down the fault searching range. Once the faulted feeder is identified, information from each line section is transmitted to the central decision-making unit to locate the specific faulted line section. The system platform for faulted line-section location, developed using LabVIEW and MATLAB, is described in Supplementary Material A Fig. SA2.

V. Case Studies and Discussion

The proposed method is validated through simulations in PSCAD/EMTDC using a 10 kV radial distribution network and a modified IEEE 33-node system with distributed energy resources. Details of the network configuration, parameters, and simulation setup are provided in Supplementary Material A Fig. SA3 and Fig. SA4.

A. Results Under Different Fault Conditions in a Radial Distribution Network

1) Effects of Fault Location

A single-phase-to-ground fault is set at the middle of line section *NP* on feeder 3, as shown in Supplementary Materials A Fig. SA3, with a fault inception angle of 30° and a fault resistance of 50 Ω. The zero-sequence current waveforms for the four detection points are subsequently obtained

and depicted in Supplementary Material A Fig. SA5. As can be observed from the waveforms, there is a high similarity between the waveforms of detection points *M* and *N*, located upstream of the fault point, and between those of detection points *P* and *Q*, located downstream of the fault point. However, the waveforms between detection points *N* and *P* do not exhibit similarity. The zero-sequence currents provide the intuitive fault characteristics for discriminating faulted line section from healthy line sections. A more intuitive demonstration of the computational process including the LabVIEW front panel visualization and email notification interface is provided in Supplementary Material A Fig. SA6.

The performance of the proposed method under different fault locations is examined by simulating single-phase-to-ground faults at different sections of feeder 3, using a fault inception angle of 30° and a fault resistance of 1000 Ω. Table I presents the simulation results for different fault locations at the front (1 km), middle (3 km), and end (5 km) of the line sections when a single-phase-to-ground fault occurs. The calculation results presented in Table I demonstrate that when a fault occurs at different line sections and fault locations within the same line section, the faulted line section exhibits a significantly different NQHD value compared with the healthy line sections. Therefore, the proposed method has good generality against variations in fault location.

TABLE I
SIMULATION RESULTS FOR DIFFERENT FAULT LOCATIONS

Fault location	Similarity metric			Faulted line section
	MN	*NP*	*PQ*	
First	0.3696	0.0007	0.0011	*MN*
Middle	0.3678	0.0015	0.0008	*MN*
End	0.2733	0.0010	0.0009	*MN*
First	0.0006	0.3453	0.0012	*NP*
Middle	0.0003	0.2947	0.0046	*NP*
End	0.0012	0.3247	0.0009	*NP*
First	0.0008	0.0012	0.2861	*PQ*
Middle	0.0020	0.0015	0.3140	*PQ*
End	0.0007	0.0012	0.3940	*PQ*

2) Effects of Fault Resistance

To further investigate the impact of fault resistances on the proposed method, a series of single-phase-to-ground faults with different fault resistances are simulated at fault points f_1, f_2, and f_3, as shown in Supplementary Material A Fig. SA3, with a fault inception angle of 30°. Note that f_1, f_2, and f_3 are located in different line sections and are situated at distances of 4, 10, and 16 km from the main busbar, respectively. Table II demonstrates that the proposed method consistently and correctly identifies the faulted line section for different fault resistances.

3) Effects of Fault Inception Angles

Single-phase-to-ground faults occurring at different fault inception angles at fault points f_1, f_2, and f_3 are simulated, with a fault resistance set to be 1000 Ω. The QHD values for different fault inception angles are calculated, as shown in Table III.

TABLE II
SIMULATION RESULTS FOR DIFFERENT FAULT RESISTANCES

Fault location	Fault resistance (Ω)	Similarity metric			Faulted line section
		MN	*NP*	*PQ*	
f_1	5	1.8614	0.0139	0.0060	*MN*
	50	0.2263	0.0148	0.0052	*MN*
	500	0.3954	0.0053	0.0013	*MN*
	1000	0.2891	0.0018	0.0013	*MN*
f_2	5	0.0062	2.0036	0.0021	*NP*
	50	0.0152	1.7431	0.0139	*NP*
	500	0.0023	0.5280	0.0047	*NP*
	1000	0.0010	0.3083	0.0017	*NP*
f_3	5	0.0080	0.0043	1.8110	*PQ*
	50	0.0058	0.0068	2.5164	*PQ*
	500	0.0005	0.0013	0.5619	*PQ*
	1000	0.0015	0.0005	0.3267	*PQ*

TABLE III
SIMULATION RESULTS FOR DIFFERENT FAULT INCEPTION ANGLES

Fault location	Inception angle (°)	Similarity metric			Faulted line section
		MN	*NP*	*PQ*	
f_1	0	0.2997	0.0008	0.0008	*MN*
	60	0.2533	0.0031	0.0013	*MN*
	90	0.1155	0.0035	0.0010	*MN*
	150	0.3593	0.0077	0.0130	*MN*
f_2	0	0.0001	0.2739	0.0025	*NP*
	60	0.0018	0.3178	0.0019	*NP*
	90	0.0014	0.1767	0.0025	*NP*
	150	0.0001	0.3788	0.0121	*NP*
f_3	0	0.0046	0.0006	0.2940	*PQ*
	60	0.0011	0.0011	0.2801	*PQ*
	90	0.0006	0.0007	0.1289	*PQ*
	150	0.0001	0.0003	0.3982	*PQ*

The results shown in Table III reveal that the zero-sequence currents exhibit similar features for healthy line sections, with a smaller QHD value. Conversely, for the faulted line section, zero-sequence currents display dissimilar features, with the largest QHD value. Consequently, it can be concluded that the performance of the proposed method remains unaffected by the fault inception angle.

B. Results Under HIFs

The detection of HIFs is vital for ensuring distribution network safety, reliability, and efficiency while protecting human lives and valuable assets. The HIF model is adopted by employing parallel connections of two sets of series-connected elements, i.e., a variable resistance, a diode, and an adjustable DC source in one set, and a variable resistance, an antiparallel diode, and a reverse adjustable DC source in the other set. This model effectively captures the characteristics of asymmetry, intermittence, non-linearity, randomness, shoulder, and buildup, as described in [30]. The HIF model used in this paper, including its structure and parameter set-

tings, is detailed in Supplementary Material A Fig. SA7. The evaluation of the startup criterion under HIF conditions is presented in Supplementary Material A Fig. SA8.

Figure 3 shows the distorted waveforms of zero-sequence currents resulting from the occurrence of the HIF at the f_2 location on feeder 3. The fault has a fault inception angle of 0°, and the variable resistances R_p and R_n randomly vary within a range of 450-550 Ω every 0.1 ms. As shown in Fig. 3, these zero-sequence currents still effectively capture the distinct fault characteristics between the faulted and healthy line sections, even under HIF conditions. This is because the local nonlinearity distortion at the zero-crossing point, caused by the HIF, does not alter the overall waveform trend. The results for different HIF conditions with variable resistances R_p and R_n, ranging from 450 to 550 Ω, 700 to 800 Ω, and 900 to 1000 Ω, at a fault inception angle of 0°, are presented in Table IV. As demonstrated in Table IV, the proposed method can accurately detect the HIFs.

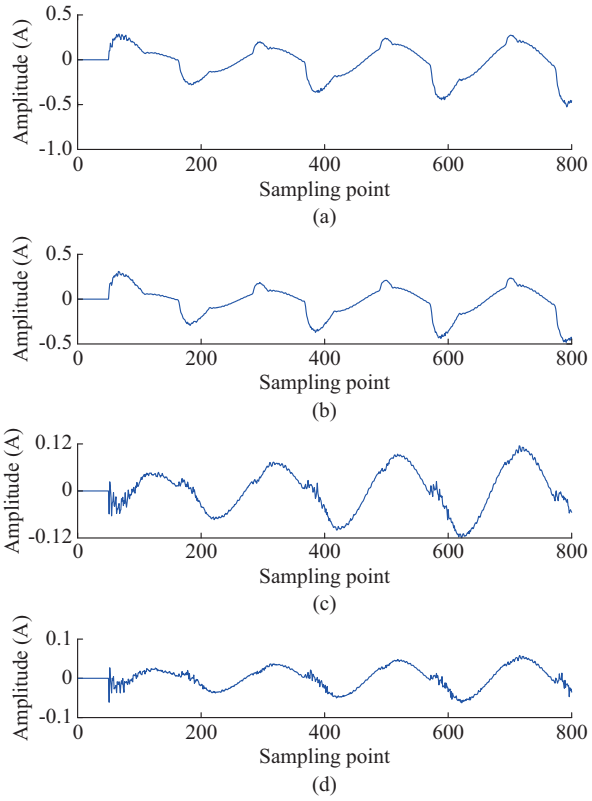

Fig. 3. Zero-sequence current resulting from occurrence of HIF. (a) Point *M*. (b) Point *N*. (c) Point *P*. (d) Point *Q*.

C. Verification of Modified IEEE 33-node Test System

To assess the performance of the proposed method in a multi-node distribution network with DGs, the modified IEEE 33-node test system shown in Supplementary Material A Fig. SA4 was employed.

1) Effects of Different Load Levels

A single-phase-to-ground fault with a 1000 Ω fault resistance is introduced 1 km from endpoint 14 at line section $L_{14,15}$. The performance of the proposed method is evaluated under three different load conditions: normal (100%), light (50%), and heavy (120%).

TABLE IV
SIMULATION RESULTS FOR DIFFERENT HIF CONDITIONS

HIF condition (Ω)	Fault location	Similarity metrics			Faulted line section
		MN	*NP*	*PQ*	
450-550	f_1	0.1708	0.0008	0.0011	*MN*
	f_2	0.0004	0.2201	0.0006	*NP*
	f_3	0.0003	0.0002	0.2358	*PQ*
700-800	f_1	0.1518	0.0002	0.0005	*MN*
	f_2	0.0002	0.1378	0.0004	*NP*
	f_3	0.0005	0.0003	0.1642	*PQ*
900-1000	f_1	0.1466	0.0003	0.0003	*MN*
	f_2	0.0002	0.1168	0.0003	*NP*
	f_3	0.0002	0.0001	0.1496	*PQ*

The test results in Table V demonstrate that the proposed method remains unaffected by load variations and accurately identifies the faulted line section.

TABLE V
SIMULATION RESULTS UNDER DIFFERENT LOAD CONDITIONS

Load condition (%)	Similarity metric			Faulted line section
	$L_{13,14}$	$L_{14,15}$	$L_{15,16}$	
50	0.0005	1.6450	0.0014	$L_{14,15}$
100	0.0032	1.6230	0.0004	$L_{14,15}$
120	0.0064	1.6632	0.0002	$L_{14,15}$

2) Effects of Grounding Types

The performance of the proposed method is evaluated under different grounding types including arc suppression coil grounding, direct grounding, low-resistance grounding (5 Ω), high-resistance grounding (500 Ω), and ungrounded systems. A single-phase-to-ground fault with a 1000 Ω fault resistance is set to be 2 km from endpoint 13 in line section $L_{13,14}$. As shown in Table VI, the proposed method accurately identifies $L_{13,14}$ as the faulted line section, demonstrating its effectiveness across different grounding types.

TABLE VI
RESULTS UNDER DIFFERENT GROUNDING TYPES

Grounding type	Similarity metric			Faulted line section
	$L_{13,14}$	$L_{14,15}$	$L_{15,16}$	
Arc suppression coil	1.5681	0.0004	0.0009	$L_{13,14}$
Direct grounding	1.5819	0.0002	0.0007	$L_{13,14}$
Low-resistance (5 Ω)	1.5696	0.0005	0.0004	$L_{13,14}$
High-resistance (500 Ω)	1.5257	0.0005	0.0005	$L_{13,14}$
Ungrounded	1.5686	0.0004	0.0013	$L_{13,14}$

3) Effects of Network Topologies

The tie-switch states in Supplementary Material A Fig. SA4 are configured to evaluate the performance of the proposed method under different topologies such as meshed and radial networks. A single-phase-to-ground fault with 1000 Ω fault resistance is set to be 1 km from endpoint 15 in line

section $L_{15,16}$.

In Table VII, $S1$-$S5$ represent different tie-switches. A tie-switch state of 1 indicates a closed switch, while a state of 0 indicates an open switch. The first three cases correspond to meshed topologies, whereas the last represents a radial topology. The results in Table VII indicate that the proposed method maintains its performance regardless of the network topology.

TABLE VII
RESULTS UNDER DIFFERENT TOPOLOGIES

Tie-switch state	Similarity metric			Faulted line section
	$L_{13,14}$	$L_{14,15}$	$L_{15,16}$	
$S1=1$, $S2=1$, $S3=1$, $S4=1$, $S5=1$	0.0001	0.0005	2.2430	$L_{15,16}$
$S1=1$, $S2=0$, $S3=1$, $S4=0$, $S5=1$	0.0005	0.0009	2.2287	$L_{15,16}$
$S1=0$, $S2=1$, $S3=0$, $S4=1$, $S5=1$	0.0018	0.0012	1.7057	$L_{15,16}$
$S1=0$, $S2=0$, $S3=0$, $S4=0$, $S5=0$	0.0008	0.0012	1.5783	$L_{15,16}$

D. Resistance to Interference and Comparative Analysis

1) Resistance to Outlier Interference

During data transmission to the decision center, factors such as electromagnetic interference or malicious human manipulation may cause outliers to occur. These outliers can lead to erroneous fault detection results, making it necessary to test the safety of faulted line-section location in such scenarios. A single-phase-to-ground fault is set at f_3 location on feeder 3, situated at line section PQ, with a fault inception angle of 60° and a fault resistance of 1000 Ω. Four consecutive maximum values multiplied by −2.6 and −1.5 are added at detection points M and P as outliers, respectively. Given the scenario with the presence of outliers, a comparative study between the proposed method and several other waveform-similarity-based fault detection methods is conducted, as shown in Table VIII.

TABLE VIII
SIMULATION RESULTS FOR DIFFERENT FAULT DETECTION METHODS UNDER INFLUENCE OF OUTLIERS

Method	Outlier	Similarity metric			Faulted line section	Correctness
		MN	*NP*	*PQ*		
CS	No	0.9830	0.9858	0.1485	*PQ*	Correct
	Yes	−0.4301	−0.1929	−0.3298	*MN*	Incorrect
PCC	No	0.9707	0.9756	0.1248	*PQ*	Correct
	Yes	−0.5034	−0.3299	−0.3396	*MN*	Incorrect
HD	No	0.1181	0.1259	0.8226	*PQ*	Correct
	Yes	2.7367	1.4239	1.3706	*MN*	Incorrect
DTW	No	1.9024	1.9869	9.2202	*PQ*	Correct
	Yes	15.5074	10.4803	12.8622	*MN*	Incorrect
MHD	No	0.0104	0.0125	0.1308	*PQ*	Correct
	Yes	0.2236	0.1208	0.1950	*MN*	Incorrect
KTC	No	0.8710	0.8743	−0.0661	*PQ*	Correct
	Yes	0.5900	0.5900	−0.1678	*PQ*	Correct
Proposed	No	0.0011	0.0011	0.2801	*PQ*	Correct
	Yes	0.0018	0.0019	0.3832	*PQ*	Correct

For comparing the proposed method with state-of-the-art fault detection methods, we chose CS, PCC, HD, DTW, MHD, and KTC based on a thorough survey of current widely-used methods in power system fault detection, especially those based on similarity comparison. These methods have been extensively and frequently employed in a variety of power system contexts, including transformers, DC lines, and transmission lines. The proposed method, which focuses on distribution networks, extends to these areas and ensures that the comparative analysis remains relevant and comprehensive. In the Table VIII, these methods with and without the presence of outliers are evaluated. It is clear from the table that all methods are capable of correctly identifying the faulted line section when no outlier interference is present in the data. Nevertheless, when outliers are encountered, the calculations derived from the CS [31], PCC [32], HD [15], DTW [16], and MHD [21] methods deviate from the anticipated value, inevitably resulting in maloperation. The KTC method in [25] can correctly identify faults even in the presence of outliers. However, compared with the case without outliers, the calculated results deteriorate. This is because although the transformation of outliers into ranks mitigates their adverse effects, they still exist within the dataset. In contrast, the proposed method achieves reliable results by effectively excluding outliers through the utilization of quantiles, thereby preserving its performance.

Table IX provides a comparison of different robust fault detection methods across different aspects. The methods presented in [17]-[20] utilize multi-source information to deal with outliers in decision information but do not directly handle outliers in electrical signals. In contrast, the methods presented in [22]-[25] along with the proposed method primarily focus on managing outliers in electrical signals to obtain reliable decision information and can be utilized as an initial stage for those focusing on dealing with outliers in decision information. Clearly, adeptly handling outliers in the initial stage can effectively alleviate the cost increments and efficiency reductions brought about by the second-stage methods, which invoke multi-source information and delayed handling of outliers. Both the method in [25] and the proposed method effectively deal with outliers in electrical signals without requiring extensive data for model training, as is the case with the learning-based methods in [22]-[24], providing a straightforward solution. As analyzed in Table VIII, it is evident that the proposed method exhibits better robustness compared with the method in [25].

TABLE IX
COMPARISON OF DIFFERENT ROBUST FAULT DETECTION METHODS

Method	Aspect of comparison			
	Multi-source information	Extensive data involved	Dealing with outliers in electrical signals	Dealing with outliers in decision information
[17]-[20]	Yes	No	No	Yes
[22]-[24]	No	Yes	Yes	No
[25] and proposed	No	No	Yes	No

The comparative analysis in Tables VIII and IX highlights the distinct advantage of the proposed method in handling outliers. The NQHD algorithm, in contrast to traditional similarity-based algorithms, exhibits enhanced robustness against outlier interference. This is achieved by strategically using quantiles to effectively disregard outliers. Unlike multi-stage processing methods that rely on additional outlier detection mechanisms or data fusion from multiple sources, the proposed method directly and efficiently removes outliers from electrical signals, thereby improving the efficiency of the fault detection process.

2) Resistance to Noise Interference

The measured current waveforms are usually affected by random noise. To validate the effectiveness of the proposed method in the presence of noise interference, harmonic noise (HN) consisting of the 3rd (150 Hz) and 5th (250 Hz) components scaled to 5% of the root-mean-square (RMS) value of the original current signals, α-stable noise (α-SN) ($\alpha = 1.8$) ensuring a heavy-tailed distribution with intensity adaptively scaled to 5% of the RMS value, and Gaussian noise (GN) with magnitudes of 40, 30, 20, and 10 dB are superimposed onto the signals at both ends of the line section. Upon observing that both the KTC and the proposed method maintain robustness against outliers while the other methods fail, we opt to explore further by comparing these two methods in a subsequent analysis under noisy conditions.

Table X presents the calculated results for both the KTC method and the proposed method under single-phase-to-ground faults, considering the effects of HN and α-SN, as well as different signal-to-noise ratios (SNRs) for GN.

TABLE X
RESULTS FOR KTC METHOD AND PROPOSED METHOD UNDER SINGLE-PHASE-TO-GROUND FAULTS

Method	Noise type	SNR (dB)	Similarity metric			Faulted line section
			MN	*NP*	*PQ*	
KTC	HN		–0.1184	0.4988	0.5706	*MN*
	α-SN		–0.0792	0.5086	0.5200	*MN*
	GN	40	–0.1282	0.5167	0.5739	*MN*
		30	–0.0727	0.3796	0.4563	*MN*
		20	0.0449	0.1869	–0.0384	*PQ*
		10	0.0302	–0.0073	0.1510	*NP*
Proposed	HN		0.3081	0.0027	0.0017	*MN*
	α-SN		0.2901	0.0025	0.0015	*MN*
	GN	40	0.2963	0.0016	0.0009	*MN*
		30	0.2751	0.0010	0.0013	*MN*
		20	0.1636	0.0004	0.0011	*MN*
		10	0.0247	0.0011	0.0019	*MN*

The fault has a fault inception angle of 30° and a fault resistance of 1000 Ω, occurring at fault point f_1. From Table X, both the proposed method and the KTC method correctly identify the faulted line section under HN and α-SN interferences. This is because while HN and α-SN introduce distortions, they do not significantly affect the relative similarity differences between faulted and healthy line sections. The fault-induced zero-sequence current variations remain dominant, ensuring that these noise types do not obscure the underlying fault characteristics. When increasing the noise intensity such as by introducing GN at different SNR levels, the proposed method exhibits superior noise immunity performance compared with the KTC method. The KTC method is susceptible to producing false results when strong noise interference alters the ordering of data points, while the proposed method, based on the original amplitude differences, effectively maintains a low level of changes in amplitude caused by noise interference.

3) Comparison Analysis with Data-driven and Hybrid Methods

Data-driven and hybrid methods, which combine physical models with machine learning or deep learning (DL) techniques, enhance fault detection by leveraging historical data for training. While they can recognize complex fault patterns, they typically require large labeled datasets, which may be sensitive to system variations and often involve high computational costs. Additionally, these methods require periodic retraining to maintain performance in evolving distribution networks, making the real-time deployment challenging. They also struggle with open-set fault diagnosis [33], as they typically only detect faults present in the training data, and outliers altering the sample distribution may lead to misclassification. Furthermore, data-driven and hybrid methods pose security risks, as they are susceptible to adversarial attacks [34] and backdoor attacks [35], which can compromise the reliability of fault detection. In contrast, the proposed method requires no training data, ensures consistent performance in evolving networks, and operates with low computational complexity, enabling real-time deployment while avoiding security risks associated with data-driven methods.

E. Verification Using Practical Field Data

The proposed method is further validated using a set of practical field data recorded on-site in a 10 kV MV distribution network. On September 10, 2020, at 05:35:10, a single-phase-to-ground fault occurred at a line section in the distribution network, and zero-sequence currents were captured by three sensors, forming two line sections. The zero-sequence current waveforms at detection points were sampled at a rate of 6400 Hz. Field fault recording data, presented in Supplementary Material A Fig. SA9, were obtained in COMTRADE format. From Fig. SA9, it is evident that the zero-sequence current waveform exhibits characteristics typical of an HIF, including asymmetry, intermittence, non-linearity, and randomness. The zero-sequence current waveforms on both ends of the line section upstream of the fault point exhibit high similarity, while the waveforms on both ends of the line section where the fault point is located show significant differences. To implement the proposed method, data from a one-quarter cycle after the fault occurrence are utilized, corresponding to 32 sample points for the 50 Hz system with a sampling rate of 6400 Hz. The calculated NQHD values are 0.0017 and 0.7309 for the healthy and faulted line sections, respectively. The results affirm the efficacy of the proposed method when applied to field data.

To address the limited availability of field data and en-

hance real-world applicability, artificial outliers are introduced into the dataset to evaluate the effectiveness of the proposed method under data contamination, with a comparative analysis against existing methods provided in Supplementary Material A Fig. SA10 and Table SAI.

F. Discussion

While the proposed method demonstrates robustness against noise and outlier interference, its performance under cyber-attacks remains a limitation. This is primarily due to its reliance on communication networks for exchanging measurements, making it susceptible to malicious attacks. For instance, an attacker could reverse the direction of electrical signals, leading to misjudgment of faulted line-sections as normal and vice versa, which could result in severe consequences. Existing studies [36]-[38] suggest verification mechanisms, often incorporating DL-based techniques to distinguish between attack and normal states, to enhance the reliability of fault detection under cyber threats. Integrating such mechanisms into the proposed method could improve its resilience against cyber-attacks, which is considered as an important direction for future research.

Regarding sensor failures, partial data losses or measurement errors are handled as outliers. An outlier suppression method reduces the influence of such issues on the accuracy of faulted line-section location. Since the proposed method already incorporates the suppression of outliers, small portions of missing or erroneous data will not impact fault detection accuracy. If a sensor becomes completely unavailable, fault detection for that section will fail. Nevertheless, leveraging neighboring sensors for similarity comparison can still provide detection capability, though this may lead to identification of an expanded fault section. Multi-source fusion-based fault detection methods are more effective in the scenarios where a sensor completely fails, as they integrate multiple sources to compensate for missing data. However, the proposed method excels in situations where a single-source measurement contains outliers, as it can maintain fault detection reliability without relying on multi-source fusion. This makes the proposed method more practical when additional measurement sources are unavailable or costly to implement.

The fault detection and fault location are two pivotal techniques essential for timely identifying and addressing electrical faults. The fault detection method serves to identify the faulted line or line section, ensuring quick actions for fault isolation or warnings and usually constitutes the first response to a fault. Meanwhile, the fault location method focuses on accurately pinpointing the fault position to guarantee that the fault point can be located and repaired promptly.

Typically, fault detection and fault location are two distinct research areas due to the utilization of different methods and data sources. Some studies such as the one presented in [39] have proposed the seamless integration of fault detection and fault location to enhance the efficiency of the fault management system effectively. However, there remain two sequential tasks: fault detection first, followed by fault location. In this paper, fault detection for a line can be accomplished by evaluating the similarity between the sampled current and the reference current. Once fault detection is concluded, the genetic algorithm randomly generates various combinations of fault location and fault resistance, and the corresponding fault current is calculated. Subsequently, the similarity between the sampled current and the calculated fault current is computed. In this context, the higher the similarity, the more the enumerated fault location and fault resistance are considered to be consistent with the actual situation, thereby determining the fault location. This provides a valuable reference for our research, allowing the NQHD algorithm to be extended for simultaneous application in fault detection and location. It can be seamlessly integrated into such systems through the addition of substitutable software, requiring no hardware modifications to the existing system. Given the robust similarity assessment performance of the NQHD algorithm amidst the presence of outliers, it presents certain prospects for practical implementations.

Deploying multiple measuring devices along a feeder involves cost considerations. The proposed method leverages the existing FTU infrastructure in distribution automation (DA) systems, reducing the need for additional installations. A key drawback of the proposed method is its reliance on a relatively dense deployment of FTUs to achieve accurate faulted line-section location. In areas where FTUs are sparse, the proposed method becomes ineffective, necessitating the use of alternative fault detection methods based on limited measurement devices. Such methods have been adopted in both transmission and distribution networks and offer a practical solution in low-infrastructure environments [40] - [42]. However, despite their cost advantages, these methods also face notable limitations. For example, phasor-based methods [40] typically rely on post-fault phasor averaging over several cycles to obtain stable measurements, which inevitably delays fault detection. Additionally, due to signal attenuation over long-distance transmissions, these methods are often only validated under low fault resistance conditions. Traveling wave-based methods [41] utilize transient signals and are resistant to fault resistance, but require ultra-high sampling rates, typically exceeding 1 MHz and up to 100 MHz for short lines, which are beyond the capabilities of current feeder FTUs. AI-based methods [42] can operate with limited measurements and leverage both transient and steady-state data, but they demand extensive training datasets and are highly sensitive to changes in network topology, limiting their applicability in real-world systems. In summary, in the networks where FTUs are sufficiently deployed, the proposed method enables more sensitive, faster, and highly reliable fault detection, especially under high fault resistance and weak fault signal conditions. Limited-sensor-based methods remain valuable in under-instrumented systems and can serve as a complement. By integrating both methods within a hybrid protection framework, utilities can enhance fault detection coverage while balancing cost and performance considerations.

VI. Conclusion

A robust faulted line-section location method based on

NQHD algorithm is proposed to identify faulted and healthy line sections during single-phase-to-ground faults in distribution networks. To effectively disregard extremes or outliers and provide consistent normalized scaling across different scenarios, the NQHD algorithm is applied by employing quantile and robust normalization techniques. The NQHD value is calculated based on zero-sequence currents from both ends of the line section, providing a representation of fault characteristics for single-phase-to-ground faults in faulted and healthy line sections. The effectiveness of the proposed method is evaluated through simulation tests and the analysis of on-site recorded data. The results demonstrate that the proposed method performs well under different fault distances, fault resistances, fault inception angles, and HIFs. The proposed method demonstrates more tolerance to outliers and noise than existing methods, which supports reliable and secure faulted line-section location in distribution networks.

REFERENCES

[1] J. Lin, J. You, M. Guo *et al*., "Amplitude-phase-characteristics-based fault phase detection method for grounding fault in distribution network," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, p. 3527712, Jul. 2024.

[2] K. Liu, S. Zhang, B. Li *et al*., "Flexible grounding system for single-phase to ground faults in distribution networks: a systematic review of developments," *IEEE Transactions on Power Delivery*, vol. 37, no. 3, pp. 1640-1649, Jun. 2022.

[3] L. He, Y. Li, X. Chu *et al*., "Single-phase to ground fault line identification for medium voltage islanded microgrids with neutral ineffectively grounded modes," *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4312-4326, Nov. 2022.

[4] S. K. Pirmani and M. A. Mahmud, "Detection of faulty feeders and phases in resonant grounded power distribution networks for bushfire prone areas using measurements from existing field devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, p. 9001016, Jan. 2024.

[5] Z. Li, Y. Ye, X. Ma *et al*., "Single-phase-to-ground fault section location in flexible resonant grounding distribution networks using soft open points," *International Journal of Electrical Power & Energy Systems*, vol. 122, p. 106198, Nov. 2020.

[6] F. Shi, L. Zhang, H. Zhang *et al*., "Diagnosis of the single phase-to-ground fault in distribution network based on feature extraction and transformation from the waveforms," *IET Generation, Transmission & Distribution*, vol. 14, no. 25, pp. 6079-6086, Jan. 2020.

[7] M. Guo, J. Gao, X. Shao *et al*., "Location of single-line-to-ground fault using 1-D convolutional neural network and waveform concatenation in resonant grounding distribution systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, p. 3501009, Jan. 2021.

[8] Z. Li, S. Lin, M. Guo *et al*., "A decentralized fault section location method using autoencoder and feature fusion in resonant grounding distribution systems," *IEEE Systems Journal*, vol. 16, no. 4, pp. 5698-5707, Dec. 2022.

[9] J. Li, Y. Liu, C. Li *et al*., "An FTU-based method for locating single-phase high-impedance faults using transient zero-sequence admittance in resonant grounding systems," *IEEE Transactions on Power Delivery*, vol. 37, no. 2, pp. 913-922, Apr. 2022.

[10] Y. Chen, J. Yin, Z. Li *et al*., "Location for single-phase grounding fault in distribution network based on equivalent admittance distortion rate," *IET Generation, Transmission & Distribution*, vol. 15, no. 11, pp. 1716-1729, Feb. 2021.

[11] J. Fang, Y. Yan, H. Zhang *et al*., "Research on fault section location method of distribution networks with arc suppression coil grounding based on energy leakage function," *IET Generation, Transmission & Distribution*, vol. 14, no. 25, pp. 6097-6106, Dec. 2020.

[12] P. Bountouris, H. Guo, D. Tzelepis *et al*., "MV faulted section location in distribution systems based on unsynchronized LV measurements," *International Journal of Electrical Power & Energy Systems*, vol. 119, p. 105882, Jul. 2020.

[13] J. Li, G. Wang, D. Zeng *et al*., "High-impedance ground faulted line-section location method for a resonant grounding system based on the zero-sequence current's declining periodic component," *International Journal of Electrical Power & Energy Systems*, vol. 119, p. 105910, Jul. 2020.

[14] Q. Wang, T. Jin, M. A. Mohamed *et al*., "A novel linear optimization method for section location of single-phase ground faults in neutral noneffectively grounded systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, p. 3513410, Jan. 2021.

[15] Z. Li, J. Wan, P. Wang *et al*., "A novel fault section locating method based on distance matching degree in distribution network," *Protection and Control of Modern Power Systems*, vol. 6, p. 20, Apr. 2021.

[16] D. Xiao, T. He, R. Xiao *et al*., "Segment location for single-phase-to-ground fault in neutral non-effectively grounded system based on distributed electric-field measurement," *Electric Power Systems Research*, vol. 184, p. 106321, Jul. 2020.

[17] K. Sun, Q. Chen, and Z. Gao, "An automatic faulted line section location method for electric power distribution systems based on multisource information," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1542-1551, Aug. 2016.

[18] I. Kiaei and S. Lotfifard, "Fault section identification in smart distribution systems using multi-source data based on fuzzy petri nets," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 74-83, Jan. 2020.

[19] Q. Wang, T. Jin, and M. A. Mohamed, "A fast and robust fault section location method for power distribution systems considering multisource information," *IEEE Systems Journal*, vol. 16, no. 2, pp. 1954-1964, Jun. 2022.

[20] Z. Tang, P. Zhang, R. Liang *et al*., "Multi-source data-cooperated neutral low-resistance grounding cable grid faulty segment identification," *IEEE Transactions on Power Systems*, vol. 37, no. 2, pp. 1413-1424, Mar. 2022.

[21] J. Zhao, G. Zhang, X. Shi *et al*., "Fault section location method using wide area centralized structure based on modified Hausdorff distance," in *Proceedings of 2017 IEEE Conference on Energy Internet and Energy System Integration*, Beijing, China, Nov. 2017, pp. 1-6.

[22] A. Ameli, A. Ayad, E. F. El-Saadany *et al*., "A learning-based framework for detecting cyber-attacks against line current differential relays," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2274-2286, Aug. 2021.

[23] Y. M. Khaw, A. A. Jahromi, M. F. M. Arani *et al*., "A deep learning-based cyberattack detection system for transmission protective relays," *IEEE Transactions on Smart Grid*, vol. 12, no. 3, pp. 2554-2565, May 2021.

[24] A. M. Saber, A. Youssef, D. Svetinovic *et al*., "Anomaly-based detection of cyberattacks on line current differential relays," *IEEE Transactions on Smart Grid*, vol. 13, no. 6, pp. 4787-4800, Nov. 2022.

[25] G. Zhang, X. Tong, Q. Hong *et al*., "Waveform similarity-based robust pilot protection for transmission lines," *IEEE Transactions on Power Delivery*, vol. 37, no. 3, pp. 1856-1865, Jun. 2022.

[26] M. Deng, Z. Li, and X. Chen, "Extended Hausdorff distance for spatial objects in GIS," *International Journal of Geographical Information Science*, vol. 21, no. 4, pp. 459-475, Apr. 2007.

[27] M. Elsisi, J. Yu, C. Lai *et al*., "A drone-assisted deep learning-based IoT system for monitoring ship emissions in ports considering adversarial attacks," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, p. 9506111, Jan. 2024.

[28] J. Yuan, Y. Hu, Y. Liang *et al*., "Faulty feeder detection for single line-to-ground fault in distribution networks with DGs based on correlation analysis and harmonics energy," *IEEE Transactions on Power Delivery*, vol. 38, no. 2, pp. 1020-1029, Apr. 2023.

[29] T. Jin, F. Zhuo, and M. A. Mohamed, "A novel approach based on CEEMDAN to select the faulty feeder in neutral resonant grounded distribution systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4712-4721, Jul. 2020.

[30] X. Wang J. Gao, X. Wei *et al*., "High impedance fault detection method based on variational mode decomposition and Teager-Kaiser energy operators for distribution network," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6041-6054, Nov. 2019.

[31] A. Saadat, R. A. Hooshmand, A. Kiyoumarsi *et al*., "Voltage sag source location in distribution networks with DGs using cosine similarity," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, p. 3529610, Jan. 2022.

[32] K. Jia, Y. Li, Y. Fang *et al*., "Transient current similarity based protection for wind farm transmission lines," *Applied Energy*, vol. 225, pp. 42-51, Sept. 2018.

[33] P. Peng, J. Lu, T. Xie *et al*., "Open-set fault diagnosis via supervised contrastive learning with negative out-of-distribution data augmentation," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 3, pp.

2463-2473, Mar. 2023.
[34] C. Ren, X. Du, Y. Xu *et al*., "Vulnerability analysis, robustness verification, and mitigation strategy for machine learning-based power system stability assessment model under adversarial examples," *IEEE Transactions on Smart Grid*, vol. 13, no. 2, pp. 1622-1632, Mar. 2022.
[35] Z. Yao, H. Zhang, Y. Guo *et al*., "Reverse backdoor distillation: towards online backdoor attack detection for deep neural network models," *IEEE Transactions on Dependable and Secure Computing*, vol. 21, no. 6, pp. 5098-5111, Nov.-Dec. 2024.
[36] A. Ameli, A. Ayad, E. F. El-Saadany *et al*., "A learning-based framework for detecting cyber-attacks against line current differential relays," *IEEE Transactions on Power Delivery*, vol. 36, no. 4, pp. 2274-2286, Aug. 2021.
[37] L. Wang, K. Zhang, W. Bi *et al*., "Indirect coordinated attack against relay via load-side power electronics and its defense strategy," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 4, pp. 5112-5124, Apr. 2024
[38] A. M. Saber, A. Youssef, D. Svetinovic *et al*., "Cyber-immune line current differential relays," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 3597-3608, Mar. 2024.
[39] L. Kong and H. Nian, "Fault detection and location method for mesh-type DC microgrid using Pearson correlation coefficient," *IEEE Transactions on Power Delivery*, vol. 36, no. 3, pp. 1428-1439, Jun. 2021.
[40] G. Zhang, G. Xiao, X. Liu *et al*., "A robust wide-area fault location method for transmission lines with uncertainties in measurements and network parameters," *IEEE Transactions on Power Delivery*, vol. 39, no. 1, pp. 687-690, Feb. 2024.
[41] C. Galvez and A. Abur, "Fault location in active distribution networks containing distributed energy resources (DERs)," *IEEE Transactions on Power Delivery*, vol. 36, no. 5, pp. 3128-3139, Oct. 2021.
[42] W. Li, D. Deka, M. Chertkov *et al*., "Real-time faulted line localization and PMU placement in power systems through convolutional neural networks," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4640-4651, Nov. 2019.

Guangxiao Zhang received the Ph.D. degree in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2022. He was a Visiting Researcher in University of Strathclyde, Glasgow, U.K., from April 2021 to July 2022. He is currently a Research Fellow with the Nanyang Technological University and Singapore-ETH Centre, Singapore. His research interests include application of wide-area measurement, state estimation, and artificial intelligence in power system protection.

Gaoxi Xiao received the Ph.D. degree in computing from The Hong Kong Polytechnic University, Hong Kong, China, in 1998. He was a Postdoctoral Research Fellow with Polytechnic University, Brooklyn, USA, in 1999, and a Visiting Scientist with the University of Texas at Dallas, Dallas, USA, from 1999 to 2001. In 2001, he joined the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently an Associate Professor. His research interests include complex system and complex network, communication network, smart grid, and system resilience and risk management.

Xinghua Liu received the B.S. degree in mathematics from Jilin University, Changchun, China, in 2009, and the Ph.D. degree in control science and engineering from University of Science and Technology of China, Hefei, China, in 2014. From 2014 to 2015, he was a Visiting Fellow with Royal Melbourne Institute of Technology (RMIT) University, Melbourne, Australia. From 2015 to 2018, he was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He is currently a Professor with the Department of Power Grid Information and Control Engineering, School of Electrical Engineering, Xi'an University of Technology, Xi'an, China. His research interests include state estimation and control, intelligent system, power system, autonomous vehicle, and cyber-physical system.

Yan Xu received the B.E. and M.E degrees from South China University of Technology, Guangzhou, China, in 2008 and 2011, respectively, and the Ph. D. degree from University of Newcastle, Newcastle, Australia, in 2013. After the postdoctoral training with University of Sydney Postdoctoral Fellowship, he joined Nanyang Technological University, Singapore, with the Nanyang Assistant Professorship. He was promoted to Associate Professor in 2021 and appointed as the Cham Tao Soon Professor in Engineering in 2024. His research interests include power system stability and control, microgrid, and data analytics for smart grid.

Peng Wang received the B.Sc. degree in electronic engineering from Xi'an Jiaotong University, Xi'an, China, in 1978, the first M.Sc. degree in electrical engineering from the Taiyuan University of Technology, Taiyuan, China, in 1987, and the second M.Sc. and Ph.D. degrees in electrical engineering from University of Saskatchewan, Saskatoon, Canada, in 1995 and 1998, respectively. He is currently a Full Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His current research interests include power system planning and operation, renewable energy planning, solar or electricity conversion system, and power system reliability.