Detectability Based Data-Driven Fault Diagnosis Method for Multiple Device Faults of Converters

Fan Wu, Kai Chen, *Member, IEEE*, Gen Qiu, Hao Ying, Hanmin Sheng, and Yifan Wang

***Abstract*—The data-driven fault diagnosis method, which eliminates the need for additional sensors while preserving the system's original structure, offers a promising approach to enhancing converter reliability. However, diagnosing multiple device faults presents significant challenges due to difficulties in fault sample acquisition, detectability uncertainty, and unexplained decisions. This article addresses these challenges by presenting a detectability-based data-driven fault diagnosis method. First, a model-based fault detectability analysis method is proposed to establish the measurement conditions necessary for reliably detecting various fault types. Utilizing these measurement condition constraints, a mechanism-enhanced neural network is designed to locate faults by fitting the changes in fault parameters. The consistency between the fitting fault parameters and the actual fault process ensures the interpretability of the diagnosis results. Additionally, by guaranteeing the identification of fault parameters, the fault circuit model assists in training, significantly reducing the number of actual fault samples required for mechanism-enhanced neural network training. Finally, experiments on a representative converter are conducted to verify the effectiveness of the proposed method. Comparisons with state-of-the-art techniques show the proposed scheme's superiority in terms of diagnostic accuracy, decision explainability, sample dependence.**

***Index Terms*—Fault detectability analysis, mechanism enhanced data-driven, multiple device faults, power converters.**

I. Introduction

POWER converters, comprising main power circuits and auxiliary control circuits, play a crucial role in various power conversion applications such as renewable energy generation [1], motor drives [2], and intelligent control [3]. However, the main power circuits, including power switching devices, capacitors, and inductors, are prone to faults caused by wear, aging, corrosion, and continuous operation in harsh environments [4]. According to industry feedback, these faults account for approximately 77% of converter failures. If not promptly detected and addressed, they can pose significant safety risks and result in costly accidents, downtime, and repair expenses [5]. Installing monitoring systems for each type of fault is impractical, and fault diagnosis should not compromise the system's performance. Therefore, output characteristic fault diagnosis, which identifies faults by analyzing output characteristics without need for additional sensors, has garnered widespread research attention [6].

Currently, two types of output characteristic fault diagnosis methods for converters exist: model-based [7], [8], [9] and data-driven [10], [11], [12], [13], [14], [15], [16], [17], [18], [19]. Model-based approaches diagnose faults by comparing measurable information from the converter system with estimates from mathematical models [10]. Existing research has successfully applied model-based methods for output characteristic diagnosis of various faults, including open-circuit faults [9], wire bonding faults [7] in power switch devices, and capacitor aging faults [8]. In [9], a model-based state estimator approach is introduced for detecting and identifying open-circuit faults in power converters. However, diagnosing wire bonding and capacitor aging faults is more challenging due to weak fault output characteristics, necessitating more precise inspection methods. For instance, in [7], in-situ diagnosis of wire bonding faults for multichip power switch modules is achieved based on the crosstalk effect. Similarly, in [8], a high-precision capacitor observer based on the *LC* resonance process is developed for diagnosing dc capacitor aging faults in three-phase inverters.

In recent years, due to significant advancements in machine learning (ML) technology, data-driven fault diagnosis methods have attracted considerable research attention [11]. The principle of data-driven methods lies in extracting the mapping relationship between measurement data and fault labels [10]. During the offline stage, an ML-based diagnostic model is trained using a fault database. In online applications, this well-trained model can generate diagnostic results based on inputs such as sampled currents and voltages. Unlike model-based approaches, data-driven methods are independent of system models, making them more robust and providing better generalization capabilities across variable system conditions [15]. Current research on data-driven fault diagnosis methods primarily focuses on capacitor faults and power switch open-circuit faults.

A range of machine learning methodologies [12], [13], [14], [15] utilizing time-series data have been explored for predicting capacitor aging failures. In the domain of power switch open-circuit failures, deep learning algorithms such

Received 24 July 2024; revised 26 October 2024; accepted 30 November 2024. Date of publication 4 December 2024; date of current version 28 January 2025. This work was supported in part by the National Natural Science Foundation of China under Grant U2230206 and Grant 6237308, in part by the China Postdoctoral Science Foundation under Grant 2021M690560, and in part by the Sichuan Provincial Science and Technology Program under Grant 2023NSFSC1962. Recommended for publication by Associate Editor H. S. Krishnamoorthy. *(Corresponding author: Kai Chen.)*

Fan Wu, Kai Chen, Gen Qiu, Hanmin Sheng, and Yifan Wang are with the College of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 610000, China (e-mail: 202212060922@std.uestc.edu.cn; kaichen@uestc.edu.cn; qgen615@uestc.edu.cn; hmsheng@uestc.edu.cn; yifanwang@uestc.edu.cn).

Hao Ying is with The 60th Research Institute, China Rongtong Group, Nanjing 210000, China (e-mail: ahjsyh@126.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2024.3510749.

Digital Object Identifier 10.1109/TPEL.2024.3510749

as long short-term memory networks (LSTM), convolutional neural networks (CNN), and their derivatives have demonstrated effectiveness in diagnosing faults in power converters [16]. For instance, Yuan et al. [17] employ a one-dimensional CNN to directly extract critical features from data for identifying open-circuit faults in converters. Similarly, Si et al. [18] utilize a deep LSTM network with attention mechanisms to enhance diagnostic performance by extracting discriminative features from raw data across diverse conditions. Additionally, Hang et al. [19] propose a robust approach for diagnosing open-circuit faults in permanent magnet synchronous motor (PMSM) drives, leveraging a wavelet CNN with limited samples of normalized current vector trajectory graphs.

Despite significant advancements, current studies have predominantly focused on individual device failures. Industrial feedback data indicate that power converters can experience various device failures, including power switch failures, as well as inductance and capacitance failures, which collectively account for up to 77% of all failures [20]. These faults impact output characteristics, suggesting that a diagnostic model designed for a single device may lead to a high misdiagnosis rate [21].

Research on simultaneous fault diagnosis for multiple devices remains limited. In [29], a state-space model (SSM) for a buck converter is developed to track system parameter changes and degradation levels using dynamic mode decomposition. Similarly, Kimura and Maruta [30] proposed a parameter estimation method for fault detection in dc/dc converters, relying on a discretized state-space model (DSSM) to estimate key parameters such as inductance (L), capacitance (C), and equivalent series resistance (ESR). However, these methods heavily depend on accurate mathematical modeling, requiring precise circuit models for reliable fault detection. To address these limitations, Kareem and Hur [31] employed time-domain feature extraction and filter-based selection on parameters such as C, ESR, dissipation factor, and impedance. Various machine learning classifiers are trained using the selected features to diagnose capacitor soft faults. Nonetheless, these approaches extract diagnostic features through Fourier transforms, making the process reliant on human expertise and potentially insufficient in capturing essential features. In response, Miao et al. [32] proposed a contrastive learning (CL) method for soft fault detection in dc/dc buck converters, which identifies shared features among similar data and highlights differences between dissimilar data. However, this approach demands a large quantity of fault samples to train the model effectively, limiting its practical applicability [33].To further address these challenges, Zhao et al. [23] introduced a physics-informed machine learning (PIML) method for parameter estimation in dc/dc buck converters, combining data-driven and model-based approaches. Despite its novelty, the method remains sensitive to modeling inaccuracies and evolving fault parameters over time. Moreover, the existing studies often overlook critical power switch faults, such as open-circuit and bond wire faults, resulting in incomplete fault coverage, as highlighted in Table I.

Utilizing output characteristic data-driven fault diagnosis methods to simultaneously diagnose multiple device faults presents several challenges.

TABLE I
COMPARISON OF FAULT MODE ADAPTATION OF EXISTING METHODS

Methods	Faulty Modes	Methods	Faulty Modes
Crosstalk effect model [7]	Power switch open circuit faults	LSTM[18]	IGBT and diodes open circuit fault
LC resonance model [8]	DC-link capacitor faults	WCNN[19]	Power switch open circuit faults
Normalized cost function model [9]	Power switch open circuit faults	PIML[23]	Inductor and capacitor faults
LSTM[12]	Electrolytic capacitors	SSM[29]	Inductor and capacitor faults
DNN[13]	Electrolytic capacitors	DSSM[30]	Inductor and capacitor faults
Extended relevance vector machine[14]	Dc-link capacitor	Time-domain feature extraction [31]	Inductor and capacitor faults
Similarity measurement[15]	Electrolytic capacitor	CL[32]	Inductor and capacitor faults
1-D CNN [17]	Power switch open circuit faults	This paper	Power switch open circuit, power switch bonding wire, inductor and capacitor faults

1) *Fault Sample Collection:* Power converters exhibit complex, multidimensional variations during operation [26], making the weak fault characteristics of power devices highly sensitive to these changes. Consequently, the need for diverse fault samples to train fault diagnosis models increases exponentially with the number of device fault types. However, obtaining fault samples by enumerating all possible operating conditions in a real system is both costly and challenging [25]. Additionally, numerical simulations cannot accurately replicate the nonlinear changes in fault parameters, rendering these simulated fault samples unsuitable for direct use in model training.
2) *Lack of Detectability Analysis Method*: Existing methods rely on building a sufficiently rich database to verify fault detectability, but the difficulty in obtaining fault samples for industrial processes leads to unreliable fault detectability assessments. When detectability is unknown, we cannot determine whether the fault diagnosis model constructed under limited measurement conditions is reliable. Nor is it possible to determine what measurement conditions should be required to meet specific fault diagnosis performance.
3) *Lack of Interpretability in Decision-Making*: Data-driven models utilize multilayer nonlinear neural networks to fit the function mapping between input signals and fault labels. However, the specific signal features upon which the model bases its fault detection decisions remain unknown, thereby reducing the credibility of the fault diagnosis results.

This article addresses these challenges by presenting a detectability based data-driven fault diagnosis method with fault detectability analysis and mechanism-enhanced neural network (MENN). First, a model-based fault detectability analysis method is proposed to determine the measurement condition requirements for reliable multidevice fault diagnosis. Based on the results of diagnosability analysis, MENN is explored to achieve multidevice fault diagnosis with low sample dependence

and decision interpretability. The main contributions of this article are as follows.

1) *Model-Based Fault Detectability Analysis:* The proposed method demonstrates that parameter identification is a sufficient condition for fault diagnosability in converters. Consequently, the relationship between measurement conditions and fault diagnosability can be characterized by fault parameter identification. This approach determines the necessary measurement conditions, providing a theoretical guarantee for reliably detecting multiple device fault types.
2) *Mechanism-Enhanced Neural Network*: Utilizing these measurement condition constraints, fault detectability and fault parameter identifiability are guaranteed theoretically. Hence, MENN can be designed to fit nonlinear changes in fault parameters and classify faults. This method utilizes neural networks to predict nonlinear variations in fault parameters and incorporates fault equivalent circuit model and fault detectability analysis to guide sample collection and model training. As a result, it achieves accurate fault diagnosis with improved interpretability in decision-making and reduces dependency on extensive real fault data.

II. Model-Based Fault Diagnosability Analysis Method

This section presents a model-based fault detectability analysis method, which analyzes fault detectability through nonlinearity discrete input-output model. Firstly, fault detectability is defined by the independence of fault parameter value space (Section II-A). Then, a computable fault detectability analysis method is proposed by revealing that fault parameter identification is a sufficient condition for fault detectability (Section II-B). Finally, a fault detectability degree analysis method is proposed considering the influence of sensor noise and sampling rate on fault detectability (Section II-C).

A. Diagnosability Definition

Considering the nonlinearity discrete input-output model

$$Y = h(\theta; U, x_0) \tag{1}$$

$Y\epsilon\mathbb{R}^{N\times n\times q}, Y := [y_1, y_2, \ldots y_N]$ represent output signal; $U\epsilon\mathbb{R}^{N\times n\times q}, U := [u_1, u_2, \ldots u_N]$ represent input signal; $y_k\epsilon\mathbb{R}^{n\times q}$ and $u_k\epsilon\mathbb{R}^{n\times m}$ respectively represent q output variables and m input variables, with n observations in the k-neighborhood; x_0 represent initial state; $\theta\epsilon\Theta \subset \mathbb{R}^{N\times n_p}$, $\theta := [\theta_1, \theta_2, \ldots \theta_N]$ represent the fault parameter vector; $\theta_k\epsilon\Theta_k \subset \mathbb{R}^h$ represent an observation of n_p fault parameters in the k-neighborhood; $\Theta = \cup_{i=0}^{p} \Theta^i$, Θ^i represent the fault parameter value space of the ith fault F^i; Θ^0 represent the fault parameter value space of fault-free F^0; p represent the number of device fault types. For any two different kinds of faults in the converter: F^i and $F^j, i \neq j$, the fault parameter value space is different, so

$$\forall k, k = 1, 2 \ldots N; s.t: {\Theta_k}^i \cap {\Theta_k}^j = \emptyset. \tag{2}$$

Assuming that output characteristic fault diagnosis uses Y as input to predict the changes of θ for fault diagnosis. Based on (2), the fault detectability is defined as follows.

Definition 1: For $Y = h(\theta; U, x_0)$, given U and x_0, taking Y as the input signal to estimate θ for diagnosing F^i, F^i has global fault detectability, if $\forall\theta_{(1)}, \theta_{(2)}\epsilon\Theta$, hold that

$$\text{Condition}: \left\{h\left(U, \theta_{(1)}; x_0\right) = h\left(U, \theta_{(2)}; x_0\right)\right\} \Rightarrow \text{Result}:$$
$$\left\{\exists i, \text{make}: \theta_{(1)}, \theta_{(2)} \in \Theta^i; \nexists j, j \neq i, \text{ make}: \theta_{(1)} \text{ or } \theta_{(2)} \in \Theta^j\right\}. \tag{3}$$

Definition 1 gives a general definition of fault detectability, using the output signals. However, due to the influence of nonlinear model, the calculation of Definition 1 is difficult. Therefore, in order to obtain a fault detectability analysis model which is convenient for qualitative analysis, the nonlinear model is locally linearized. Considering the linearization of $h(\theta; U, x_0)$ in the k-neighborhood, the linear discrete input–output model is obtained as follows:

$$y_k = G(\theta_k; d)\, u_k \tag{4}$$

where d represents the order of the delay element from u_k to y_k. Through the $G(\theta_k; d)u_k$, the local fault detectability of F^i is defined as follows.

Definition 2: For $y_k = G(\theta_k; d)u_k$ linearized from $Y = h(U, \theta; x_0)$, given u_k and d, taking y_k as the input signal to estimate θ_k for diagnosing F^i, F^i has local fault detectability.

If $\forall\theta_{k,(1)}, \theta_{k,(2)}\epsilon\Theta_k$, hold that

$$\text{Condition}: \left\{G\left(\theta_{k,(1)}; d\right) u_k = G\left(\theta_{k,(2)}; d\right) u_k\right\} \Rightarrow \text{Result}:$$
$$\left\{\exists i, \text{make}: \theta_{k,(1)}, \theta_{k,(2)} \in {\Theta_k}^i; \nexists j, j \neq i, \text{make}: \theta_{k,(1)} \text{ or } \theta_{k,(2)} \in {\Theta_k}^j\right\} \tag{5}$$

Next, this article explains the connection between Definition 1 and Definition 2 through Theorem 1.

Theorem 1: If $\forall k, k = 12 \ldots N$, $y_k = G(\theta_k; d)u_k$, F^i has local fault detectability, then F^i has global fault detectability.

Proof: Because $\forall k, k = 12 \ldots N$, $y_k = G(\theta_k; d)u_k$, F^i has local fault detectability, condition equivalence (CE) between the Definition 1 and Definition 2 is obtained.

Due to $\theta_{(1)} = [\theta_{1,(1)}, \ldots \theta_{N,(1)}]$, $\theta_{(2)} = [\theta_{1,(2)}, \ldots \theta_{N,(2)}]$, $h(\theta_{(1)}; U, x_0) = [G(\theta_{1,(1)}; d)u_1, \ldots, G(\theta_{N,(1)}; d)u_N]$, $h(\theta_{(2)}; U, x_0) = [G(\theta_{1,(2)}; d)u_1, \ldots, G(\theta_{N,(2)}; d)u_N]$, So,CE:

$$\left\{\forall \text{k}, \theta_{k,(1)}, \theta_{k,(2)}\epsilon\Theta_k, \text{hold that}: \left\{G\left(\theta_{k,(1)}; d\right) u_k = G\left(\theta_{k,(2)}, d\right) u_k\right\}\right\}$$
$$\Leftrightarrow \left\{\forall\theta_{(1)}, \theta_{(2)}\epsilon\Theta, \text{hold that}: \left\{h\left(\theta_{(1)}; U, x_0\right) = h\left(\theta_{(2)}; U, x_0\right)\right\}\right\}.$$

Assuming N is finite, the nature of the F^i, remains unchanged, hence

If $\exists k$, make : $\theta_{k,(1)}, \theta_{k,(2)} \in {\Theta_k}^i$, then, $\theta_{(1)}, \theta_{(2)} \in \Theta^i$.

Due to $\Theta^i = [\,{\Theta_1}^i, \ldots, {\Theta_N}^i]$, $\Theta^j = [\,{\Theta_1}^j, \ldots, {\Theta_N}^j]$, result equivalence between the between the Definition 1 and Definition 2 is obtained

$$\left\{\forall k, \left\{\exists i, \text{make} : \theta_{k,(1)}, \theta_{k,(2)} \in {\Theta_k}^i; \nexists j \neq i, \text{make} : \theta_{k,(1)} \text{ or } \theta_{k,(2)} \in {\Theta_k}^j\right\}\right\}$$

$$\Leftrightarrow \left\{\exists i, \text{make} : \theta_{(1)}, \theta_{(2)} \in \Theta^i; \nexists j, j \neq i, \text{make} : \theta_{(1)} \text{ or } \theta_{(2)} \in \Theta^j\right\}.$$

Hence, for $\forall k$, the conditions and results of Definition 1 and Definition 2 are equal. Theorem 1 is proved.

Theorem 1 shows that if a fault is local fault detectability at all discrete sampling points, then the fault is global fault detectability. Therefore, in the computable fault detectability analysis method proposed in Section II-B, global fault detectability can be derived from local fault detectability.

B. Computable Fault Detectability Analysis Method

To make the local fault detectability (5) computable, this section derives the sufficient condition for the local fault detectability through parameter identifiability. First, parameter identifiability [27] is defined as follows: For $y_k = G(\theta_k; d)u_k$, given u_k and d, taking y_k as the input signal to obtain θ_k for diagnosing F^i, θ_k has parameter identifiability, if $\forall \theta_{k,(1)}, \theta_{k,(2)} \epsilon \Theta_k$, hold that

$$\begin{aligned}&\text{Condition} : \left\{G\left(\theta_{k,(1)}; d\right)\ u_k = G\left(\theta_{k,(2)}; d\right) u_k\right\} \Rightarrow \\ &\quad \text{Result} : \left\{\theta_{k,(1)} = \theta_{k,(2)}\right\}\end{aligned} \tag{6}$$

Theorem 2: For $y_k = G(\theta_k; d)u_k$ linearized from $Y = h(\theta; U, x_0)$, given u_k and d, taking y_k as the input signal to obtain θ_k for diagnosing F^i, parameter identifiability of θ_k is a sufficient condition for local fault detectability of F^i.

Proof: When θ_k is parameters identifiability, $\forall \theta_{k,(1)}, \theta_{k,(2)} \epsilon \Theta_k$, hold that

$\{G(\theta_{k,(1)}; d)\ u_k = G(\theta_{k,(2)}; d)u_k\} \Rightarrow \{\ \theta_{k,(1)} = \theta_{k,(2)}\ \}$. Due to $\Theta_k = \cup_{i=0}^{p} {\Theta_k}^i$, hence, $\exists i \in \{0, 1, ..p\}$, make: $\theta_{k,(1)}, \theta_{k,(2)} \epsilon {\Theta_k}^i$. If $\exists j \in \{0, 1, ..p\}, j \neq i$, make : $\theta_{k,(1)}$ or $\theta_{k,(2)} \epsilon {\Theta_k}^j$,then $\theta_{k,(1)}, \theta_{k,(2)} \in {\Theta_k}^i \cap {\Theta_k}^j$. This contradicts (2) ${\Theta_k}^i \cap {\Theta_k}^j = \emptyset$. Thus, $\forall \theta_{k,(1)}, \theta_{k,(2)} \epsilon \Theta_k$, when $G(\theta_{k,(1)}; d)\ u_k = G(\theta_{k,(2)}; d)u_k$, we can derive that $\{\exists i$, make : $\theta_{k,(1)}, \theta_{k,(2)} \in {\Theta_k}^i; \nexists j, j \neq i$, make : $\theta_{k,(1)}$ or $\theta_{k,(2)} \in {\Theta_k}^j\}$. Theorem 2 is proved.

Next, a calculable fault detectability analysis method of F^i is derived from Theorem 2. $y_k = G(d, \theta_k)u_k$ can be expressed in the following matrix form:

$${Y_W}^{\mathrm{T}} = \Psi {\theta_k}^{\mathrm{T}}$$

$$\begin{aligned}Y_W &= (Y_1, Y_2, \ldots Y_{w-h-d+1}) = (y_{k,d+h},\ y_{k,d+h+1}, \ldots y_{k,w}) \\ \Psi &= (\Psi_1, \Psi_2, \ldots \Psi_{w-h-d+1})^{\mathrm{T}} \\ &= \begin{pmatrix} -y_{k,h+d-1}, \ldots - y_{k,h+d-p_y}, u_{k,h-1}, \ldots u_{k,h-p_u} \\ -y_{k,h+d}, \ldots - y_{k,h+d-p_y+1}, u_{k,h}, \ldots u_{k,h-p_u+1} \\ \cdot \quad\quad \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \quad\quad \cdot, E \\ \cdot \quad\quad \cdot \quad\quad \cdot \\ -y_{k,w-1}, \ldots - y_{k,w-p_y}, u_{k,w-d-1}, \ldots u_{k,w-d-p_u} \end{pmatrix}\end{aligned} \tag{7}$$

where h and w represent the beginning time and end time of the observation, respectively; $y_{k,w}$ and $u_{k,w}$ represent the observations of y_k and u_k at time w; p_y and p_u denote the number of parameters in θ_k related to y_k and u_k; E is the identity matrix of dimension $(w - h - d + 1) \times p_e$, p_e represents the number of unknown variables. The optimal θ_k is estimated by minimizing the objective function $\mathrm{V}(\theta_k)$

$$\begin{cases} \theta_{k,\text{opt}} = \text{argmin}_{\theta_k} \mathrm{V}(\theta_k) \\ V(\theta_k) = \left(\widetilde{{Y_W}^{\mathrm{T}}} - \Psi {\theta_k}^{\mathrm{T}}\right)^{\mathrm{T}} \left(\widetilde{{Y_W}^{\mathrm{T}}} - \Psi {\theta_k}^{\mathrm{T}}\right) \end{cases} \tag{8}$$

where $\widetilde{{Y_W}^{\mathrm{T}}}$ represents the observation values of ${Y_W}^{\mathrm{T}}$. $\theta_{k,\text{opt}}$ represents the optimal value of θ_k, which is achieved by $\left.\frac{\partial \mathrm{V}(\theta_k)}{\partial \theta_k}\right|_{\theta_k = \theta_{k,\text{opt}}} = 0$

$$\begin{aligned}\frac{\partial \mathrm{V}(\theta_k)}{\partial \theta_k} &= -2\Psi^{\mathrm{T}}\left(\widetilde{{Y_W}^{\mathrm{T}}} - \Psi {\theta_k}^{\mathrm{T}}\right),\ \theta_{k,\text{opt}} \\ &= \left(\Psi^{\mathrm{T}}\Psi\right)^{+} \Psi^{\mathrm{T}} \widetilde{{Y_W}^{\mathrm{T}}}\end{aligned} \tag{9}$$

where $\left(\Psi^{\mathrm{T}}\Psi\right)^{+}$ represents the pseudoinverse of $\Psi^{\mathrm{T}}\Psi$. If θ_k has parameter identifiability, then $\theta_{k,\text{opt}}$ is uniquely determined. Therefore, $\left.\frac{\partial^2 \mathrm{V}(\theta_k)}{\partial {\theta_k}^2}\right|_{\theta_k = \theta_{k,\text{opt}}} = \Psi^{\mathrm{T}}\ \Psi > 0$. When $\Psi^{\mathrm{T}}\Psi > 0$, the rank ($r$) of $\Psi^{\mathrm{T}}\Psi$ satisfies: $r = n_p = p_y + p_u + p_e$. This is a sufficient condition for the parameter identifiability [27]. Where n_p represents the parameter number of θ_k. Thus, the following theorem can be derived.

Theorem 3: For $y_k = G(\theta_k; d)u_k$ linearized from $Y = h(\theta; U, x_0)$, given u_k and d, taking y_k as the input signal to obtain θ_k for diagnosing F^i. When $\left.\frac{\partial \mathrm{V}(\theta_k)}{\partial \theta_k}\right|_{\theta_k = \theta_{k,\text{opt}}} = 0$ form (8), r represents the rank of $\Psi^{\mathrm{T}}\Psi$, n_p represents the number of θ_k, if $r = n_p$, then F^i has local fault detectability.

Based on Theorem 3, a computable fault detectability analysis method is obtained that does not require actual data and model training, as Ψ can be computed numerically from fault equivalent circuit model (25).

C. Fault Detectability Degree Analysis Method

Due to the insulated gate bipolar transistor (IGBT) wire bonding fault, inductor turns short circuit, and capacitor aging fault, the fault parameters change less relative to the rating parameters,

fault detectability is easily affected by sampling rate and SNR. To address this issue, this section proposes a fault detectability degree analysis method.

Considering the different sampling rates and SNR, the estimation method of θ_k in (8) is redefined as follows:

$$\begin{cases} \theta_{k,\mathrm{opt}} = \mathrm{argmin}_{\theta_k} V(\theta_k) \\ V(\theta_k) = \left(\overline{Y_W{}^{\mathrm{T}}} - \Psi\theta_k{}^{\mathrm{T}}\right)^{\mathrm{T}} \mathrm{P_V}^{-1} \left(\overline{Y_W{}^{\mathrm{T}}} - \Psi\theta_k{}^{\mathrm{T}}\right) \end{cases} \tag{10}$$

where $\overline{Y_W{}^{\mathrm{T}}}$ represents the measured values of $Y_W{}^{\mathrm{T}}$ with certain sampling, $\mathrm{P_V}$ denotes the estimated covariance matrix of SNR interference affecting the measured output. Therefore, when fault *i* occurs, the estimated fault parameter value can be obtained as follows:

$$\theta_{k,\mathrm{opt}}{}^{i} = \left(\Psi^{i\mathrm{T}}\mathrm{P_v}^{-1}\Psi^{i}\right)^{+} \Psi^{i\mathrm{T}}\mathrm{P_v}^{-1}\overline{Y_W{}^{i\mathrm{T}}}. \tag{11}$$

Due to the uncertainties associated with $\mathrm{P_v}$ and $\overline{Y_W{}^{\mathrm{T}}}$, $\theta_{k,\mathrm{opt}}{}^{i}$ is considered a random matrix. Consequently, each fault type can be characterized by a set of probability density functions $p(\theta_{k,\mathrm{opt}}{}^{i})$

$$\mathcal{F}^{i} = \left\{p\left(\theta_{k,\mathrm{opt}}{}^{i}\right), \forall\theta_{k,\mathrm{opt}}{}^{i} \subset \Theta^{i}\right\} \tag{12}$$

The fault detectability degree of F^i is defined based on the significant differences of the estimated $\theta_{k,\mathrm{opt}}$ for different fault types. The more significant differences indicate less difficulty in locating these faults

$$d^{i} = \min\{\mathrm{K}(\mathcal{F}^{i}, \mathcal{F}^{j}), \forall j, j \neq i\} \tag{13}$$

The detectability degree of fault *i* relative to fault *j* is defined as follows:

$$d^{i,j} = \mathrm{K}(\mathcal{F}^{i}, \mathcal{F}^{j})$$

$$\mathrm{K}\left(\mathcal{F}^{i}, \mathcal{F}^{j}\right) := \sqrt{\left(\overline{\theta_{k,\mathrm{opt}}{}^{i}} - \overline{\theta_{k,\mathrm{opt}}{}^{j}}\right)^{\mathrm{T}} \Sigma_{i,j}{}^{-1} \left(\overline{\theta_{k,\mathrm{opt}}{}^{i}} - \overline{\theta_{k,\mathrm{opt}}{}^{j}}\right)} \tag{14}$$

where $\mathrm{K}(\mathcal{F}^{i}, \mathcal{F}^{j})$ represents the relative entropy of the *probability* density function. To estimate $\mathrm{K}(\mathcal{F}^{i}, \mathcal{F}^{j})$, Mahalanobis distance [34] is introduced, which is used to estimate the deviation of probability distribution based on mean and variance. $\overline{\theta_{k,\mathrm{opt}}{}^{i}}$ and $\overline{\theta_{k,\mathrm{opt}}{}^{j}}$ represent the means of $\theta_{k,\mathrm{opt}}{}^{i}$ and $\theta_{k,\mathrm{opt}}{}^{j}$, respectively, while $\Sigma_{i,j}{}^{-1}$ is the inverse of the covariance matrix of $\theta_{k,\mathrm{opt}}{}^{i}$ and $\theta_{k,\mathrm{opt}}{}^{j}$.

Remark 1: By performing numerical calculations using (25), the parameters Ψ and $\mathrm{P_v}$ under various operating conditions and sampling rates can be obtained. These parameters are then substituted into (11) to generate multiple datasets of $\theta_{k,\mathrm{opt}}{}^{i}$ for different faults. Finally, the detectability degree is computed by substituting the multiple datasets into (14).

III. Fault Diagnosability Analysis of Converter

In this section, based on the proposed fault detectability analysis method, the fault diagnosability of key power devices of typical converter structure is analyzed.

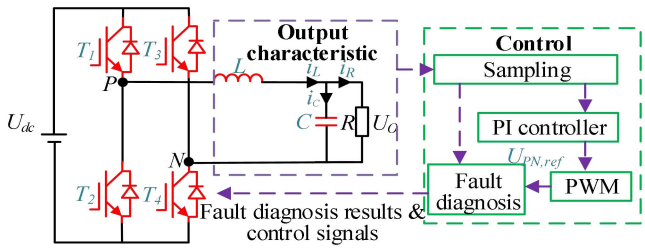

Fig. 1. Fault diagnosis system of the full-bridge converter.

TABLE II
Fault Labels for Multidevice Fault Location

Fault type	Fault device	Labels	Fault type	Fault device	Labels
normal	normal	0		normal	0
IGBT wire bonding fault	T_1	1	IGBT open-circuit fault	T_1	12
	T_2	2		T_2	13
	T_3	3		T_3	14
	T_4	4		T_4	15
	T_1&T_2	5		T_1&T_2	16
	T_1&T_3	6		T_1&T_3	17
	T_1&T_4	7		T_1&T_4	18
	T_2&T_3	8		T_2&T_3	19
	T_2&T_4	9		T_2&T_4	20
	T_3&T_4	10		T_3&T_4	21
Inductor turns short circuit fault	L	11	Capacitor aging fault	C	22

A. Fault Diagnosis System Description of Converter

This article focuses on a widely used single-phase full-bridge PWM converter system, while the proposed method retains theoretical applicability to other converter topologies. The circuit topology, illustrated in Fig. 1, comprises a main power circuit and an auxiliary control unit. The main power circuit consists of IGBTs $T_1 - T_4$, a filter inductor L, and a filter capacitor C. The control circuit collects output characteristics and achieves stable control of the circuit through a proportional integral (PI) controller and pulsewidth modulation (PWM) modulation.

Focusing on easily damaged devices such as IGBTs, filter capacitor, and filter inductor, this article explores an output characteristic data-driven fault diagnosis method. In this method, the function mapping between the output characteristics and the fault labels is derived from the fault data, enabling fault location by referencing the fault labels in Table II. Based on the analysis of the fault equivalent circuit model of converter, the proposed fault diagnosability analysis determines the necessary measurement conditions of output characteristics for reliable multidevice fault type diagnosis, including which output signals need to be measured, the sampling rate, and SNR.

B. Fault Equivalent Circuit Model

Considering the IGBT wire bonding fault, IGBT open-circuit fault, inductor turns short-circuit fault, and capacitor aging fault, the fault equivalent circuit is modeled in this section.

IGBT faults mainly include bond wire faults and open-circuit faults [24]. These faults affect the bridge arm output voltage

TABLE III
VALUE OF S_P WHEN THE IGBT IS OPEN-CIRCUIT FAULT

S_1	S_2	S_{FO1}	S_{FO2}	S_i	S_P
1	0	0	1/0	1	1
1	0	1	1/0	1	−1
1	0	1/0	1/0	0	1
0	1	1/0	1/0	1	−1
0	1	1/0	1	0	1
0	1	1/0	0	0	−1

and the IGBT's ON-state resistance. Therefore, the device fault model for IGBTs can be equivalently represented by a controlled voltage source U_{PN} and internal resistance R_{PN}, as described in the following:

$$\begin{cases} U_{PN} = \frac{U_{\rm dc}}{2}\ (S_P - S_N) \\ S_P,\ S_N = \begin{cases} 1\ P, N \text{ connect to the positive } U_{\rm dc} \\ -1\ P, N \text{ connect to the negative } U_{\rm dc} \end{cases} \end{cases} \tag{15}$$

where $U_{\rm dc}$ represents the dc bus voltage, S_P and S_N respectively denote the connection states of P and N to $U_{\rm dc}$, which are determined by the IGBT's conduction status $S_1 - S_4$, open-circuit fault status $S_{FO1} - S_{FO4}$, and the output current polarity S_i, as illustrated in Table III

$$S_x = \begin{cases} 1,\ T_x \text{ Turn on} \\ 0,\ T_x \text{ Turn off} \end{cases} \quad x = 1, 2, 3, 4 \tag{16}$$

$$S_{FOx} = \begin{cases} 1,\ T_x \text{ Open} - \text{circuit fault} \\ 0, \qquad\qquad T_x \text{ Normal} \end{cases} \quad x = 1, 2, 3, 4 \tag{17}$$

$$S_i = \begin{cases} 1,\ i_L \geq 0 \\ 0,\ i_L < 0. \end{cases} \tag{18}$$

The following equation can be obtained by simplification of the Table III based on Carnaugh diagram:

$$\begin{cases} S_P = S_{P+} - S_{P-} \\ \qquad S_{P+} = S_1\left(\overline{S_{FO1}} S_i + \overline{S_i}\right) + S_2 S_{FO2} \overline{S_i} \\ \qquad S_{P-} = S_2\left(\overline{S_{FO2}} S_i + S_i\right) + S_1 S_{FO1} S_i. \end{cases} \tag{19}$$

In the same way

$$\begin{cases} S_N = S_{N+} - S_{N-} \\ \qquad S_{N+} = S_3\left(\overline{S_{FO3}} S_i + \overline{S_i}\right) + S_4 S_{FO4} \overline{S_i} \\ \qquad S_{N-} = S_4\left(\overline{S_{FO4}} S_i + S_i\right) + S_3 S_{FO3} S_i. \end{cases} \tag{20}$$

The internal resistance R_{PN} is equivalent as follows:

$$\begin{cases} R_{PN} = R_{P+} + R_{P-} + R_{N+} + R_{N-} \\ \quad R_{P+} = S_{P+}\left(R_{\rm on1} + S_{FB1} R_{FB1}\right) \\ \quad R_{P-} = S_{P-}\left(R_{\rm on2} + S_{FB2} R_{FB2}\right) \\ \quad R_{N+} = S_{N+}\left(R_{\rm on3} + S_{FB3} R_{FB3}\right) \\ \quad R_{N-} = S_{N-}\left(R_{\rm on4} + S_{FB4} R_{FB4}\right) \end{cases} \tag{21}$$

where $R_{{\rm on}x}$ represents the conduction impedance of T_x, R_{FBx} represents the change in conduction resistance caused by bond wire faults in T_x, and S_{FBx} represents the bond wire fault status of T_x

$$S_{FBx} = \begin{cases} 1, \quad T_x \text{ bond wire fault} \\ 0,\ T_x \text{ bond wire normal.} \end{cases} \tag{22}$$

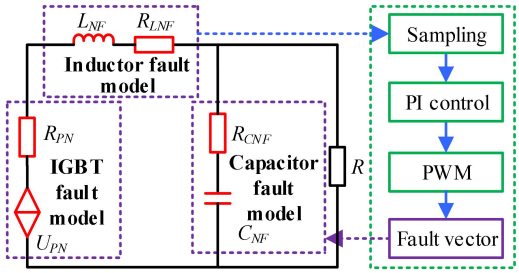

Fig. 2. Unified circuit model of converter fault.

The failure modes of filter inductors primarily include inter-turn short circuits and degradation of magnetic materials [24]. Therefore, their fault equivalent circuit model can be represented using equivalent inductance L_{NF} and equivalent resistance R_{LNF}

$$\begin{cases} L_{NF} = L_N + S_{FL} L_F \\ R_{LNF} = R_{LN} + S_{FL} R_{LF} \\ S_{FL} = \begin{cases} 1, \quad \text{Inductor fault} \\ 0, \ \text{Inductor normal} \end{cases} \end{cases} \tag{23}$$

where L_N and R_{LN} represent the normal values of inductance and equivalent resistance, respectively, L_F and R_{LF} represent the change values in inductance and equivalent resistance during faults, respectively, and S_{FL} denotes the fault status of the inductor.

Capacitor fault primarily manifests as a decrease in capacitance value and an increase in equivalent impedance [24], making its fault equivalent circuit model representable by equivalent capacitance C_{NF} and equivalent resistance R_{CNF}

$$\begin{cases} C_{NF} = C_N + S_{FC} C_F \\ R_{CNF} = R_{CN} + S_{FC} R_{CF} \\ S_{FC} = \begin{cases} 1, \quad \text{Capacitor fault} \\ 0, \ \text{Capacitor normal} \end{cases} \end{cases} \tag{24}$$

where C_N and R_{CN} respectively represent the normal capacitance value and equivalent resistance value, C_F and R_{CF} respectively denote the changes in capacitance value and equivalent resistance value during faults, and S_{FC} represents the fault status of the capacitor. Thus, a unified circuit model for converter faults can be constructed as shown in Fig. 2.

Its numerical computational model can be represented as follows:

$$\begin{cases} L_{NF} \frac{di_L}{dt} = U_{PN} - i_L R_{PN} - i_L R_{LNF} - U_O \\ C_{NF} \frac{dU_c}{dt} = i_L - i_R \\ \qquad \frac{dU_{PN,\rm ref}}{dt} - \frac{k_p dU_o}{dt} = k_i\left(U_o - U_{\rm ref}\right) \\ U_{PN} = f_{{\rm PWM},U}(U_{PN,\rm ref}, S) \\ R_{PN} = f_{{\rm PWM},R}(U_{PN,\rm ref}, S) \\ \qquad S = \Big[S_{FO1}, S_{FO2}, S_{FO3}, S_{FO4}, S_{FB1}, \\ S_{FB2}, S_{FB3}, S_{FB4}, S_{FL}, S_{FC}, S_i, S_x\Big] \\ U_O = i_C R_{CNF} + U_c \end{cases} \tag{25}$$

where $U_{PN,\text{ref}}$ represents the reference modulation signal of the PWM modulation. $f_{\text{PWM},U}$ and $f_{\text{PWM},R}$ represent functions consisting of (19)–(21) and PWM modulation, S represents the fault vector composed of the multiple device fault states, k_p and k_i indicates PI controller parameters, and $L_{NF}, C_{NF}, R_{CNF}, R_{PN}$, etc., are the fault parameters of the devices, obtained by fitting experimental data to model the nonlinear changes in device parameters during faults.

Therefore, based on the unified fault circuit model (25), the transmission characteristics of faults to outputs can be analyzed through numerical calculation, as proposed in the fault diagnosability analysis method in Section III-C. Additionally, rich simulation samples are generated from (25) to assist in the training of the proposed MANN in Section IV.

C. Fault Detectability Analysis Results

In the fault detectability analysis, different output characteristics used as input signals for fault diagnosis will lead to variations in the matrix $\Psi^{\mathrm{T}}\Psi$, as shown in (7). According to Theorem 3, changes in the rank of $\Psi^{\mathrm{T}}\Psi$ will affect fault detectability. Therefore, this section first analyzes the impact of using different output characteristics as input signals for fault detectability, determining which output characteristics should be measured to achieve reliable multi-device fault diagnosis.

1) Fault Detectability Analysis: First, the local linearization and discretization expressions of (25) can be obtained using forward Euler's difference formula as follows:

$$\begin{cases} i_L(w+1) = \frac{\Delta T k_{S,u} U_{PN,\text{ref}}(w)}{L_{NF}} \\ \quad + \frac{(\Delta T(-R_{PN}-R_{LNF})+L_{NF}) i_L(w)}{L_{NF}} + \frac{\Delta T U_o(w)}{L_{NF}} \\ U_c(w+1) - U_c(w) = \frac{\Delta T}{C_{NF}} (i_L(w) - i_o(w)) \\ U_{PN,\text{ref}}(w+1) - U_{PN,\text{ref}}(w) = k_p U_o(w+1) \\ \quad + U_o(w)(\Delta T k_i - k_p) - \Delta T k_i U_{\text{ref}}(w) \\ U_c(w+1) = \frac{\Delta T}{C_{NF}R_{CNF}} U_O(w) + \frac{\Delta T - C_{NF}R_{CNF}}{C_{NF}R_{CNF}} U_c(w) \end{cases} \tag{26}$$

where w represents the current discrete time; ΔT represents the sampling period of the measured signal; $k_{S,u}$ represents the linearization coefficient of $f_{\text{PWM},U}$. Expressing (26) as the matrix form of (7) yields the following:

$$(Y_1, Y_2, ..Y_w)^{\mathrm{T}} = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \cdot \\ \cdot \\ \cdot \\ \Psi_w \end{pmatrix} \theta^{\mathrm{T}} \tag{27}$$

where

$$Y_w = \begin{pmatrix} i_L(w+1) \\ U_c(w+1) - U_c(w) \\ U_{PN,\text{ref}}(w+1) - U_{PN,\text{ref}}(w) \\ U_c(w+1) \end{pmatrix} \tag{28}$$

TABLE IV
FAULT DETECTABILITY ANALYSIS RESULTS OF DIFFERENT CASE

Case	Input signal of fault diagnosis	n_p	r	Global fault detectability
1	$Y=[U_O, i_R, i_L, U_{PN,ref}]$	9	9	√
2	$Y=[i_L, U_c, U_O, U_{PN,ref}]$	9	9	√
3	$Y=[U_O, i_L, U_{PN,ref}]$	9	9	√
4	$Y=[U_C, i_L, U_{PN,ref}]$	9	9	√
5	$Y=[U_O, i_R, U_{PN,ref}]$	9	7	×
6	$Y=[U_O, U_{PN,ref}]$	9	6	×
7	$Y=[i_R, U_{PN,ref}]$	9	6	×
8	$Y=[i_L, U_{PN,ref}]$	9	6	×
9	$Y=[U_{PN,ref}]$	9	3	×
10	$Y=[U_O]$	9	3	×

$$\Psi_w = \begin{pmatrix} U_{PN,\text{ref}}(w), i_L(w), U_o(w), 0,0,0,0,0,0 \\ 0,0,0, i_L(w) - i_o(w), 0,0,0,0,0 \\ 0,0,0,0, U_o(w+1), U_o(w), U_{\text{ref}}(w), 0,0 \\ 0,0,0,0,0,0,0, U_O(w), U_c(w) \end{pmatrix} \tag{29}$$

$$\theta = \left[\frac{\Delta T k_{S,u}}{L_{NF}}, \frac{\Delta T(-R_{PN}-R_{LNF})+L_{NF}}{L_{NF}}, \frac{\Delta T}{L_{NF}}, \frac{\Delta T}{C_{NF}}, k_p, \Delta T k_i - k_p, -\Delta T k_i, \frac{\Delta T}{C_{NF}R_{CNF}}, \frac{\Delta T - C_{NF}R_{CNF}}{C_{NF}R_{CNF}} \right] \tag{30}$$

where the dimensionality of θ is 9, hence $n_p = 9$. Runge-Kutta [35] numerical calculation method is used to solve (25) to obtain the numerical solution of $i_L, U_c, U_o, U_{PN,\text{ref}}$. By substituting $i_L, U_c, U_o, U_{PN,\text{ref}}, U_{\text{ref}}$ into $\Psi = (\Psi_1, \Psi_2, ..\Psi_w)$, the rank of $\Psi^T\Psi$ is obtained: $r = 9$. According to Theorem 3, when $r = n_p$, the fault mentioned above has local fault detectability. Under the different operating point, the linearization of (25) can also obtain $r = n_p = 9$. Hence, according to Theorem 1, the fault mentioned above has global fault detectability. Therefore, taking $Y = [i_L, U_c, U_{PN,\text{ref}}]$ as the fault diagnosis input signal can realize the fault location of the IGBT wear bonding fault, IGBT open-circuit fault, inductor turns short circuit, and capacitor aging fault. Note that using different output characteristics as input signals for fault diagnosis will change the structure of (25) and affect the rank of $\Psi^T\Psi$. Therefore, this section analyzes the fault diagnosability of using different output characteristics as input signals for fault diagnosis. The fault detectability for different cases is summarized in Table IV. It shows that when the output characteristics of Cases 1–4 are measured as input signals of fault diagnosis, the faults mentioned above have global fault detectability.

2) Fault Detectability Degree Analysis: When the input signals of fault diagnosis are determined according to Table IV, the detectability of fault diagnosis will be affected by different measured SNR and sampling rate. For example, in (11), different sensor SNR will affect $\mathrm{P_V}$, and different sampling rates will change $\overline{Y_W{}^{i^{\mathrm{T}}}}$. This results in different fault detectability degree in (13). Therefore, this section uses the fault diagnosis input signal of Case 3 in Table IV as an example. Based on the fault

TABLE V
FAULT DETECTABILITY DEGREE OF CASE 3

$d^{i,j}$	$\mathcal{F}^1$ Open circuit fault	$\mathcal{F}^2$ Wire bonding fault	$\mathcal{F}^3$ Inductor fault	$\mathcal{F}^4$ Capacitor aging fault	d^i
$\mathcal{F}^1$	0	1.993	1.896	2.191	1.896
$\mathcal{F}^2$	1.993	0	1.114	1.108	1.108
$\mathcal{F}^3$	1.896	1.114	0	1.117	1.114
$\mathcal{F}^4$	2.191	1.108	1.117	0	1.108

Note: *Sampling rate of 165 kHz and SNR of 45 dB for $Y = [U_O, i_L, U_{PN,ref}]$.

TABLE VI
MAIN PARAMETERS OF THE REAL SYSTEM

Parameter	Value
Rated output voltage/frequency	220 V/50 Hz
DC bus voltage	500 V/DC
Rated power	3 kW
Switching frequency	10 kHz
Inductor L_N;R_{LN}	29.1 μH ;36.2 mΩ
Capacitor C_N;R_{CN}	156.8 μF;0.234 Ω
IGBT;R_{PN}	F4-50R12KS4;21.4 mΩ

diagnosability degree analysis method proposed in Section II-C, the influence of different sampling rates and SNR on the fault diagnosability degree is analyzed. This analysis aims to determine the sampling rates and SNR necessary for achieving reliable multidevice fault diagnosis.

First, the change of sampling rate and SNR is equivalent by adjusting $\overline{Y_W{}^i}^{\mathrm{T}}$ and $\mathrm{P_V}$. Different sampling rates and SNR are adjusted using the grid search method to ensure that the minimum $d^{i,j}$ is greater than 1. This is because in (14), $(\overline{\theta_{k,op}{}^i} - \overline{\theta_{k,op}{}^j})^{\mathrm{T}} (\overline{\theta_{k,op}{}^i} - \overline{\theta_{k,op}{}^j})$ represents the average distance of fault parameters between different faults, and $\Sigma_{i,j}$ represents the covariance of different faults affected by disturbances. When the minimum $d^{i,j}$ is greater than 1, the distance between the parameters of different fault types is greater than the disturbance distance. Hence, the sampling rate and SNR can be determined using this method to satisfy the performance requirements for fault diagnosis. The results, shown in Table V, is obtained at a sampling rate of 165 kHz and an SNR of 45 dB. The minimum $d^{i,j} = 1.114$ indicates that if $Y = [U_O, i_L, U_{PN,\mathrm{ref}}]$ is used as the input signal for fault diagnosis, the sampling rate should be greater than 165 kHz and the SNR should be greater than 45 dB. Under these measurement conditions, the fault diagnosability requirements for diagnosing multiple device types can be met.

IV. PROPOSED DETECTABILITY BASED DATA-DRIVEN FAULT DIAGNOSIS METHOD

A. Fault Diagnosis Model Construction Process

The fault detectability analysis results obtained through the proposed fault detectability analysis method are shown in Tables V and VI. Table VI indicates that measuring the input signal of fault diagnosis of Cases 1–4 can achieve multidevice fault diagnosis. The results in Table V show that consider measurement condition interference, using $Y = [U_O, i_L, U_{PN,\mathrm{ref}}]$ as the input signal for fault diagnosis, the fault detectability degree greater than 1 when the sampling rate and signal-to-noise ratio are greater than 165 kB and 45 dB, respectively. Therefore, the proposed detectability based data-driven fault diagnosis method optimizes the measurement conditions for fault diagnosis through fault detectability analysis, thereby providing a theoretical guarantee for reliable multidevice fault diagnosis.

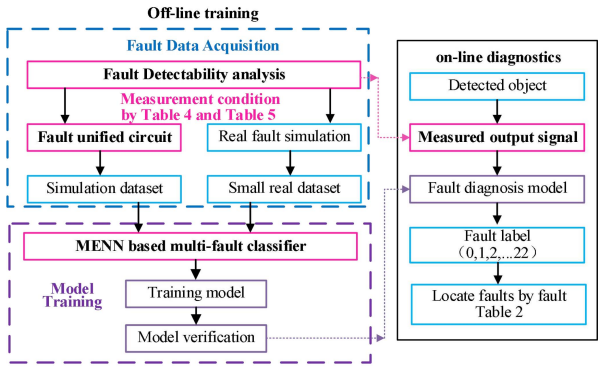

Fig. 3. Fault diagnosis model construction process of the proposed detectability based data-driven fault diagnosis method.

The fault diagnosis model construction process for the proposed detectability-based data-driven fault diagnosis method is illustrated in Fig. 3. It involves two main stages: offline training and online fault diagnosis. During offline training, the fault detectability analysis results from Tables V and VI are used to establish the measurement conditions for fault sample collection. A large number of simulation samples are generated to create a simulation dataset through numerical calculations based on the fault equivalent circuit model (25). Additionally, a small amount of actual fault data is collected from a physical system of a single-phase full-bridge PWM converter to form the real system database. The proposed MENN is then trained using these datasets, learning the functional mapping between the fault diagnosis input signals and fault labels. Finally, the trained fault diagnosis model is deployed on the detected object for online fault diagnosis.

B. Network Structure and Training Methods of MENN

The measurement conditions obtained through fault detectability analysis (see Tables V and VI) ensure fault detectability and fault parameter identifiability. Therefore, the core idea of the MENN is to fit the variations of fault parameters over multiple time intervals based on input signals, aiming to localize faults by understanding the patterns of these parameter changes. This approach offers several significant benefits. 1) Given the identifiable nature of fault parameters, a reliable fault parameter predictor can be developed. Since changes in fault parameters correspond to actual physical principles and are measurable, fault localization through fault parameter changes provides interpretability for fault diagnosis decisions. 2) Because fault parameters are identifiable, (25) can assist in training neural networks for fault parameter prediction according to the physical informed neural network principle [22], [23]. This reduces the need for large amounts of real fault samples during model training.

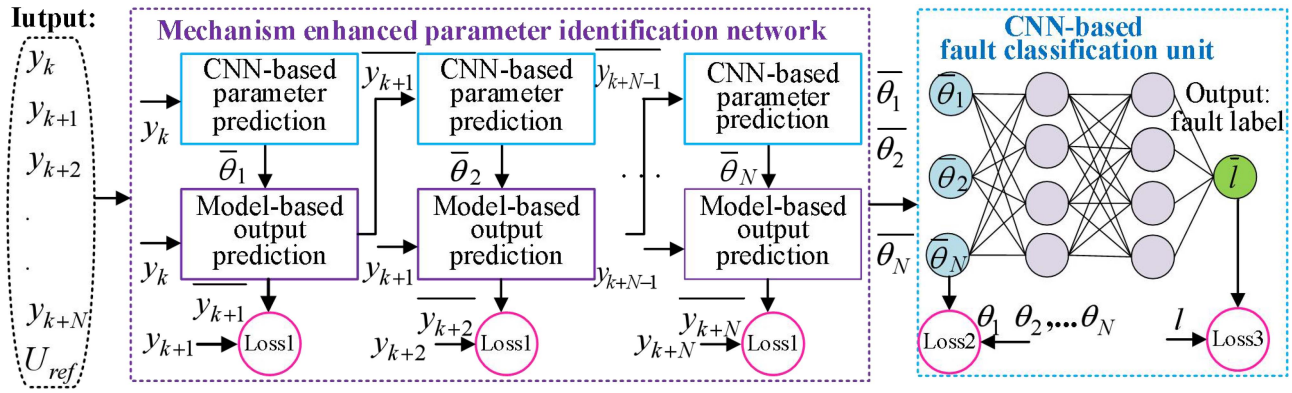

Fig. 4. Network structure of the proposed MENN.

The network structure of the proposed MENN is shown in Fig. 4. It comprises a mechanism-enhanced parameter identification network and a CNN-based fault classification unit. The mechanism-enhanced parameter identification network predicts the nonlinear fault parameters based on the input signals, while the fault classification unit predicts the fault labels based on these predicted nonlinear fault parameters. The faults are then located according to the fault labels refer to the Table II. The parameter identification network is composed of CNN-based parameter prediction unit and model-based output prediction unit. The parameter prediction unit utilizes the strong nonlinear fitting ability of neural networks to predict fault parameters, and its mathematical model is expressed as follows:

$$\left(\overline{\theta_1},\ldots,\overline{\theta_N}\right) = P_\theta\ \left(y_k,\overline{y_{k+1}},\ldots,\overline{y_{k+N-1}},U_{\text{ref}};\Theta_p\right). \quad (31)$$

The model-based output prediction unit utilizes numerical computations based on the fault unified circuit model described by (15) to predict output. The mathematical formulation is as follows:

$$\left(\overline{y_{k+1}},\ldots,\overline{y_{k+N}},\right)=P_y\ \left(y_k,\ldots,y_{k+N-1},U_{ref};\overline{\theta_1},\ldots,\overline{\theta_N}\right). \quad (32)$$

The fault classification unit is a CNN network commonly employed in time-domain signal fault diagnosis. Its mathematical model is expressed as follows:

$$\bar{l} = P_c\ \left(\overline{\theta_1},\ldots,\overline{\theta_N};\Theta_c\right). \quad (33)$$

Θ_p and Θ_c represent the model parameters to be trained

$$\Theta_p,\ \Theta_c = \underset{\Theta_p,\Theta_c}{\operatorname{argmin}}\ \lambda_1\text{loss1}\ +\lambda_2\text{loss2}+\lambda_3\text{loss3} \quad (34)$$

where loss1 ensures that the predicted output closely matches the actual output, loss2 ensures that the predicted fault parameters approximate the actual fault parameters, and loss3 ensures that the predicted fault label approximates the actual fault label. λ represents the weight coefficient. For more detailed training specifics, refer to paper [27].

Remark 1: The proposed MENN predicts the nonlinear variation of fault parameters by forecasting these parameters across different time segments. The neural network's strong nonlinear modeling capability ensures high accuracy in fault parameter prediction. The identified fault patterns correspond closely to real physical processes, facilitating model generalization and reducing the need for extensive training samples. Furthermore, fault parameter variations can be directly measured with instruments, providing interpretability for fault diagnosis decisions. The training process is guided by a model-based output prediction unit, as defined in (34) (loss1), which allows the use of extensive numerical data from (25) to train the CNN-based parameter predictor. Hence, this approach mitigates the reliance on large datasets typically required by purely data-driven models.

V. Developments and Validations

In this article, a real system database and a numerical calculation simulation database based on (25) are constructed to compare the fault diagnosis performance of the proposed method with advanced techniques. These techniques include data-driven methods such as BN [26], RFs [25], 1D-CNN [17], LSTM [18], CL [32]. Additionally, to analyze the validity of the MENN structure, the MENN1 method is introduced, which replaces the CNN-based parameter prediction model with model-based parameter identification (least squares method), based on the proposed MENN.

BN and RFs: Fault features are extracted form input signals by fast Fourier transform. According to the fault features output fault labels to locate faults. 1D-CNN, LSTM, and CL obtain the fault label directly from the input signal. The training principle of the comparison method is the same as that of the literature.

A. Datasets Setup

1) Real System Database: A physical system of a single-phase full-bridge PWM converter, as shown in Fig. 5, is built to simulate faults and construct a real system database. The system comprises power supply, electronic load, main power circuit of converter, LCR tester, control and measurement instruments. The main parameters of the physical system are detailed in Table VI. The power supply and electronic load are used to simulate multiple operating conditions of the converter in actual operation. LCR tester is used to verify the parameter identification accuracy of the proposed MENN. Measurement instruments are used to test and display fault data.

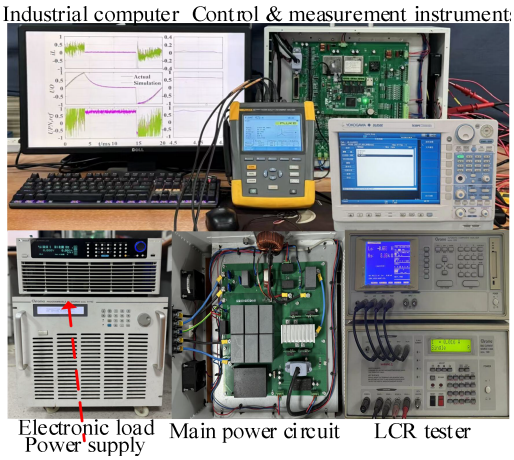

Fig. 5. Physical system of single-phase full-bridge PWM converter.

To construct a rich database, the following device fault type are simulated: single/double open-circuit faults, wire bonding faults (1/8, 2/8, …, 7/8) for IGBTs T_1–T_4, inductor turns short circuit faults (10%, 20%, 30%), and capacitor aging fault with a decrease in capacitance (10%, 15%, 20%, 25%). These failures encompass a total of 23 fault types across 3 key power devices. Each fault type is simulated by replacing the corresponding faulty device. During the fault data collection process, the following operating conditions are considered [17]: output voltage amplitude, output voltage frequency, output load power, dc bus voltage amplitude, and dc bus voltage ripple. An equal number of fault samples are collected for each fault type, resulting in a total of 160 samples per fault type, culminating in 23∗160 fault samples for real system database.

2) Numerical Calculation Simulation Database: The simulation database is derived from numerical calculations using the fault equivalent circuit model (25). The circuit parameters and sample acquisition process are identical to those used in the real system database. This process results in a total of 512 samples per fault type, amounting to 23∗512 fault samples for the simulation database. Fault device parameters in (25), L_{NF},R_{LNF},C_{NF},R_{CNF}, R_{PN}, are obtained from fitting data collected from device fault experiments.

3) Nonlinear Fitting of Fault Parameters: The fault parameters in (25) are derived from experimental data fitting. The experimental results for IGBT equivalent on-resistance R_{PN} and equivalent inductance resistance R_{LNF}, which exhibit the most significant nonlinear changes, are shown in Fig. 6. These results indicate that without fitting the multiple time interval nonlinear changes of fault parameters, a parameter confusion area (highlighted in the orange shaded part of the Fig. 6) may occur, leading to potential misdiagnosis. Therefore, the proposed MENN improves the fault diagnosis performance by fitting the multiple time interval nonlinear change of fault parameters.

4) Fault Equivalent Circuit Model Analysis: The proposed method is based on the fault equivalent circuit model (25), so it is necessary to verify the accuracy of the model. According to Table IV, the input signal of fault diagnosis of Case 3 is used as measured output characteristic. The fault amplitude

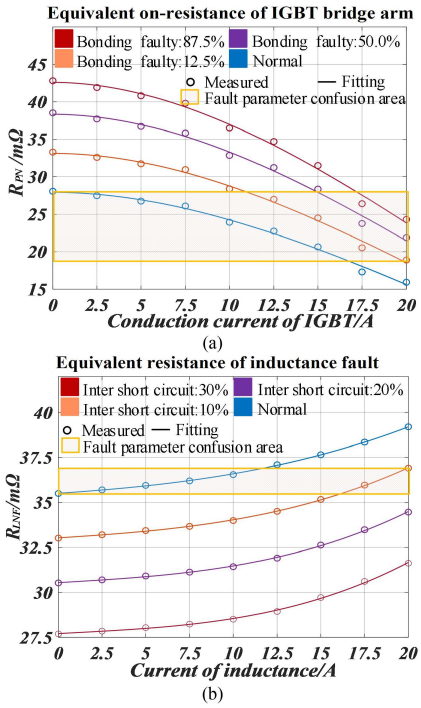

Fig. 6. Most significant nonlinear changes of fault device parameters. (a) IGBT wire bonding fault. (b) Inductor turns short-circuit fault.

normalization waveforms of the device types in both the real system database and the simulation database are shown in Fig. 7. Fig. 7 indicates that the fault equivalent circuit model effectively reflects the changing trends of the output characteristics when actual faults occur. Additionally, the distinct fault waveforms of different fault types show that the output characteristics of various fault types are very similarity, especially the output voltage U_o. Consequently, traditional methods struggle to simultaneously diagnose multiple device types. It is worth noting that through the proposed model-based fault detectability analysis, different fault types in some output characteristics are distinguishable, particularly in i_L and $U_{PN,\mathrm{ref}}$. This highlights the important role of fault detectability analysis in diagnosing multiple device types.

B. Fault Detectability Verification Experiment

1) Verification of Fault Detectability Analysis Results: In order to verify the correctness of the proposed model-based fault detectability analysis method, 70% of database is randomly sampled to train the different methods based on the analysis results in Table IV, with the remaining 30% of the real system database used as test data. This experiment is repeated ten times, and the results are shown in Fig. 8. The experimental results indicate that the fault diagnosis accuracy for diagnosable Cases 1–4 exceeds 90% across different methods, while the accuracy for undiagnosable Cases 5–10 is lower than 65%. These findings are consistent with the results of the proposed fault detectability analysis method in Table IV, verifying the correctness of the proposed method.

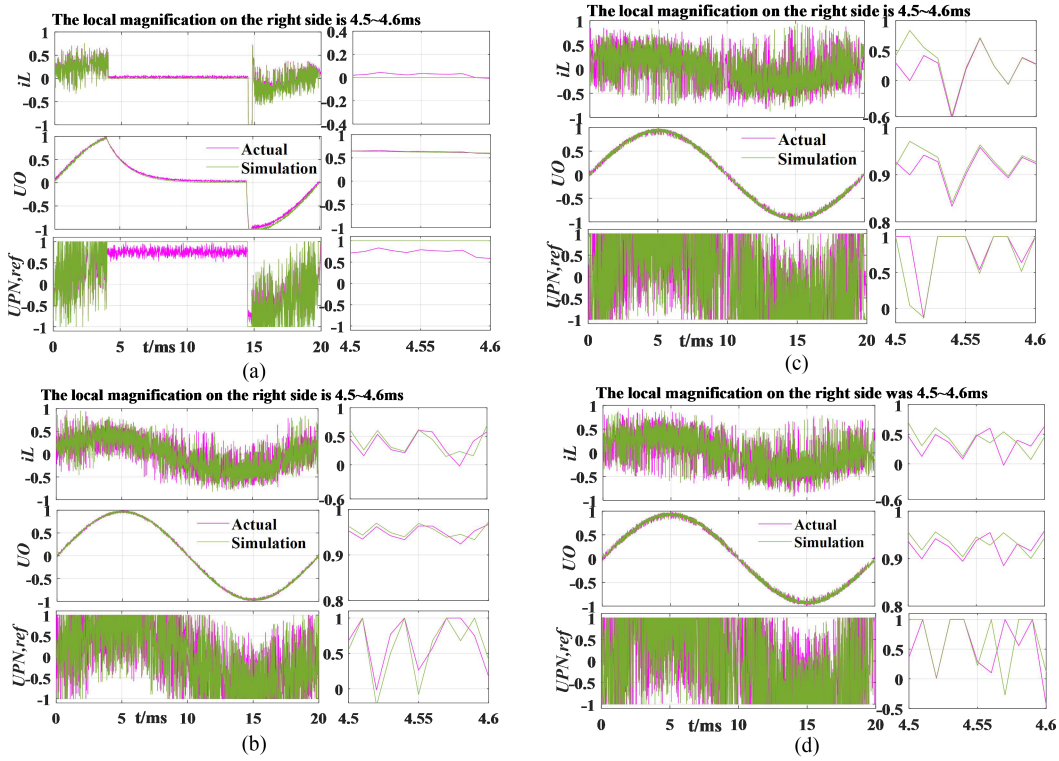

Fig. 7. Fault waveforms of the device types in both the real system database and simulation database. (a) $\mathcal{F}^1$ IGBT open-circuit fault. (b) $\mathcal{F}^2$ IGBT wire bonding fault 50.0%. (c) $\mathcal{F}^3$ Inductor turns short-circuit fault 30%. (d) $\mathcal{F}^4$ Capacitor aging fault 20%.

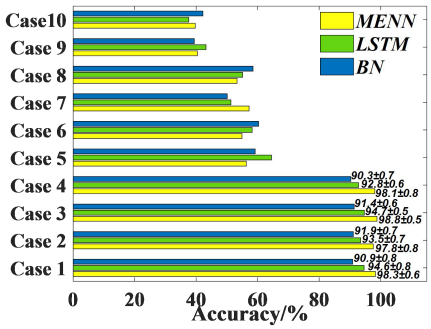

Fig. 8. Average fault diagnosis accuracy of different cases in the fault detectability analysis results (see Table IV).

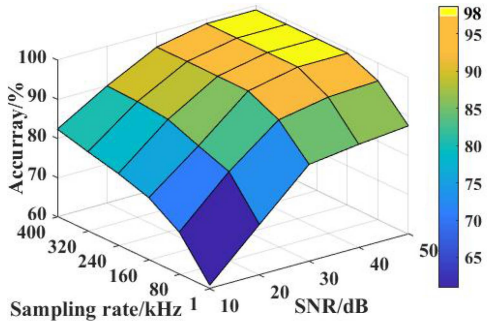

Fig. 9. Average fault diagnosis accuracy results under different measurement conditions (sampling rate and SNR).

2) Verification of Fault Detectability Degree Analysis Results: According to the fault detectability analysis results of Case 3 in Table IV, 70% of the database is randomly sampled to train the model, and the remaining 30% is used as test data. The average fault diagnosis accuracy from ten experiments under different measurement conditions (sampling rate and SNR) is shown in Fig. 9. Fig. 9 illustrates that when the sampling rate exceeds 160 kHz and the SNR is above 40 dB, the fault diagnosis accuracy of the proposed method exceeds 98%. This closely matches the measurement conditions (165 kHz and 45 dB in Table V) determined by the proposed fault detectability degree analysis method, validating its accuracy. Since the proposed fault detectability analysis method does not require real fault sample collection and model training, it significantly reduces the cost associated with model training while ensuring reliable diagnosis of multidevice equipment failures.

C. Fault Diagnosis Performance Evaluation

1) Fault Diagnosis Accuracy Analysis: To evaluate the fault diagnosis accuracy of the proposed method, the same fault measurement conditions in Table V are employed. 70% of the real system database and simulation database are randomly allocated for model training, while the remaining 30% is reserved for testing model accuracy. The average results from ten experiments are calculated. Fig. 10 illustrates the proposed MENN training process, demonstrating strong convergence within the

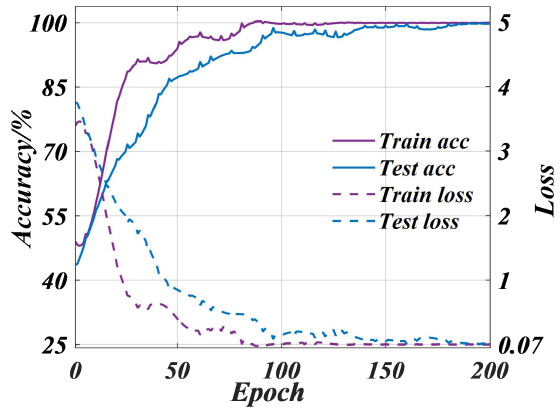

Fig. 10. Model training process of the proposed MENN.

TABLE VII
AVERAGE DIAGNOSIS ACCURACY OF DIFFERENT METHODS

Method	Precision	Accuracy	Recall	Parameters
BN [26]	90.53%	91.62%	91.13%	0.64M
RFs [25]	90.33%	91.42%	91.58%	1.56M
1D-CNN [17]	92.65%	93.73%	92.16%	8.92M
LSTM [28]	93.27%	93.64%	93.75%	13.62M
CL [32]	92.97%	93.77%	92.67%	9.82M
MENN1	95.73%	95.49%	95.59%	1.62M
MENN	98.23%	98.79%	98.59%	4.26M

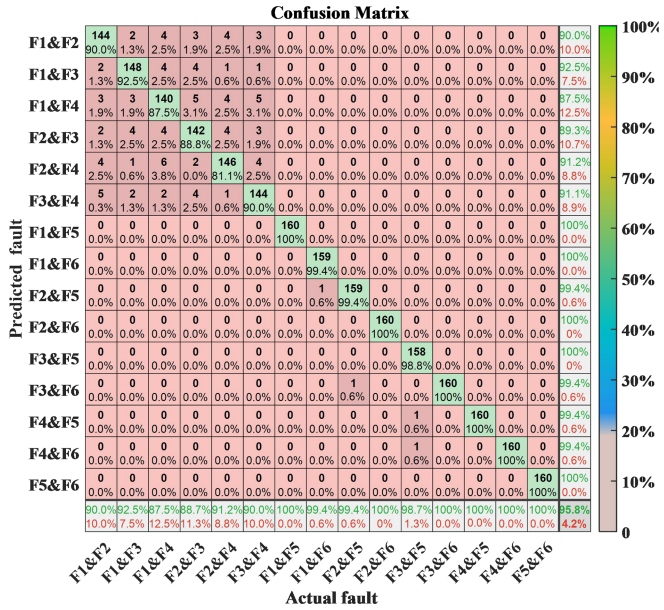

Fig. 11. Confusion matrix of simultaneous faults in multiple devices by the proposed MENN. ∗Remark: In the first row and first column of the confusion matrix, the black number "144" indicates that 144 out of the 160 actual fault samples were correctly predicted as F_1&F_2, corresponding to a 90% prediction accuracy. The green percentage values in the last column represent the recall rates, while the red values indicate the error rates. The green percentages in the last row denote the precision for each fault type. The green "95.8%" in the bottom-right corner indicates the overall accuracy of the proposed MENN.

training dataset. The model also exhibits excellent generalization performance on the test dataset. Therefore, the proposed MENN can identify the fault parameters effectively.

The average results from ten experiments using different methods are presented in Table VII. All methods achieved fault diagnosis accuracy above 90%, demonstrating that reliable multiple device fault diagnosis can be accomplished through the proposed fault detectability analysis method. The proposed MENN captures multiple time interval changes in fault parameters over time and uses these changes as fault features for fault diagnosis. Since variations in fault parameters follow a consistent tendency with the actual physical process during fault occurrence, the MENN exhibits strong generalization capabilities, achieving the highest fault diagnosis accuracy index in Table VII. The proposed method leverages model-based prior knowledge for feature extraction, resulting in a simpler network structure compared to data-driven feature extraction methods. Additionally, the parameter predictions at different time steps in MENN share the same CNN-based parameter predictor, which further reduces redundancy. Consequently, the proposed method requires fewer parameters (4.26M) and has lower model complexity than conventional deep learning approaches. This efficiency makes it suitable for real-time applications and resource-constrained environments.

2) Simultaneous Faults in Multiple Devices: To validate the proposed methods capability for diagnosing simultaneous faults in multiple devices, this experimental evaluation is conducted. The bond wire faults of IGBT, along with inductor and capacitor faults, are labeled as F_1–F_6. These faults are paired to create 15 distinct simultaneous fault types. Following the previously described methodology, samples are collected, with 160 fault samples obtained for each type, resulting in a total of 160×15 samples. Of these, 70% are used to train the model, while the remaining samples are employed to test the model's fault diagnosis accuracy. The confusion matrix for the test results is shown in Fig. 11. The proposed method demonstrates robust fault diagnosis accuracy (95.8%) even in the presence of simultaneous faults across two devices. This is attributed to the model's ability to localize faults based on individual fault parameters, which remain independent of each other. As a result, simultaneous faults across multiple devices do not lead to severe parameter confusion. However, when two bond wire faults occur simultaneously between IGBTs, some misdiagnosis occurs. This is because the fault parameters $U_{PN,\text{ref}}$ and R_{PN} are jointly determined by switching states and multiple IGBT fault states, as described by (15) and (21). These parameters exhibit time-varying behavior and are coupled with different IGBT fault conditions. The proposed method mitigates the risk of misdiagnosis by identifying the temporal patterns of $U_{PN,\text{ref}}$ and R_{PN}, leveraging their differences across time sequences to accurately localize faults.

3) Switching Frequency and Load Effect Analysis: To verify the applicability of the proposed method under varying switching frequencies and load conditions, offline fault data corresponding to these scenarios are used as the test dataset. The samples in the test dataset are evenly divided into 36 subtest sets based on switching frequencies ranging from 1 to 16 kHz

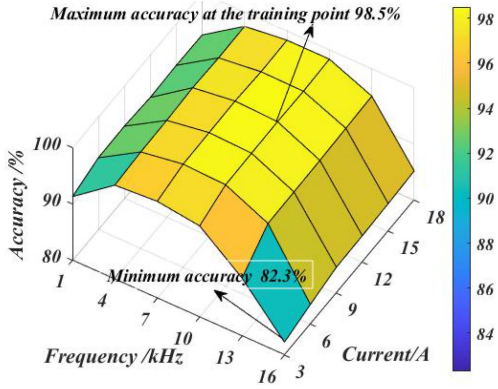

Fig. 12. Impact of switching frequency and load variations on the fault diagnosis accuracy of the proposed method.

TABLE VIII
CIRCUIT PARAMETERS OF THE FAULT DEVICE MEASURED BY LCR TESTER

Device parameter	Fault value	Fault type
R_{PN}	32.7 mΩ	IGBT wire bonding fault 50.0%
L_{NF}/ R_{LNF}	18.6 μH/28.8 mΩ	Inductor turns short circuit 30%
C_{NF} /R_{CNF}	125.4 μH/0.307 Ω	Capacitor aging fault 20%

(in 3 kHz increments) and load currents ranging from 3 to 18 A (in 3 A increments). The fault diagnosis accuracy across these subtest sets is shown in Fig. 12. The results indicate that the proposed method exhibits adaptability to both switching frequency and load current variations. This adaptability is achieved by modeling the transient switching behavior in the parameter prediction model (25) and accounting for the impact of current on component parameters, as shown in Fig. 7. However, fault diagnosis performance declines at higher switching frequencies. This is because the optimal sampling rate of 160 kHz, determined through the detectability analysis at the nominal switching frequency of 10 kHz, becomes insufficient to capture high-frequency fault information at elevated switching frequencies. Additionally, a slight reduction in diagnostic accuracy is observed at lower switching frequencies and smaller load currents. This occurs because reduced switching frequency and current degrade the control performance of the output voltage, affecting its quality and leading to occasional misdiagnoses. Nevertheless, these performance declines do not significantly affect the method's adaptability, as converters typically operate outside ranges that impair output performance. Furthermore, the decline in high-frequency fault detection can be mitigated by optimizing the sampling rate using the detectability analysis proposed in this paper.

4) Decision Interpretability Analysis: The decision interpretability of the proposed method depends on the accuracy of fault parameter identification. Hence, to verify the nonlinear fault parameter fitting ability of MENN and MENN1, the circuit parameters of the faulty device measured using an LCR tester are presented in Table VIII. The error rates of fault parameters identified by MENN and MENN1 are shown in Table IX. A comparison between MENN1 and MENN reveals that adjusting

TABLE IX
ERROR RATES OF FAULT PARAMETERS IDENTIFIED BY MENN AND MENN1

Identification error rate %	Method: MENN /MENN1				
	$\mathcal{F}^2$	$\mathcal{F}^3$	$\mathcal{F}^4$	$\mathcal{F}^0$	Average
R_{PN}	**1.2**/1.8	**1.0**/1.5	**1.0**/1.5	**1.0**/1.5	**1.05**/1.58
R_{LNF}	**1.1**/1.6	**1.3**/2.0	**1.1**/1.6	**1.1**/1.6	**1.15**/1.70
L_{NF}	**1.3**/1.9	**1.5**/2.2	**1.3**/1.9	**1.3**/1.8	**1.35**/1.95
R_{CNF}	**0.4**/0.7	**0.4**/0.7	**0.5**/0.8	**0.4**/0.7	**0.43**/0.73
C_{NF}	**0.5**/0.8	**0.5**/0.8	**0.6**/0.9	**0.5**/0.8	**0.53**/0.83
Average	**0.9**/1.4	**0.9**/1.4	**0.9**/1.3	**0.9**/1.3	**0.90**/1.36

TABLE X
COMPARISON RESULTS OF ERROR RATE OF FAULT PARAMETER IDENTIFICATION

Methods	Average Identification error rate %				
	R_{PN}	R_{LNF}	L_{NF}	R_{CNF}	C_{NF}
DSSM [30]	6.16	6.42	6.86	5.43	4.91
PILM [23]	5.07	5.26	5.52	3.94	3.86
MENN1	1.58	1.70	1.95	0.73	0.83
MENN	1.05	1.15	1.35	0.43	0.53

the CNN-based parameter prediction model structure according to actual measurement data (as done in MENN) provides a more accurate method for fitting nonlinear fault parameters than solely relying on model parameter identification (as done in MENN1). Consequently, MENN achieves better fault diagnosis accuracy, as shown in Table VII. Furthermore, Table IX demonstrates that the proposed method accurately identifies parameters under multiple fault condition. Because the fault location law derived from the change in fault parameters can be measured by LCR tester and applied to any similar converter scenario, the fault location results of the proposed MENN are interpretable and reliable, under the guarantee of high-precision parameter identification.

5) Comparative Analysis of Parameter Identification Accuracy: To compare and analyze the accuracy of fault parameter identification, parameter identification models from papers [23] and [30] are implemented based on the circuit model in (25). The identified fault parameters are compared with the measurements obtained using an *LCR* meter, and the average identification error rates are summarized in Table X. The proposed MENN shows superior accuracy in identifying IGBT bond wire faults, inductance, and capacitance parameters. This is because the proposed approach constrains the neural network training process using a mechanism-based model, mitigating the inaccuracy issues caused by insufficient samples. The integration of data-driven methods further addresses the inaccuracies associated with mechanism model-based parameter estimation. Additionally, the optimized measurement conditions, achieved through fault detectability analysis, ensure parameter identifiability. As a result, the proposed method offers this advantage in fault parameter identification accuracy.

6) Sample Dependence Analysis: To assess the fault sample dependence of the proposed MENN, various proportions of the real system dataset and 70% simulation dataset are randomly selected as training data, with the remaining dataset used for testing. The average fault diagnosis accuracy of different proportions of real system samples from ten experiments is depicted in Fig. 13. The results show that the fault diagnosis accuracy of different methods increases with the proportion of actual

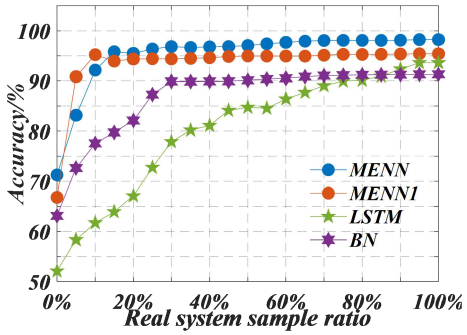

Fig. 13. Average fault diagnosis accuracy of different-proportion real system samples from ten experiments.

system samples. However, the proposed method maintains high fault diagnosis accuracy even with a small number of real system samples. Specifically, when training with only 10% real system dataset and 70% simulation dataset, the fault diagnosis accuracy remains above 92%. This is because the proposed method constrains the CNN training process through the fault equivalent circuit model (25), allowing a large amount of numerical calculation data to assist in the CNN training process (loss1), thus significantly reducing the dependence on actual system fault samples. Additionally, as previously analyzed, the generalization ability of fault parameter identification further reduces the need for samples in the fault classifier.

D. On-Line Fault Diagnosis Experiment

To assess the real-time fault diagnosis performance of the proposed method, the trained MENN model (Flash/RAM usage: 4.26/1.04 MiB) is deployed on the STM32H747I-disco platform to diagnose faults in a single-phase full-bridge PWM converter (see Fig. 5). The control system employs a DSP28377S with a sampling rate of 160 kHz and a control frequency of 10 kHz. The DSP captures the output characteristic $Y = [U_O, i_L, U_{PN,\text{ref}}]$, buffers them for one cycle, and transmits the data to the deployed fault diagnosis system. Upon fault detection, the system returns the fault parameters and labels to the DSP, which then relays the information to an industrial computer via a serial port for visualization.

The transient behavior of inductor fault diagnosis is shown in Fig. 14. At 0.04 s, a fault inductor is injected into the main power circuit through a relay triggered by a timer. The proposed method successfully identifies the inductance parameter and tracks its variation over time. Since the inductance parameter is influenced by the alternating current, the equivalent resistance R_{LNF} changes dynamically. Following the fault injection, parameters R_{LNF} and L_{NF} display significant shifts, and the method accurately detects the fault at about 0.06 s. These findings confirm that the proposed method enables fast and precise real-time inductance fault diagnosis by accurately identifying and monitoring fault parameters.

The transient behavior of capacitor fault diagnosis is shown in Fig. 15. At 0.04 s, a fault capacitor is injected into the

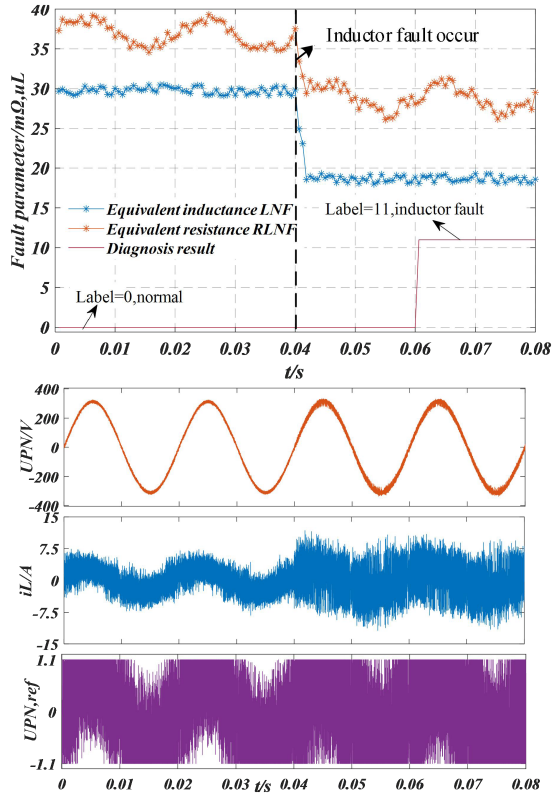

Fig. 14. Inductor online fault diagnosis experiment.

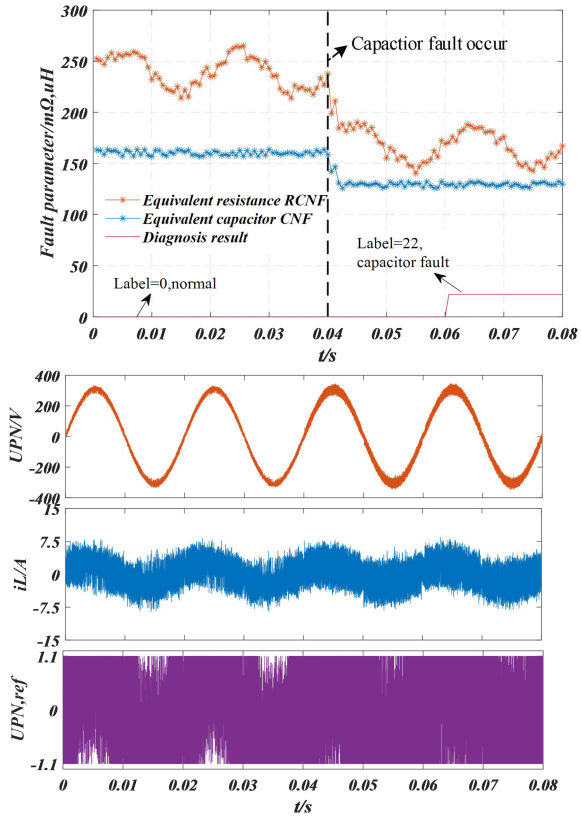

Fig. 15. Capacitor online fault diagnosis experiment.

TABLE XI
COMPARISON OF THE PERFORMANCE OF THE RELEVANT METHODS IN THE MULTIDEVICE FAULT DIAGNOSIS EXPERIMENTS

Methods	Diagnostic accuracy	Decision explainability	Sample dependence
BN [26]	Low	No	Medium
RFs [25]	Low	No	Medium
1D-CNN [17]	Medium	No	Severe
LSTM [28]	Medium	No	Severe
CL [32]	Medium	No	Severe
DSSM [30]	Medium	Yes	Slight
PILM [23]	Medium	Yes	Slight
Proposed Method	High	Yes	Slight

main power circuit through a relay triggered by a timer. The proposed method successfully identifies the capacitor parameter and tracks its variation over time. Since the capacitor parameter is influenced by the alternating current, the equivalent resistance R_{CNF} changes dynamically. Following the fault injection, parameters R_{CNF} and C_{NF} display significant shifts, and the method accurately detects the fault at about 0.06 s. These findings confirm that the proposed method enables fast and precise real-time capacitor fault diagnosis by accurately identifying and monitoring fault parameters.

E. Discussions and Potentials

The performance comparison of various fault diagnosis methods used in the multidevice fault diagnosis experiment for converters is summarized in Table XI. Comparisons with state-of-the-art techniques confirm the proposed method's superiority in terms of diagnostic accuracy, decision explainability, and reduced sample dependence. This superiority is attributed to several key factors.

1) Enhanced diagnostic accuracy is achieved through model-based fault detectability analysis, which identifies the essential measurement conditions required for reliable fault diagnosis.
2) Time-series fault parameter identification enables decision interpretability by leveraging high-precision parameter estimation to support accurate fault localization. This time-series fault parameter identification approach ensures that the temporal dynamics of fault parameters are captured, allowing the proposed method to distinguish faults more effectively over time.
3) The CNN-based parameter predictor is trained with model-constrained output predictions, which significantly reduces the need for large sample datasets, ensuring efficient learning even with limited data.

By integrating model-based and data-driven approaches, the proposed method overcomes the challenges of high-precision modeling and extensive data requirements, making it suitable for scenarios with known circuit fault modes and limited fault data. Notably, increased converter complexity has minimal impact on computational efficiency, as it primarily depends on the order of the state equation in model (25). Although the proposed method shows promising results, several open challenges remain.

1) Optimal Time Window Selection: The model complexity and fault diagnosis speed are affected by the length of the time window used for fault detection. Considering the periodic nature of fault characteristics, a 0.02 s window is adopted in this study. However, determining the optimal time window remains an open question. How can the ideal balance between diagnostic accuracy and computational efficiency be achieved?
2) Model Discovery under Structural Uncertainty: How can data-driven methods be employed to discover unknown model structures when the system model is uncertain?
3) Conditions for Fault Sample Collection: What operating conditions are necessary to ensure the collection of effective fault samples?

VI. CONCLUSION

The data-driven fault diagnosis method, which eliminates the need for additional sensors while preserving the system's original structure, offers a promising approach to enhancing converter reliability. However, diagnosing multiple device faults presents significant challenges due to difficulties in fault sample acquisition, detectability uncertainty, and unexplained decisions. This article addresses these challenges by presenting a detectability-based data-driven fault diagnosis method with fault detectability analysis and a mechanism-enhanced neural network. First, fault detectability is defined based on the independence of the fault parameter value space. It is then proved that fault parameter identification is a sufficient condition for fault parameter detection. Based on the property of parameter identification, a calculable fault diagnosability analysis method is derived to determine which output characteristics can be measured to achieve device fault diagnosis. Considering the influence of sampling rate and SNR on fault detectability, a fault detectability degree analysis method is proposed to optimize the measurement conditions.

Utilizing these measurement condition constraints, fault detectability and fault parameter identifiability are guaranteed theoretically. Hence, a mechanism-enhanced neural network can be designed to locate faults by fitting the changes in fault parameters. The consistency between the fitting fault parameters and the actual fault process ensures the interpretability of the diagnosis results. Additionally, by guaranteeing the identification of fault parameters, the fault equivalent circuit model assists in training, significantly reducing the number of actual fault samples required for mechanism-enhanced neural network training.

Finally, experiments on a representative converter are conducted to verify the effectiveness of the proposed method. Comparisons with state-of-the-art techniques show the proposed scheme's superiority in terms of diagnostic accuracy, decision explainability, and sample dependence.

REFERENCES

[1] J. Stuyts, G. Horn, W. Vandermeulen, J. Driesen, and M. Diehl, "Effect of the electrical energy conversion on optimal cycles for pumping airborne wind energy," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 2–10, Jan. 2015.

[2] J. Hang, W. Sun, Q. Hu, X. Ren, and S. Ding, "Integration of interturn fault diagnosis and fault-tolerant control for PMSM drive system," *IEEE Trans. Transp. Electrific.*, vol. 8, no. 2, pp. 2825–2835, Jun. 2022.

[3] W. Ni et al., "Study on optimization of passenger flow at a metro station based on any logic—Case study of Youfangqiao station of Nanjing metro line 2," *Complex Syst. Model. Simul.*, vol. 1, no. 3, pp. 242–252, 2021.

[4] M. Yaghoubi, J. S. Moghani, N. Noroozi, and M. R. Zolghadri, "IGBT open-circuit fault diagnosis in a quasi-Z-source inverter," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 2847–2856, Apr. 2019.

[5] Y. Fassi, V. Heiries, J. Boutet, and S. Boisseau, "Towards physics-informed machine learning-based predictive maintenance for power converters–A review," *IEEE Trans. Power Electron.*, vol. 39, no. 2, pp. 2692–2720, Feb. 2024.

[6] J. Liang et al., "A state-of-the-art review on wind power converter fault diagnosis," *Energy Rep.*, vol. 8, pp. 5341–5369, 2022.

[7] W. Zhang et al., "In situ diagnosis of wire bonding faults for multichip IGBT modules based on the crosstalk effect," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 10, no. 5, pp. 5107–5117, Oct. 2022.

[8] H. Li, D. Xiang, X. Han, X. Zhong, and X. Yang, "High-accuracy capacitance monitoring of DC-link capacitor in VSI systems by LC resonance," *IEEE Trans. Power Electron.*, vol. 34, no. 12, pp. 12200–12211, Dec. 2019.

[9] W. Huang et al., "Current-based open-circuit fault diagnosis for PMSM drives with model predictive control," *IEEE Trans. Power Electron.*, vol. 36, no. 9, pp. 10695–10704, Sep. 2021.

[10] D. Peng, H. Wang, Z. Liu, W. Zhang, M. J. Zuo, and J. Chen, "Multibranc and multiscale CNN for fault diagnosis of wheelset bearings under strong noise and variable load condition," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4949–4960, Jul. 2020.

[11] B. Li et al., "Application of artificial neural networks to photovoltaic fault detection and diagnosis: A review," *Renewable Sustain. Energy Rev.*, vol. 138, 2021, Art. no. 110512.

[12] D. K. Kulevome et al., "A bidirectional LSTM-based prognostication of electrolytic capacitor," *Prog. Electromagn. Res. C*, vol. 109, pp. 139–152, 2021.

[13] S. H. Jeong, J. W. Park, and H. S Kim, "Deep neural network-based lifetime diagnosis algorithm with electrical capacitor accelerated life test," *J. Power Sources*, vol. 599, 2024, Art. no. 234182.

[14] X. Wang, B. Jiang, S. X. Ding, N. Lu, and Y. Li, "Extended relevance vector machine-based remaining useful life prediction for DC-link capacitor in high-speed train," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9746–9755, Sep. 2020.

[15] J. Zhao et al., "A remaining useful life prediction method of aluminum electrolytic capacitor based on wiener process and similarity measurement," *Microelectron. Rel.*, vol. 142, 2023, Art. no. 114928.8.

[16] A. Moradzadeh, B. Mohammadi-Ivatloo, K. Pourhossein, and A. Anvari-Moghaddam, "Data mining applications to fault diagnosis in power electronic systems: A systematic review," *IEEE Trans. Power Electron.*, vol. 37, no. 5, pp. 6026–6050, May 2022.

[17] W. Yuan, Z. Li, Y. He, R. Cheng, L. Lu, and Y. Ruan, "Open-circuit fault diagnosis of NPC inverter based on improved 1-D CNN network," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 3510711.

[18] Y. Si, R. Wang, S. Zhang, W. Zhou, A. Lin, and Y. Wang, "Fault diagnosis based on attention collaborative LSTM networks for NPC three-level inverters," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 3512416.

[19] J. Hang, X. Shu, S. Ding, and Y. Huang, "Robust open-circuit fault diagnosis for PMSM drives using wavelet convolutional neural network with small samples of normalized current vector trajectory graph," *IEEE Trans. Ind. Electron.*, vol. 70, no. 8, pp. 7653–7663, Aug. 2023.

[20] S. Yang, D. Xiang, A. Bryant, P. Mawby, L. Ran, and P. Tavner, "Condition monitoring for device reliability in power electronic converters: A review," *IEEE Trans. Power Electron.*, vol. 25, no. 11, pp. 2734–2752, Nov. 2010.

[21] G. Yating, W. Wu, L. Qiongbin, C. Fenghuang, and C. Qinqin, "Fault diagnosis for power converters based on optimized temporal convolutional network," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 3501910.

[22] G. E. Karniadakis et al., "Physics-informed machine learning[J]," *Nature Rev. Phys.*, vol. 3, no. 6, pp. 422–440, 2021.

[23] S. Zhao, Y. Peng, Y. Zhang, and H. Wang, "Parameter estimation of power electronic converters with physics-informed machine learning," *IEEE Trans. Power Electron.*, vol. 37, no. 10, pp. 11567–11578, Oct. 2022.

[24] K. Hu, Z. Liu, Y. Yang, F. Iannuzzo, and F. Blaabjerg, "Ensuring a reliable operation of two-level IGBT-based power converters: A review of monitoring and fault-tolerant approaches," *IEEE Access*, vol. 8, pp. 89988–90022, 2020.

[25] F. Wu, K. Chen, G. Qiu, and W. Zhou, "Robust open circuit fault diagnosis method for converter using automatic feature extraction and random forests considering nonstationary influence," *IEEE Trans. Ind. Electron.*, vol. 71, no. 10, pp. 13263–13273, Oct. 2024.

[26] B. Cai, Y. Zhao, H. Liu, and M. Xie, "A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems," *IEEE Trans. Power Electron.*, vol. 32, no. 7, pp. 5590–5560, Jul. 2017.

[27] J. F. M. Van Doren et al., "Identifiability: From qualitative analysis to model structure approximation," *IFAC Proc. Volumes*, vol. 42, no. 10, pp. 664–669, 2009.

[28] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.

[29] P. Suskis et al., "Converter state-space model estimation using dynamic mode decomposition," in *Proc. IEEE 7th Int. Energy Conf.*, 2022, pp. 1–5.

[30] T. Kimura and H. Maruta, "Parameter estimation of DC-DC converters for failure detection based on linear approximation model," in *Proc. IEEE 30th Int. Symp. Ind. Electron*, 2021, pp. 1–6.

[31] A. B. Kareem and J.-W. Hur, "Towards data-driven fault diagnostics framework for SMPS-AEC using supervised learning algorithms," *Electronics*, vol. 11, no. 16, 2022, Art. no. 2492.

[32] J. Miao, Y. Liu, Q. Yin, B. Ju, G. Zhang, and H. Wang, "A novel soft fault detection and diagnosis method for a DC/DC buck converter based on contrastive learning," *IEEE Trans. Power Electron.*, vol. 39, no. 1, pp. 1501–1513, Jan. 2024.

[33] A. V. D. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," 2018, *arXiv:1807.03748*.

[34] F. Fu, T. Xue, Z. Wu, and D. Wang, "A fault diagnosability evaluation method for dynamic systems without distribution knowledge," *IEEE Trans. Cybern.*, vol. 52, no. 6, pp. 5113–5123, Jun. 2022.

[35] J. C Butcher, "A history of Runge-Kutta methods," *Appl. Numer. Math.*, vol. 20, no. 3, pp. 247–260, 1996.