A Data-Driven Approach for High-Impedance Fault Localization in Distribution Systems

Preprint

Yuqi Zhou,[1] Yuqing Dong,[2] and Rui Yang[1]

1 National Renewable Energy Laboratory
2 University of Tennessee, Knoxville

To be presented at the 2024 IEEE Power and Energy Society General Meeting
Seattle, Washington
July 21–25, 2024

A Data-Driven Approach for High-Impedance Fault Localization in Distribution Systems

Preprint

Yuqi Zhou,[1] Yuqing Dong,[2] and Rui Yang[1]

1 National Renewable Energy Laboratory
2 University of Tennessee, Knoxville

NOTICE

A Data-Driven Approach for High-Impedance Fault Localization in Distribution Systems

Yuqi Zhou*, Yuqing Dong†, and Rui Yang*
*National Renewable Energy Laboratory, Golden, CO, USA
†The University of Tennessee, Knoxville, TN, USA

Abstract—Accurate and quick identification of high-impedance faults (HIFs) is critical for the reliable operation of distribution systems. Unlike other faults in power grids, HIFs are very difficult to detect by conventional overcurrent relays due to the low fault current. Although HIFs can be affected by various factors, the voltage–current characteristics can substantially imply how the system responds to the disturbance and thus provides opportunities to effectively localize HIFs. In this work, we propose a data-driven approach for the identification of HIF events. To tackle the nonlinearity of the voltage–current trajectory, first, we formulate optimization problems to approximate the trajectory with piecewise functions. Then we collect the function features of all segments as inputs and use the support vector machine approach to efficiently identify HIFs at different locations. Numerical studies on the IEEE 123-node test feeder demonstrate the validity and accuracy of the proposed approach for real-time HIF identification.

Index Terms—High-impedance fault, voltage-current trajectory, piecewise approximation, machine learning, support vector machine, explainable artificial intelligence.

I. Introduction

High-impedance faults (HIFs) occur when an energized conductor comes into contact with a high-impedance medium (e.g., sod, asphalt, tree limb). As a consequence of the low-fault current therein, HIFs are very difficult to detect by conventional overcurrent protection devices. In addition, if left unattended, such faults could pose fire risks and even cause electric shock, endangering human life. Therefore, the timely localization of HIFs in distribution systems is crucial for safe and reliable grid operation.

Because the impedance of HIFs is highly uncertain and can fluctuate over time, it is very difficult to deterministically model the dynamic process of arcing. Meanwhile, limited monitoring devices in the distribution system make it challenging to pinpoint the HIF locations. Due to these restrictions, machine learning approaches are uniquely positioned for

HIF localization. Early work used neural network algorithms [1], [2] to identify the existence of HIFs in the distribution system and achieve satisfactory performance. More recently, autoencoder-based approaches [3], [4] have been explored for HIF learning with partially labeled data sets. Albeit sufficiently accurate, in general, these approaches lack explainability and interpretability on the detection model. In contrast, algorithms like support vector machines (SVMs) are suitable for producing more explainable solutions under varying operating conditions. Earlier efforts have been made to localize faults using SVM-based approaches [5], [6], but thus far the dynamic electrical characteristics are not suitably integrated into the model.

To address this issue, we propose an efficient and explainable SVM approach for localizing HIFs in the distribution system. For dimension reduction, we formulate tractable optimization problems to approximate the voltage-current (V-I) trajectory with piecewise functions. Notably, the proposed optimization method can be easily modified to incorporate different approximation requirements. We collect the properties of the piecewise functions (e.g., slope rates in the piecewise linear approximation) and directly use them as the inputs to the SVM model. Meanwhile, the learning outputs are the HIF locations (e.g., bus number). The approximated functions can be reckoned as a unique projection from the original nonlinear trajectory, and the properties of these functions are indicative of the electrical characteristics of HIFs during arc quenching and restriking. Hence, the simplified learning inputs not only nicely conserve the unique properties of HIF arcing but also effectively reduce the dimensionality of the measurements from the original nonlinear trajectory. Admittedly, even for faults at the same location, their V-I trajectories might be different due to operation conditions and measurement noise. However, the rationale behind our proposed approach is that the electrical characteristics of faults at the same location share similarities, and these are reflected in the properties of the piecewise approximation function and can be effectively captured by the SVM algorithm.

This paper is organized as follows. Section II introduces the dynamic model of HIFs. Section III presents the optimization formulation to obtain the optimal piecewise approximation. Section IV explains the SVM algorithm and how to use approximation results for the HIF identification task. Section V shows numerical simulations on the IEEE 123-node test feeder, and Section VI concludes the paper.

II. HIGH-IMPEDANCE FAULT MODEL

Accurately modeling HIFs is crucial for the efficient identification of these events. Although various arc models of HIFs have been studied, the model using two antiparallel DC voltage sources and diodes [7], [8] is well-known to be able to effectively simulate the dynamic activities of arcs during HIF events, as shown in Fig. 1.
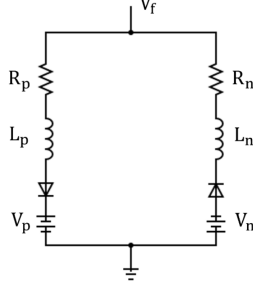

Fig. 1: HIF modeled with two antiparallel diodes

When $V_f > V_p$, the current flows from the source to the ground. On the other hand, when $V_f < V_n$, the negative cycle occurs, during which the current flows back to the source. As $V_n < V_f < V_p$, there is no current in the HIF circuit, which represents the arc quenching. When the voltage V_f falls within these ranges, the electrical characteristics at the fault location as well as the rest of system are generally different. As a matter of fact, the voltage and current at the distribution feederhead approximately feature a piecewise trajectory [9] during the HIF events. The occurrences of HIFs at different system locations will cause the voltage-current (V-I) trajectories at the feederhead to differ as well. The uniqueness of the voltage-current (V-I) trajectory has been recognized to be very helpful for HIF detection and localization tasks. Nonetheless, the randomness and nonlinearity of the HIF process make it extremely challenging to attain a closed-form expression for the V-I trajectory. This motivates us to explore more tractable and efficient algorithms for the initial approximation of the V-I trajectory before integrating it into the learning framework.

III. PIECEWISE APPROXIMATION OF V-I TRAJECTORY

Providing tractable yet accurate approximations of the V-I trajectory is crucial for later identifying the fault locations. In this section, we present optimization formulations for finding the optimal piecewise linear and piecewise quadratic approximations, both of which can be efficiently handled by mathematical solvers. Notably, the problems are shaped to provide continuous approximations over pieces instead of optimizing each one individually, which is more logical and explainable especially for the V-I trajectory approximation.

Remark 1 (Segmentation for the approximation). *In this paper, we approximate the nonlinear trajectory with three segments as an example, because the V-I characteristics of HIFs roughly feature a piecewise function composed of three pieces (see, e.g., [7], [9]). Nonetheless, the formulation can be easily adjusted to incorporate more pieces.*
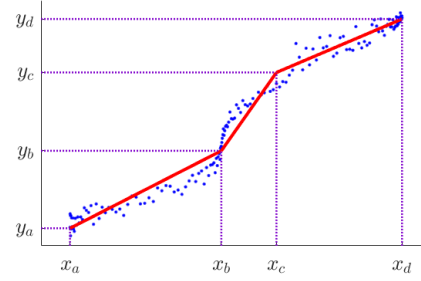

Fig. 2: Illustration of the piecewise linear approximation

A. Piecewise Linear Approximation

The piecewise linear approximation is a regression model that aims to approximate different pieces of the trajectory using affine functions (see Fig. 2). Without loss of generality, we assume that the breakpoints x_a, x_b, x_c, x_d for the linear approximation are known parameters. Our goal is to find three affine functions, as follows, that optimally fit the data points (I_i, V_i) from the trajectory:

$$y_k = s_k x + h_k, \quad \forall k = 1, 2, 3. \tag{1}$$

The slope rates for these affine functions are $s_1 = \frac{y_b - y_a}{x_b - x_a}$, $s_2 = \frac{y_c - y_b}{x_c - x_b}$, and $s_3 = \frac{y_d - y_c}{x_d - x_c}$, respectively. Since two points are sufficient to uniquely determine an affine function and the x-values for the breakpoints are given, the problem can be simplified to finding the optimal y_a, y_b, y_c, y_d values. To this end, the piecewise linear approximation can be formulated as an optimization problem to minimize the total distance between the data points and the fitting affine functions:

$$\min \quad \sum_{k=1}^{3} \sum_{i=1}^{N_k} \left(s_k I_i + y_k^* - s_k x_k^* - V_i \right)^2 \tag{2a}$$

$$\text{s.t.} \quad \underline{y}_\zeta \le y_\zeta \le \overline{y}_\zeta, \quad \zeta = a, b, c, d. \tag{2b}$$

where we use N_k to denote the number of data points in the k-th segment. The fitting distance for piece k is equal to $\sum_{i=1}^{N_k} \left(s_k I_i + h_k - V_i \right)^2$, and substituting $h_k = y_k^* - s_k x_k^*$ yields the expression in objective function (2a), where (x_k^*, y_k^*) is the breakpoint in the k-th piece function. For instance, when $k = 1$, (x_k^*, y_k^*) can be either (x_a, y_a) or (x_b, y_b). The limits of decision variables are enforced in constraints (2b). Note that these constraints are provided in the optimization formulation as a preventive measure for data anomalies, but they are not strictly necessary in general. Because I_i, V_i, x_k are known constants and slope rates s_k are linear functions of y_a, y_b, y_c, y_d, problem (2) constitutes a linear least squares (LS) problem, which can be solved very efficiently.

Specifically, the linear coefficients of $y = \begin{bmatrix} y_a & y_b & y_c & y_d \end{bmatrix}^\mathsf{T}$ for each piece can be stacked and given in matrix form:

$$A_1 = \begin{bmatrix} v_1^1 & w_1^1 & 0 & 0 \\ & \cdots & & \\ v_1^i & w_1^i & 0 & 0 \\ & \vdots & & \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & v_2^1 & w_2^1 & 0 \\ & \cdots & & \\ 0 & v_2^i & w_2^i & 0 \\ & \vdots & & \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 0 & v_3^1 & w_3^1 \\ & \cdots & & \\ 0 & 0 & v_3^i & w_3^i \\ & \vdots & & \end{bmatrix}$$

where $v_k^i, w_k^i, \forall i \in \{1, 2, \cdots, N_k\}$ represent nonzero coefficients corresponding to the i-th sample in piece k, and can be explicitly given by:

$$v_1^i = \frac{x_b - I_i}{x_b - x_a}, v_2^i = \frac{x_c - I_i}{x_c - x_b}, v_3^i = \frac{x_d - I_i}{x_d - x_c} \quad (3a)$$

$$w_1^i = \frac{I_i - x_a}{x_b - x_a}, w_2^i = \frac{I_i - x_b}{x_c - x_b}, w_3^i = \frac{I_i - x_c}{x_d - x_c} \quad (3b)$$

Concatenation of coefficients and constants in (2) leads to:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}, \quad b = \begin{bmatrix} V_1 \\ \vdots \\ V_N \end{bmatrix} = V \quad (4)$$

To this end, the closed-form solution to optimization problem (2) immediately follows:

$$y = (A^{\mathsf{T}} A)^{-1} A^{\mathsf{T}} b \quad (5)$$

Overall, the proposed LS method is very efficient in handling noisy data and outliers, and can be conveniently modified to accommodate more segments as needed.

B. Piecewise Quadratic Approximation

To improve the fitting performance, one can further consider piecewise quadratic approximation (see Fig. 3). The goal then becomes finding three optimal quadratic functions:

$$y_k = m_k x^2 + n_k x + h_k, \quad \forall k = 1, 2, 3. \quad (6)$$

where m_k, n_k, h_k denote quadratic, linear, and constant coefficients, respectively. Different from the linear approximation, uniquely determining a quadratic function requires at least three points. Therefore, an additional interpolation point needs to be introduced for each piece. For easier derivation of the coefficients m, n later, one can pick the midpoint of each piece for interpolation. Hence, the problem can be formulated as:

$$\min \quad \sum_{k=1}^{3} \sum_{i=1}^{N_k} \left(m_k I_i^2 + n_k I_i + y_k^* - m_k x_k^{*2} - n_k x_k^* - V_i \right)^2 \quad (7a)$$

$$\text{s.t.} \quad \underline{y}_\zeta \leq y_\zeta \leq \overline{y}_\zeta, \quad \zeta = a, b, c, d, e, f, g. \quad (7b)$$

where the number of decision variables is increased to include additional interpolation points x_b, x_d and x_f (see Fig. 3). Since the midpoint of each piece is selected as the interpolation point, the quadratic and linear coefficients m, n can be conveniently derived. For example, for the first piece $k = 1$, calculating $(y_c - y_b) - (y_b - y_a)$ cancels out the linear and constant coefficients, and thus:

$$m_1 = \frac{y_c - 2y_b + y_a}{x_c^2 - 2x_b^2 + x_a^2} \quad (8)$$

Further, combining this equation with $y_b - y_a = m_1(x_b^2 - x_a^2) + n_1(x_b - x_a)$ gives us the linear coefficient:

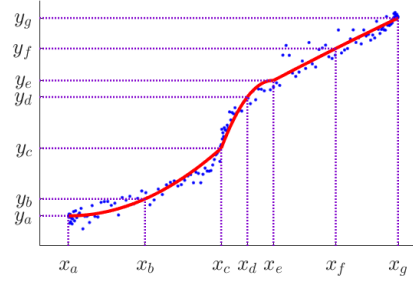$$n_1 = \frac{y_b - y_a - m_1(x_b^2 - x_a^2)}{x_b - x_a} \quad (9)$$

Fig. 3: Illustration of the piecewise quadratic approximation

Similarly, the coefficients m, n for other pieces can be obtained. Note that both quadratic and linear coefficients are linear functions of y, hence problem (7) is also a linear LS problem that is efficiently solvable (similar to (4)-(5)). Of course, the quadratic model provides superior fitting performance because the linear function is just a special case of the quadratic function. However, as higher-order modeling inevitably introduces more variables, a slightly higher computational complexity will incur as a trade-off.

Remark 2 (On parametric continuity). *Constraints on first-order parametric continuity are not enforced for the piecewise approximation because there is also no such guarantee on the V-I trajectory during the arc quenching and restriking.*

IV. SUPPORT VECTOR MACHINE APPROACH

The approximation algorithms introduced earlier simplify the nonlinear V-I trajectory of HIFs to piecewise functions (linear, quadratic), which makes it easier to fit into the machine learning framework. Specifically, the approximated piecewise functions can effectively reflect the system's electrical characteristics during the arc restriking and quenching. Therefore, a collection of features (e.g., slope rates) of these approximated functions can serve as the learning inputs to identify the HIF locations. Compared with other supervised learning algorithms such as neural networks, the SVM approach delivers more explainable and interpretable solutions for HIF identification.

SVM is an efficient supervised learning algorithm used for classification and regression tasks. The algorithm seeks to find a set of hyperplanes, $w^{\mathsf{T}} x - b = 1$ and $w^{\mathsf{T}} x - b = -1$ (e.g., linear SVM) that maximize the separation margin $\frac{2}{||w||}$, which can be formulated as an optimization problem:

$$\min \quad ||w||^2 \quad (10a)$$

$$\text{s.t.} \quad y_i(w^{\mathsf{T}} x_i - b) \geq 1, \forall i \quad (10b)$$

Going beyond the simple two-class linear model, advances in SVM algorithms in the past few decades have enabled multiclass SVM to be easily constructed. Furthermore, by applying the well-known kernel trick [10], it is also possible to perform nonlinear classification.

When SVM is used for the HIF identification task, the learning output y is the fault location label (e.g., bus number), and we propose utilizing the features from the approximation
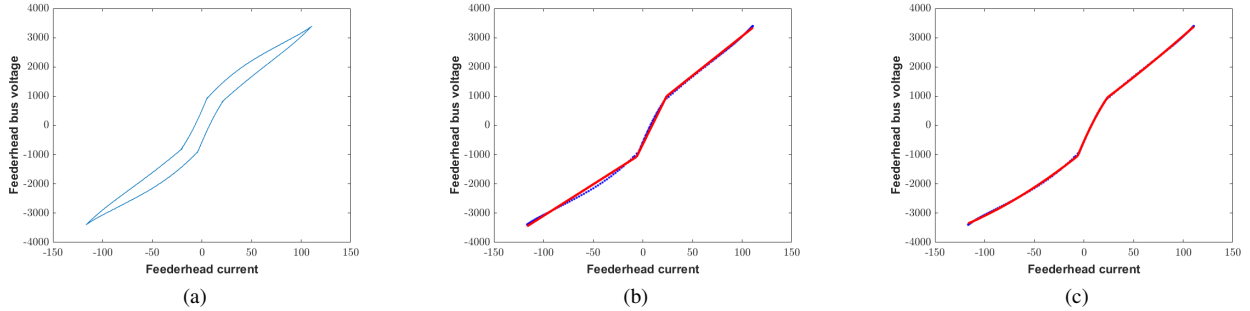
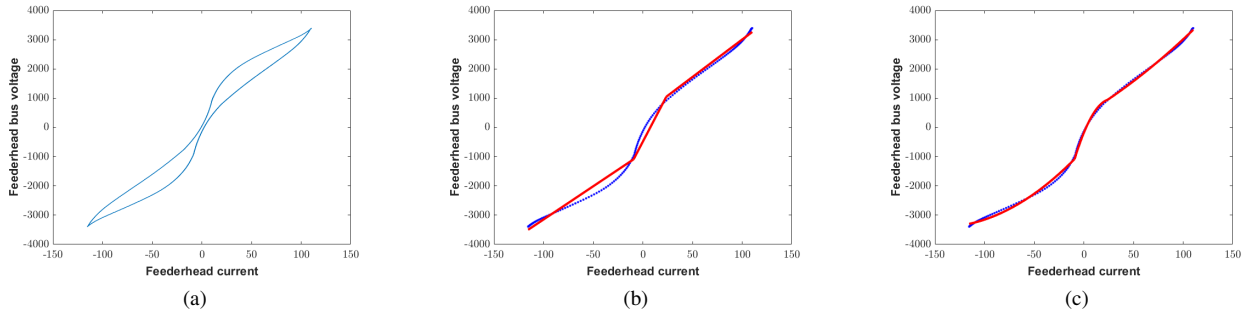Fig. 4: The (a) V-I trajectory, (b) linear approximation, and (c) quadratic approximation, for an HIF at Bus 7

Fig. 5: The (a) V-I trajectory, (b) linear approximation, and (c) quadratic approximation, for an HIF at Bus 64

functions as the input x. Specifically, for the piecewise linear approximation, the SVM input can be constructed as:

$$x_{\mathcal{L}} = \{s_1, s_2, s_3\} \tag{11}$$

which consists of slope rates for all segments in the piecewise linear function. Since quadratic functions contain more information than linear functions, SVM inputs under quadratic approximation will be lifted to a higher-dimensional space:

$$x_{\mathcal{Q}} = \{(m_1, n_1), (m_2, n_2), (m_3, n_3)\} \tag{12}$$

which include both quadratic and linear coefficients for each piece of the function. Note that for both $x_{\mathcal{L}}$ and $x_{\mathcal{Q}}$, the constant coefficient h_k is not included. This is because for the V-I trajectory, the "shape" (e.g., y', y'') in general is more meaningful than the "height" (e.g., constant h) as the former substantially reflects the change in the system's electrical characteristics over the course of HIF. The rationale behind using the approximation function features for the SVM algorithm is twofold. First, as the V-I trajectories for HIFs at different locations are generally different, their unique projections to emblematic piecewise functions are expected to be different as well. Second, thanks to the piecewise approximation, the size of the learning input is greatly reduced to the order of k, which offers a more efficient and explainable learning model.

V. NUMERICAL RESULTS

This section validates the performance of the piecewise approximation and SVM algorithm for HIF identification tasks on the IEEE 123-node test feeder. The test feeder is simulated in EMTP-ATP (ElectroMagnetic Transients Program, version Alternative Transients Program) with 14 rooftop photovoltaic (PV) units incorporated. In this model, the loading condition

($\{0.4, 1.0\}$ per unit) and PV capacity ($\{0.4, 0.6, 0.8, 1.0\}$ per unit) are varied for different steady-state scenarios. The HIF simulations span across 66 locations in total, with 63 internal nodes as well as 3 adjacent nodes outside the feeder. The generated data are sampled at 20 kHz and posted open source to the public [11]. The simulations are performed on a regular laptop with Intel® CPU @ 2.60 GHz and 16 GB of RAM. The piecewise approximation is solved using CPLEX, and SVM algorithms are implemented using the MATLAB simulator.

In this work, we primarily utilize the feederhead voltage and current measurements to validate the HIF identification algorithm. The feederhead V-I trajectories for two HIFs are shown in Fig. 4a and 5a. As the validated trajectory displays partial symmetry, we preprocess the data to extract only the lower half of the trajectory for the approximation and the SVM algorithms later. The piecewise linear and quadratic approximation results are demonstrated in Fig. 4b (5b) and Fig. 4c (5c), respectively. Clearly, the piecewise quadratic model provides a more accurate approximation. However, as the linear approximation offers a lower-dimensional SVM input, $x_{\mathcal{L}}$, this makes it a favorable choice for us, especially for the visualization of the SVM algorithm. Additionally, the slope rate for each piece of the trajectory under the linear approximation implies the "equivalent impedance" from the feederhead, according to Ohm's law, which can be also beneficial for our analysis. For these reasons, we will use the results from the piecewise linear approximation to demonstrate the learning performance. Nevertheless, one can always raise the input dimension to switch to the quadratic model.

The actual V-I trajectory might exhibit non-smoothness due to measurement errors, and thus we add noise proportional to the nominal measurement value to simulate this effect.
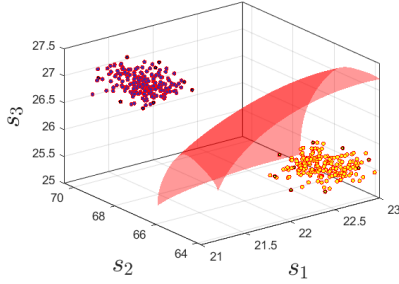
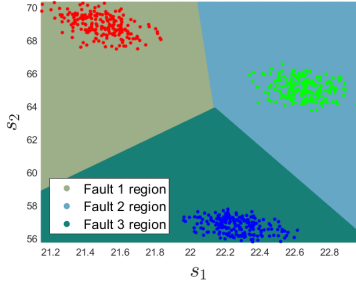Fig. 6: SVM on HIFs under the three-dimensional space

Fig. 7: SVM on HIFs under the two-dimensional space

The presence of noise in trajectories produces different approximation results and correspondingly leads to dispersed SVM inputs $x_{\mathcal{L}} = \{s_1, s_2, s_3\}$ in the three-dimensional space. The classification of two faults at buses 7 and 64 using the Gaussian kernel as an example is shown in Fig. 6. In fact, the data points are already linearly separable because the two HIFs are located far away from each other. The hyperplane in red distinctly classifies these two faults using the slope rates of V-I trajectories. Despite the similarity presented in the approximation results in Fig. 4 and Fig. 5, they are clearly identified under three-dimensional space. For the SVM under multiple HIFs, we can project the input to a two-dimensional space by keeping only two slope rates (s_1 and s_2). The performance of SVM with a linear kernel on three different HIFs is demonstrated in Fig. 7. These faults are simulated at buses 7, 64, and 82, respectively. Because they are located in different areas of the system, even with partial inputs $\{s_1, s_2\}$, the linear SVM can efficiently identify these HIFs.

In contrast, HIFs located in the same area are generally more difficult to distinguish due to their similar network topology and electrical characteristics. To that end, we also validate the identification performance on six HIFs that are located on the same branch of the system. These faults, based on their distance to the feederhead, are at buses 18, 21, 23, 25, 28, and 29 (from nearest to farthest), respectively. The results of the SVM with a polynomial kernel under two-dimensional space are shown in Fig. 8. Because these faults are adjacent, the slope rates $\{s_1, s_2\}$ of their V-I trajectories are relatively close. The regions corresponding to distinct faults are colored differently, as denoted in the figure legend. Interestingly, as the fault location moves farther from the feederhead, the slope rate, s_1, which corresponds to the arc restriking phase of the HIF, gets bigger. This is mainly because when the fault is
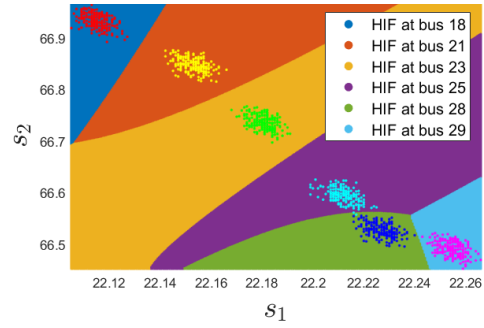

Fig. 8: SVM on HIFs located on the same branch

located deeper on that branch, fewer system components will be disconnected behind that fault, and therefore the equivalent impedance, $Z = V/I$ (or slope rate s_1), is bigger. Then again, the results have demonstrated that compared with area-based HIF identification, bus-level localization of HIFs in small areas is possible but more challenging in general.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we propose an explainable data-driven approach to localize HIFs efficiently. To handle the nonlinearity of the HIF, we approximate the V-I trajectory with piecewise linear (or quadratic) functions. Following that, we feed the simplified function features as inputs and utilize the SVM method for HIF identification. Numerical tests on the IEEE 123-bus distribution system demonstrate the validity and efficiency of the proposed algorithm. Further work includes higher-order approximation and learning algorithms under system uncertainties (e.g., renewable resources, topology change).

REFERENCES

[1] T. Sirojan, S. Lu, B. Phung, D. Zhang, and E. Ambikairajah, "Sustainable deep learning at grid edge for real-time high impedance fault detection," *IEEE Transactions on Sustainable Computing*, vol. 7, no. 2, pp. 346–357, 2018.

[2] S. Wang and P. Dehghanian, "On the use of artificial intelligence for high impedance fault detection and electrical safety," *IEEE Transactions on Industry Applications*, vol. 56, no. 6, pp. 7208–7216, 2020.

[3] K. Rai, F. Hojatpanah, F. Badrkhani Ajaei, and K. Grolinger, "Deep learning for high-impedance fault detection: Convolutional autoencoders," *Energies*, vol. 14, no. 12, 2021.

[4] W. Li and D. Deka, "Physics-informed learning for high impedance faults detection," in *2021 IEEE Madrid PowerTech*, pp. 1–6.

[5] M. Sarwar, F. Mehmood, M. Abid, A. Q. Khan, S. T. Gul, and A. S. Khan, "High impedance fault detection and isolation in power distribution networks using support vector machines," *Journal of King Saud University-Engineering Sciences*, vol. 32, no. 8, pp. 524–535, 2020.

[6] A. S. Zamzam and J. Wang, "Hierarchical data-driven protection for microgrids with 100% renewables," in *2023 IEEE Energy Conversion Congress and Exposition (ECCE)*.

[7] Q. Cui, K. El-Arroudi, and Y. Weng, "A feature selection method for high impedance fault detection," *IEEE Transactions on Power Delivery*, vol. 34, no. 3, pp. 1203–1215, 2019.

[8] S. Gautam and S. M. Brahma, "Detection of high impedance fault in power distribution systems using mathematical morphology," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1226–1234, 2012.

[9] B. Wang, J. Geng, and X. Dong, "High-impedance fault detection based on nonlinear voltage–current characteristic profile identification," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3783–3791, 2016.

[10] B. Schölkopf, "The kernel trick for distances," *Advances in neural information processing systems*, vol. 13, 2000.

[11] Y. Dong, "Transient data library of solar grid integrated distributed system," https://github.com/yuqingdong0/Transient-Data-for-OEDI, 2023.