

Article

A Tutorial on Hardware-Implemented Fault Injection and Online Fault Diagnosis for High-Speed Trains

Xiaoyue Yang ¹, Xinyu Qiao ², Chao Cheng ³, Kai Zhong ⁴ and Hongtian Chen ^{5,*}¹ School of Rail Transportation, Wuyi University, Jiangmen 529020, China; xyueoh@outlook.com² School of Astronautics, Harbin Institute of Technology, Harbin 150001, China; qiaoxinyu2018@126.com³ School of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China; chengchao@ccut.edu.cn⁴ Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601, China; kaizhong0402@ahu.edu.cn⁵ Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada

* Correspondence: chtbaylor@163.com; Tel.: +86-1326-095-0111

Abstract: Electrical drive systems are the core of high-speed trains, providing energy transmission from electric power to traction force. Therefore, their safety and reliability topics are always active in practice. Among the current research, fault injection (FI) and fault diagnosis (FD) are representative techniques, where FI is an important way to recur faults, and FD ensures the recurring faults can be successfully detected as soon as possible. In this paper, a tutorial on a hardware-implemented (HIL) platform that blends FI and FD techniques is given for electrical drive systems of high-speed trains. The main contributions of this work are fourfold: (1) An HIL platform is elaborated for realistic simulation of faults, which provides the test and verification environment for FD tasks. (2) Basics of both the static and dynamic FD methods are reviewed, whose purpose is to guide the engineers and researchers. (3) Multiple performance indexes are defined for comprehensively evaluating the FD approaches from the application viewpoints. (4) It is an integrated platform making the FI and FD work together. Finally, a summary of FD research based on the HIL platform is made.

Keywords: fault injection (FI); fault diagnosis (FD); hardware-implemented (HIL) platforms; electrical drive systems; high-speed trains



Citation: Yang, X.; Qiao, X.; Cheng, C.; Zhong, K.; Chen, H. A Tutorial on Hardware-Implemented Fault Injection and Online Fault Diagnosis for High-Speed Trains. *Sensors* **2021**, *21*, 5957. <https://doi.org/10.3390/s21175957>

Academic Editors: Steven Chatterton and Lang Xu

Received: 9 August 2021

Accepted: 2 September 2021

Published: 5 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past five decades, the rapid development of high-speed trains relying on multiple electrical traction units has been witnessed [1–3]. Nowadays, high-speed trains have become one of the most important transportation means thanks to their inherent advantages such as rapidity, comfort, and high efficiency [4]. However, unexpected railway accidents usually happen all over the world because of external and internal factors [5–7]; for example, design defects of high-speed trains and limits of usage life will induce the faults appearing in trains and may further lead to catastrophic accidents [8–10]. To the best of our knowledge, electrical drive systems can be viewed as the heart of high-speed trains, whose reliability and safety are crucially important to ensure the whole train operates safely [5]. Therefore, fault diagnosis (FD) for electrical drive systems is always an active topic from theoretical and practical researchers in the transportation field [6].

In order to investigate FD issues for electrical drive systems of high-speed trains, two key steps involved are: (1) *fault injection* (FI) which can replicate faults in a true manner [11], and (2) *FD* which can successfully and rapidly detect and diagnosis faults [5]. The first step can simulate various faults appearing in electrical drive systems on the one hand and provide sufficient data sets for fault analysis and experimental verification on the other hand [12]. Based on first principles or collected data sets from high-speed trains, to design reasonable and effective algorithms is the second step for online FD purposes [5].

Initially applied to the centralized systems, especially dedicated fault-tolerant computer architectures in the early 1970s, FI is widely used to evaluate the fault-tolerant performance and injects the artificial faults into the system of interest [13]. The FI design can speed up the occurrence of faults to observe the system changes and also provides a way for analyzing the efficiency of new fault FD mechanisms [14]. Up to now, there are three major categories of FI techniques: hardware-based, software-based and emulation-based [15]. The hardware-based FI method involving the extra hardware, which is specially designed, allows injecting faults into the target system. This technique works efficiently for large-scale-integration circuits but meanwhile introduces a high risk of damage for the target system. The software-based method is targeted to applications in operating systems, and it cannot inject faults into the inaccessible location of software. Unlike the former two methods, the detailed simulation model is crucial for the emulation-based method, which leads to a maximum amount of observability and controllability, whereas model development is time-consuming and inaccurate. These FI methods mainly concentrate on evaluating the dependability of microelectronics systems or software programs [16]. Nevertheless, apart from the microelectronics system, a traction control system (TCS) consists of the power electronic and mechanical parts [17].

Nowadays, the hardware-in-the-loop (HIL) technique is widely applied to the power system field, which requires the entire system to be modelled inside the real-time simulator and does not involve external interfacing or input/output. In fact, the primary goal of HIL simulation is testing of the physical controller (including the hardware design and control strategy). A real-time simulator needs to solve the model equations for a one-time step within the same time in a real-world clock. If the execution time for the simulation of the system is shorter than or equal to the selected time-step, the simulation is considered to be real-time. The study [18] presents an HIL platform to simulate the open-switch fault of one rectifier in TCS, which aims to verify the proposed FD method. In [12], a multiprocessor HIL-based FI is proposed for real-time simulation of faults in traction control systems, and a timing optimization method is proposed to deal with overruns induced by real-time FI.

Generally, the current FD methods for electrical drive systems of high-speed trains can be definitely divided into three categories: signal analysis-based, model-based and data-driven methods [5,19]. For signal analysis-based strategies, the difference between nominal and fault statuses reflected in time domain [20], frequency domain [21] and time-frequency domain [22], are desirable features used for the FD purposes. As pointed out in [6], the commonly used FD method in practice is the so-called human inspection, which belongs to signal analysis-based methods. Based on a well-established system model of electrical drive systems, model-based methods can effectively perform FD tasks via designing residual generators [23]. Most recently, data-driven FD methods for electrical drive systems of high-speed trains have been intensively developed and widely accepted because of their simplicity of design and ease of implementation [24].

In the previous studies, FI and FD are independently and separately designed, which poses inevitable difficulties in the validation of online FD algorithms. Motivated by the aforementioned observations, the main contributions of this paper are summarized as follows:

1. Introduce an integrated validation platform where FI and FD can work together in real time.
2. Base an evaluation system for FD systems where various performance indexes are defined.
3. Review the data-driven FD literature whose verification is based the designed HIL platform.

The rest of this study is organized as follows. Section 2 describes the preliminaries of electrical drive systems, typical fault types, and the objective pursued in this paper. Section 3 details the FI methodologies using hardware-implemented manners. Based on the existing FD research, Section 4 summarizes the FD algorithms together with imple-

mentation procedures. Section 5 concludes this paper with future work and promising research directions.

2. Background

In this section, preliminaries of electrical driven systems, associated with various faults, will be described, followed by the objectives this work is dedicated to.

2.1. Electrical Drive Systems of High-Speed Trains

In this study, the CRH2A-type high-speed train is taken into consideration in which four traction systems provide traction power of the entire train [4]. As presented in [2], the traction system is one core unit consisting of a three-level source inverter (VSI), four induction motors, a traction control unit (TCU), filters, etc. Its control strategy is space vector pulse-width-modulation (SVPWM), and the corresponding schematic diagram is depicted in Figure 1.

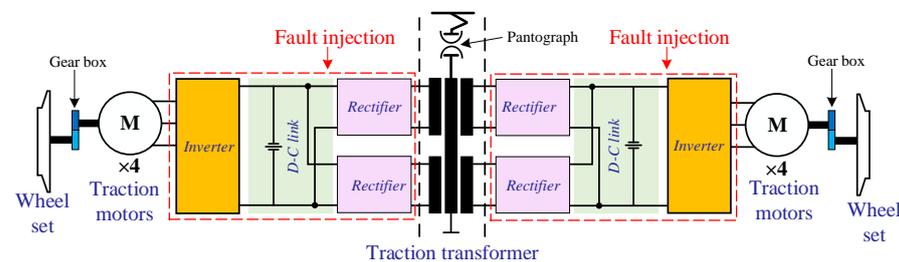


Figure 1. The schematic diagram of electrical drive systems in high-speed trains [24].

In the traction system of CRH2A-type high-speed trains, six sensors are equipped to collect real-time observations that will be used as the input of both the double proportional integral (PI) controller and the supervision unit. After setting a given traction speed, TCU can achieve expected performance by adjusting gate control signals of VSI based on the SVPWM strategy. By comparing online samplings and the corresponding pre-designed thresholds, the supervision unit adopted in the existing high-speed trains may not be effective for successful detection of parts' faults, and then an auto-protecting mode will be activated [2]. Actually, either high missing alarm ratios (MARs) or high false alarm ratios (FARs) in traction systems of high-speed trains will lead to unsatisfactory FDD results and should be prohibitive. An acceptable trade off between MAR and FAR should be at least achieved for FDD in traction systems of high-speed trains.

2.2. Fault Types

For the electrical drive systems of high-speed trains, there are different ways to judge which category one fault belongs to. For example, fault types can be determined depending upon fault amplitude [24], fault location [11] and fault-duration time [25].

According to the fault location, faults in electrical drive systems could be divided into the following four scenarios:

1. Sensor faults: Faults may happen in voltage sensors, current sensors, speed sensors, temperature sensors, etc.
2. Converter faults: Aging components such as performance degradation of capacitance, short- and open-circuit of insulated-gate bipolar transistors (IGBTs) are common faults appearing in traction converters.
3. Motor faults: Rotor-broken bar, air-gap eccentricity, together with interturn-short circuits will induce faults in traction motors.
4. Control-unit faults: Errors in both analog and digital signals are responsible for faults in traction control units.

According to the appearance time of faults, two categories are summarized as follows:

1. Permanent faults: Some hardware malfunctions such as open circuit of IGBT and gear war belong to the permanent faults.
2. Intermittent/Transient faults: These faults appear, disappear and reappear nondeterministically, and the duration time is short such that important features are difficult to be captured [26].

In addition, faults in electrical drive systems can be categorized into three types based on their amplitudes:

1. Incipient faults: This type of fault is usually characterized by small amplitudes, tiny influences and common faults as time goes on [24]. These faults in electrical drive systems of high-speed trains are, for example, sensor faults and aging components.
2. Common faults: There are some kinds of faults that have larger amplitudes than incipient faults and at the same time affect the performance of trains in a considerable means. Timely maintenance is necessary when they occur.
3. Failures: The failure means malfunctions of components or systems. It usually results in system performance far from the acceptable operation. The broken IGBT will distort three-phase currents, causing degraded traction efficiency.

2.3. Objectives

Based on preliminaries of electrical driven systems and fault types [27], the proposed HIL platform can simulate a variety of failure scenarios in electrical drive systems of high-speed trains. It can effectively avoid the following problems: (1) Design the extra hardware in hardware-based fault injection technology; (2) Real-time debugging difficulty in software-based fault injection technology. It fully provides the experimental verification convenience for peers. It contributes to the objective and comprehensive evaluation of the proposed fault diagnosis algorithm. At the same time, it is possible to transform the research results into practical applications or give technical guidance for engineers.

3. Fault Injection Methodology

3.1. Signal-Based Fault Injection Methods

As mentioned in Section 2, among the four main components of traction systems, there are some typical electronic-subsystems and mechanical equipment, such as IGBT in converters, monitoring or communication modules in TCU and rotors in traction motors. These subsystems maintain the stable operation of TCS, while they may also malfunction. Some faults are attributed to inaccuracy during the development, while others can originate from external causes such as production process defects or environmental stress (including heating, electromagnetic, mechanical stress, etc.). Once fault occurs, the electrical characteristics of faulty components are distinct from the normal status, and these changes are directly reflected in the output signals.

As mentioned in Section 2, the fault scenarios can be described qualitatively regarding the measurements. In order to generate the artificial fault scenarios, the signal-based fault injection method is introduced, providing the observable signals for FD as well as the associated fault prognosis.

For a faulty component or subsystem, its D fault types can be written as:

$$f = \{f_1, \dots, f_i, \dots, f_D\} \quad (1)$$

where $i = 1, \dots, D$. According to the fault f_i of consideration, a set of corresponding signals will be generated. Based on the characteristics such as time and amplitude, the fault signal $f(t)$ can be defined as follows:

- (1) If $f(t)$ is the transient fault, then

$$f(t) = \varepsilon(t - t_s) - \varepsilon(t - t_e) \cdot \lambda\{a_j\}. \quad (2)$$

(2) If $f(t)$ is the intermittent fault, then

$$f(t) = \sum_{j=1}^N \sum_{j=1}^{n_j} [\varepsilon(t - (T_{j+1} + \tau_j \cdot \rho_j) - T_j - \varepsilon(t - T_{j+1}))] \cdot \lambda\{a_j\}. \quad (3)$$

(3) If $f(t)$ is the permanent fault, then

$$f(t) = \varepsilon(t - t_s) \cdot \lambda\{a_j\}. \quad (4)$$

From (2) to (4), ε signifies the step function; t_s is the activation time of faults; $\lambda\{\cdot\}$ represents the threshold function for selecting analog or digital signals; a_j means the amplitude of impulses. For the intermittent-fault cases, N is the total number of impulse-sequence types; j represents the index of the category of impulse sequences; n_j means the number of pulses; τ_j denotes the period of the j -th impulse sequence; ρ is the duty cycle of impulse sequences. In fact, (2)–(4) cover a spread spectrum that can describe transient faults, intermittent faults and permanent faults.

Along with the aforementioned descriptions, the framework of the signal-based FI method is shown in Figure 2, where the FI manager controls the injection process. First, the fault scenarios of the tested subsystem or components are analyzed, and the FI point A is determined. Hence, the original signal S^0 will be obtained. Based on the fault types (which will be injected into the subsystem) confirmed by user instructions, the activation module then provides the activation time of the faults. After that, the fault signal f described by (2)–(4) is determined by the fault library. By this means, the fault injection signal S_f can be obtained by the signal conditioning operator. Specifically, the signal conditioning operator consists of three kinds of operators (i.e., the adder, multiplier and multiplexer). Because subsystems/components are connected by the wires, the FI point can be split into A and A' . As depicted in Figure 2, the FI is placed between the real and target subsystems. Consequently, the input signal of the target subsystems/components is replaced by the FI signal, and thereby the signal-based FI targets can be implemented.

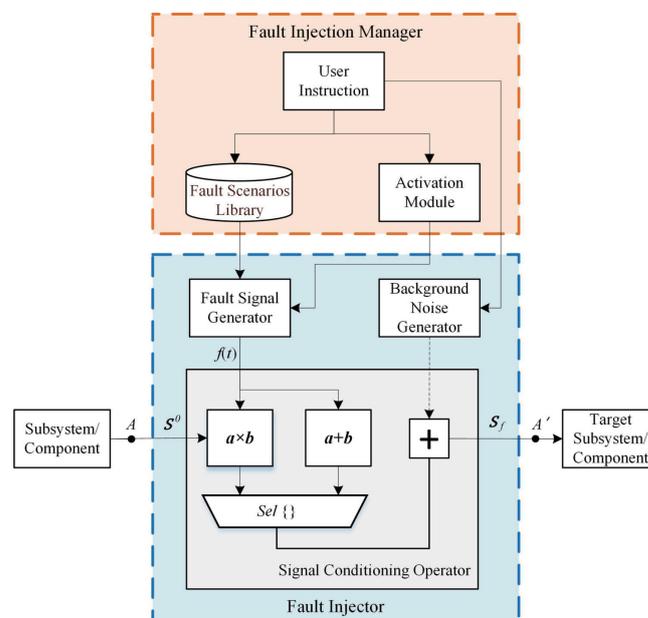


Figure 2. The framework of the signal-based FI method.

3.2. Hardware-in-the-Loop Fault Injection for Traction Systems

For the purpose of realizing the HIL fault injection, this work takes into consideration the following aspects:

- (1) A real-time simulation of models consumes a large amount of FPGA resources, especially for power electronics-based apparatus. However, not all models are required for such a short execution time.
- (2) High FPGA resources will be consumed when the FI signals are inserted.

Based on a multiprocessor structure, the work will present an HIL fault injection platform for electrical drive systems with consideration of the two above mentioned aspects. As shown in Figure 3, the HIL-based FI structure consists of a master-slave system based on multi-FPGAs/CPU, physical TCU, signal modulation module, host PC, fault scenarios library, D/A and digital I/Os, Gigalink and peripheral high-speed bus (PHS). Besides, the variable transmission between FPGA and CPU is implemented by the PHS bus.

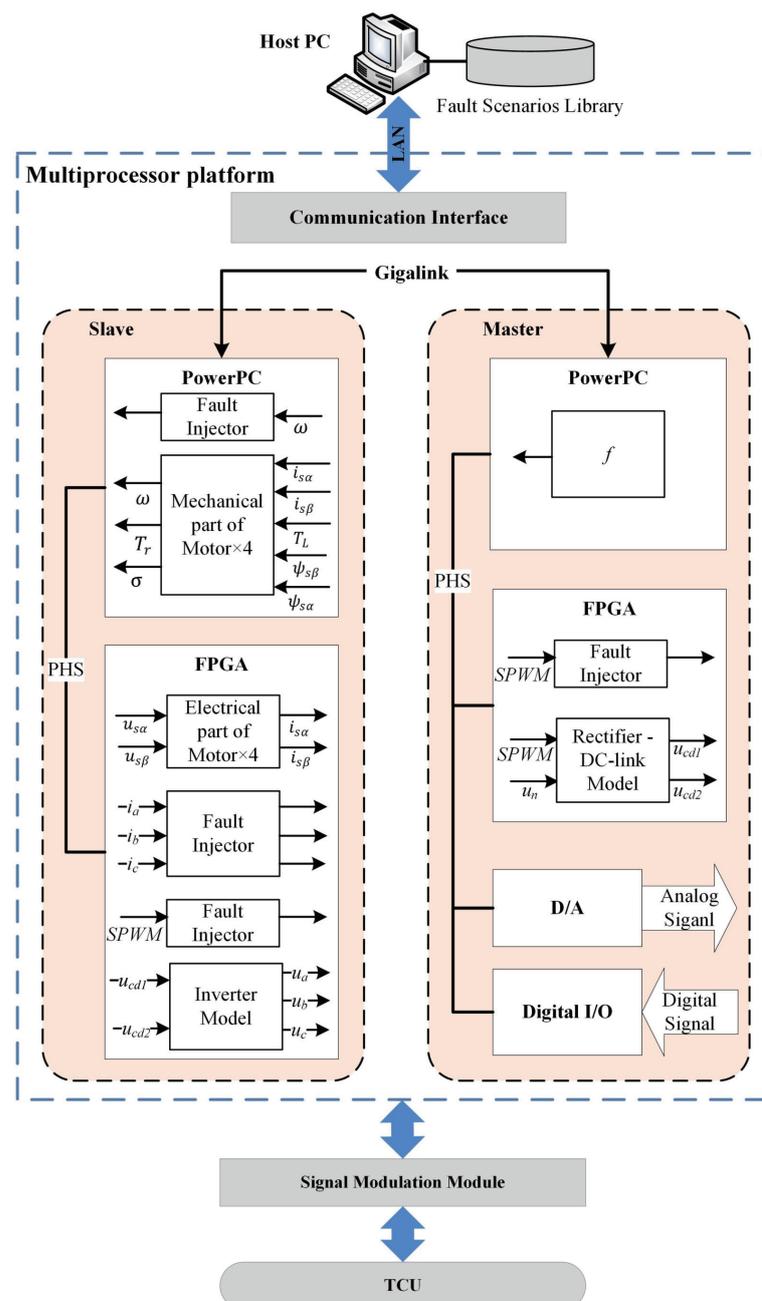


Figure 3. Schematic views of multi-FPGAs/CPU HIL platform for real-time fault injection.

4. Fault Diagnosis Methodology

This section will first develop the fault detection strategies including how to extract the fault features and how to establish the test statistics. Second, by the use of these fault features, several fault diagnosis strategies will be formulated. Based on the results including fault detection and FD, a comprehensive evaluation index will be obtained.

4.1. Fault Detection

Let the number of all measurement variables z be k_z . When a fault $f_i(k)$ appears in electrical drive systems, unexpected deviations will be reflected in z such that

$$z^f(k) = z(k) + \Omega_i f_i(k) \quad (5)$$

where k represents the sampling step, Ω_i is the matrix signifying the fault direction, $\|f_i\|$ describes the deviational magnitude caused by f_i and the subscript “ f ” means different faulty conditions [24].

In (5), z reflects the uncertainty caused by the surrounding noises and external disturbances that are unknown beforehand in practical applications. Whilst, Ωf results in unexpected deviations on z .

For the fault detection purpose, we consider the high-speed trains to work under a steady condition, i.e., at a given traction speed. Furthermore, the moving-window technique is used for enlarging the fault features. Due to the simplicity, the stacked data in a certain size of the moving window is defined as its original form. To detect the fault occurring in electrical drive systems, two necessary steps are:

- (1) To extract fault features that are helpful for addressing high-frequency (online) data.
- (2) To define a test statistic, based on which a reliable detection result of faults can be returned.

4.1.1. Data-Driven Feature Extraction

Some important features are hidden in the measurement signals, as summarized in Table 1. Among them, the mean, covariance and slope are employed in our designed platform to detect and diagnose different faults. Remarkably, the slope is an important index that can be utilized for determining incipient faults.

Table 1. Key signal features and data processing approaches.

Features	Approaches
Mean	Expectation operation
Variance/covariance	Principal component analysis
High-order statistics	Independent component analysis
Correlation	Canonical correlation analysis
Slope	First-order difference
Entropy	Mutual information
Waveform index	Wavelet transform
Root mean square	Partial least squares

For achieving online fault detection tasks where both sampling frequency and dimension of signals are high, principal component analysis is used for improving computation efficiency. Therefore, $z(k)$ on the principal and residual subspace become

$$\begin{aligned} \hat{z}(k) &= P_p P_p^T z(k) \\ \tilde{z}(k) &= P_r P_r^T z(k) \end{aligned} \quad (6)$$

where $\hat{z}(k)$ and $\tilde{z}(k)$ are respectively principal and residual components of $z(k)$; $P_p \in R^{k_z \times k_p}$ and $P_r \in R^{k_z \times (k_z - k_p)}$ are the loading matrices which are obtained via

$$\mathbb{E}((z - \mathbb{E}(z))^2) = [P_p \ P_r] \begin{bmatrix} \Lambda_p & 0 \\ 0 & \Lambda_r \end{bmatrix} [P_p \ P_r]^T. \quad (7)$$

In addition, k_p represents the number of principal components and satisfies $k_p \ll k_z$. The information of measurement variable $z(k)$ is reflected in the variance. It should be noted that $z(k)$ obeys Gaussian distribution. As mentioned in [24,28,29], the operation of traction systems is accompanied by non-Gaussian signals. In order to deal with the measured non-Gaussian signals, independent component analysis (ICA) is adopted in FDD of traction systems. Therefore, the hidden statistically independent components can be extracted by

$$\begin{aligned} \hat{s}(k) &= Wz(k) \\ \text{s.t. } \max \mathcal{M}(\hat{s}(k)) \end{aligned} \quad (8)$$

where $\hat{s}(k)$ is the estimation of independent components; \mathcal{M} is the non-Gaussian measurement function, whose function is detailed in [30].

Different from the PCA and ICA approaches, canonical correlation analysis (CCA) is to find the maximum correlation between system input and output. Then, canonical variables are extracted among measured signals. Suppose that the model has input vector $\mathbf{u}_o \in R^l$ and output vector $\mathbf{y}_o \in R^m$ such that

$$\begin{bmatrix} \mathbf{u}_o \\ \mathbf{y}_o \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} u_u \\ u_y \end{bmatrix}, \begin{bmatrix} \Sigma_u & \Sigma_{uy} \\ \Sigma_{uy}^T & \Sigma_y \end{bmatrix} \right) \quad (9)$$

where Σ is covariance matrices. Then, the correlation evaluation is defined as

$$\mathfrak{E} = \Sigma_u^{-1/2} \Sigma_{uy} \Sigma_y^{-1/2}. \quad (10)$$

Based on an SVD on \mathfrak{E} , the matrix \mathfrak{E} can be decomposed as

$$\mathfrak{E} = \Gamma \Sigma \mathbf{R}^T \quad (11)$$

with $\Gamma = (\gamma_1, \dots, \gamma_l)$, $\Sigma = \begin{bmatrix} \Sigma_\kappa & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{R} = (r_1, \dots, r_m)$. where κ is the number of principal components, Σ_κ is canonical correlation matrix [31]. Then, the canonical correlation vectors are given as

$$\begin{aligned} \mathbf{J} &= \Sigma_u^{-1/2} \Gamma \in R^{ll} \\ \mathbf{L} &= \Sigma_y^{-1/2} \mathbf{R} \in R^{mm} \end{aligned} \quad (12)$$

Obviously, CCA method maintains the following properties [32]

$$\begin{aligned} \Gamma \Gamma^T &= \mathbf{I}_{k_{zu}}, & \mathbf{R}^T \mathbf{R} &= \mathbf{I}_{k_{yy}}, & \mathbf{J}^T \Sigma_{uy} \mathbf{L} &= \Sigma \mathbf{L}^T, \\ \Sigma_{uy}^T &= \Sigma^T \mathbf{J}^T \Sigma_u, & \mathbf{J}^T \Sigma_y &= \Sigma \mathbf{L}^T \Sigma_y. \end{aligned} \quad (13)$$

The three methods mentioned above have shown superior performance in static FDD. However, as the system state (such as the controller adjust the operation), the data will have dynamic characteristics [33]. Therefore, dynamic fault diagnosis methods have been

widely used in high-speed trains [34]. The state space model of the traction system can be defined as [6]:

$$\begin{aligned}x(k+1) &= Ax(k) + Bu(k) + w(k) \\y(k) &= Cx(k) + Du(k) + v(k)\end{aligned}\quad (14)$$

where $x \in R^{m_x}$ is the state of electrical drive systems. In general, $w \in R^{m_x}$ and $v \in R^{m_y}$ are noise sequences and are normally distributed. Matrices A, B, C, D are appropriately dimensioned in electrical drive systems. Specifically, the matrices are the change in angular velocity at different times. For example, the matrix A is given as

$$A = \begin{bmatrix} \frac{-R_r L_m^2 - R_s L_r^2}{\sigma L_s L_r^2} & 0 & \frac{R_r L_m}{\sigma L_s L_r^2} & \frac{L_m \omega_r}{\sigma L_s} \\ 0 & \frac{-R_r L_m^2 - R_s L_r^2}{\sigma L_s L_r^2} & -\frac{L_m \omega_r}{\sigma L_s} & \frac{R_r L_m}{\sigma L_s L_r^2} \\ \frac{L_m R_r}{L_r} & 0 & -\frac{R_r}{L_r} & -\omega_r \\ 0 & \frac{L_m R_r}{L_r} & -\omega_r & -\frac{R_r}{L_r} \end{bmatrix}, \quad (15)$$

where ω_r is the motor rotor speed, and σ is the coefficient of magnetic leakage which can be obtained by $\sigma = (L_s L_r - L_m^2) / (L_s L_r)$. L_s, L_r and L_m are inductance in stator side, inductance in rotor side and mutual inductance of motor, respectively. R_r and R_s are resistance in stator side and in rotor side, respectively. Obviously, matrix A shows that parameters change in real time. Therefore, it is significant to adopt dynamic methods for diagnosing faults in electrical drive systems.

Next, the following three notations are introduced to describe electric drive systems of high-speed trains:

$$\begin{aligned}w_s(k) &= [w^T(k-s) \cdots w^T(k)] \in R^{(s+1)n} \\ \Omega_k &= [w(k) \cdots w(k+N-1)] \in R^{n \times N} \\ \Omega_{k,s} &= [w(k) \cdots w(k+N-1)] = [\Omega_{k-s} \cdots \Omega_k]^T \in R^{(s+1)n \times N}\end{aligned}\quad (16)$$

where s and N are some integers.

Based on the parity space approach, the state space model (14) of electric drive systems is re-written as

$$Y_{k,s} = \Gamma_s X_{k-s} + H_{u,s} U_{k,s} + H_{w,s} W_{k,s} + V_{k,s} \in R^{(s+1)m_y \times N}, \quad (17)$$

$$\begin{aligned}X_{k-s} &= [x(k-s) \cdots x(k-s+N-1)], \Gamma_s = [C^T (CA)^T \cdots (CA^s)^T]^T, \\ H_{u,s} &= \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{s-1}B & \cdots & CB & D \end{bmatrix}, H_{w,s} = \begin{bmatrix} D & 0 & \cdots & 0 \\ C & D & \ddots & \vdots \\ \vdots & \ddots & D & 0 \\ CA^{s-1} & \cdots & C & D \end{bmatrix},\end{aligned}\quad (18)$$

where $\Gamma_s \in R^{(s+1)m_y \times m_x}$, $H_{u,s} \in R^{(s+1)m_y \times (s+1)m_u}$, and $H_{w,s} \in R^{(s+1)m_y \times (s+1)m_u}$. $H_{w,s} W_{k,s} + V_{k,s}$ represents the noise of electric drive systems.

The state space model (17) can be further written into

$$\begin{bmatrix} U_{k,s} \\ Y_{k,s} \end{bmatrix} = \begin{bmatrix} I & 0 \\ H_{u,s} & \Gamma_s \end{bmatrix} \begin{bmatrix} U_{k,s} \\ X_{k-s} \end{bmatrix} + \begin{bmatrix} 0 \\ H_{w,s} W_{k,s} + V_{k,s} \end{bmatrix}. \quad (19)$$

Construct

$$Z_f = \begin{bmatrix} U_{k,s} \\ Y_{k,s} \end{bmatrix}, Z_p = \begin{bmatrix} U_{k-s_p-1, s_p} \\ Y_{k-s_p-1, s_p} \end{bmatrix}. \quad (20)$$

where Z_f and Z_p represent future and past collected data, respectively. Therefore, LQ decomposition is performed on the data set

$$\begin{bmatrix} Z_p \\ U_{k,s} \\ Y_{k,s} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix}, \quad (21)$$

where

$$\begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} [Q_1^T \quad Q_2^T \quad Q_3^T] = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}. \quad (22)$$

Finally, $U_{k,s}$ and $Y_{k,s}$ is expressed as

$$Z_f \begin{bmatrix} Z_p \\ U_{k,s} \end{bmatrix}^T = \begin{bmatrix} L_{21} & L_{22} \\ L_{31} & L_{32} \end{bmatrix} \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}^T, \quad (23)$$

$$Y_{k,s} = [L_{31} \quad L_{32}] \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} + L_{33}Q_3. \quad (24)$$

In a word, the dynamic method maps the input and output data into several stacking matrices to establish the system model. In addition, the high-speed train is vulnerable to external interference and condition switching, such as acceleration and deceleration. Therefore, the internal parameters change in real time, so that the dynamic method is more appropriate in electric drive systems. Compared with static methods, it is often not rigorous enough to deal with such situations.

4.1.2. Definition of Test Statistics

Based on the obtained features of signals, two test statistics, the T^2 and squared prediction error (SPE), are defined at k -th time instance as follows

$$\begin{aligned} T^2(k) &= z^T(k)P_p\Lambda_p^{-1}P_p^Tz(k) \\ \text{SPE}(k) &= z^T(k)(I_{k_z \times k_z} - P_pP_p^T)z(k) \end{aligned} \quad (25)$$

Then a reliable fault detection task can be achieved according to the following binary decision

$$\begin{aligned} \text{Fault and alarm} &\iff T^2(k) \geq J_{th,T^2} \text{ or } \text{SPE} \geq J_{th,\text{SPE}} \\ \text{Fault-free} &\iff \text{otherwise} \end{aligned} \quad (26)$$

where J_{th} signifies the threshold of test statistics. In practical scenarios, the flood of false alarms must be prohibited so that the detection results can provide the conductor an auxiliary and effective indicator. Therefore, determining proper thresholds is usually accomplished via abundant tests [6]. It should be pointed out that the distribution of test statistics, such as the Chi-square distribution, cannot be directly used as a look-up table because $z(k)$ is non-Gaussian.

4.2. Fault Diagnosis

In addition, taking PCA as an example, $\hat{z}(k)$ and $\tilde{z}(k)$ under faulty cases will be

$$\begin{aligned} \hat{z}^f(k) &= \hat{z}(k) + P_pP_p^T\Omega_i f_i(k), \\ \tilde{z}^f(k) &= \tilde{z}(k) + P_rP_r^T\Omega_i f_i(k). \end{aligned} \quad (27)$$

As shown in (27), the necessary condition that f_i can be detectable and diagnosable is

$$PP^T \not\perp \Omega_i \quad (28)$$

where $\not\perp$ denotes the non-orthogonality between two spaces. After the fault being detected successfully, then the objective of fault diagnosis is

$$\max \Pr(f \in f_i | PP^T \Omega_i f), \quad (29)$$

or

$$\max \Pr(f \in f_i | \Omega_i f) \quad (30)$$

where $\Pr(\cdot)$ is the probability.

In fact, the implementation of (29) is depended upon the fault feature; whilst the formulation of (30) is directly based on fault information hidden in original signals.

4.3. Comprehensive Evaluation Indices

At present, comprehensive evaluation indexes of FD can be considered from two dimensions in electrical drive systems: (1) the actual engineering indexes are analyzed to understand the engineering requirements; (2) the differences of performance indexes among the methods are verified in the laboratory. The following two dimensions are introduced in detail.

In practical engineering, FD is also called quality evaluation in electrical drive systems. The main method of quality evaluation is to determine the fault level by the score of the evaluation. Specifically, it is divided into the following steps:

Step 1: Standards of fault deduction. The identified faults include four types: A, B, C and D. Among them, 5 points will be deducted for class A, 10 points for class B, 20 points for class C and 100 points for class D. Partial score standards of faults of electrical drive systems are listed in Table 2.

Table 2. Evaluation standards of electrical drive systems.

Standards of Quality	Faults Types			
	A	B	C	D
Main transformer and its cooling	Poor cleaning	Loose wiring, and invalid desiccant	Abnormal fan rotation and low liquid level or oil spill	Traction converter faults
Traction converter and its control	Poor cleaning	Loose wiring, and invalid desiccant	Abnormal fan rotation and traction converter faults	None
Traction motor and its cooling	Poor installation of oil filling plug and unstable cooling duct	Loose fittings, broken wiring, missing oil injection plug and blocked exhaust	Loose and damaged power line, sensor and wiring	Cracked mount and bad motor
Auxiliary converter	None	None	One dysfunction	Two dysfunctions

Step 2: According to Table 2, the quality score of electrical drive systems can be given

$$\text{True score} = 1000 - \text{Fault score} \times 8. \quad (31)$$

where the true score must not be negative, and the full score for quality is 1000 points.

Step 3: Based on the true score, the evaluation list of electrical drive systems is presented in Table 3.

Table 3. Fault level results of electrical drive systems.

Fault Levels	True Scores
A	900–1000
B	800–899
C	700–799
D	699 or less

Step 4: Finally, engineers repair the damaged parts according to the true score.

The above steps can effectively provide more direct maintenance for electrical drive systems in the actual project, which is helpful to help engineers identify different fault levels. In addition, in order to further measure the superior performance of different algorithms, the following comprehensive evaluation indicators of false alarm rate (FAR) and missing alarm rate (MAR) are given in the laboratory.

FAR is defined as the probability that normal operations are judged to be faulty, which can be expressed as

$$\text{FAR} = P(J > J_{th} | f = 0), \quad (32)$$

where J is the test statistic or evaluation function to measure the system. Generally, there are different definitions of J about the requirements of the system [35]. For example, J has ℓ_2 -norm, root square mean, ℓ_∞ -norm and other forms to provide a more accurate measurement standard. J_{th} is represented as a threshold. The setting of J_{th} is directly related to J . In fault detection, FAR has important physical meaning to some extent.

MAR is defined as the probability that faults are not successfully detected, which can be expressed as

$$\text{MAR} = P(J < J_{th} | f \neq 0). \quad (33)$$

Compared with FAR, MAR is concerned with the fault detection ability of the system in case of fault.

Hence, it is helpful to the scientific research and engineering application of HIL to select the appropriate evaluation indices based on the actual engineering needs.

4.4. An Overview of FD Methods

The HIL-based platform has been widely used by a large number of experts and scholars to verify the effectiveness of the proposed FD methods [36–38]. From the perspective of model-based diagnosis method, a real-time FD method for sensors and IGBTs of the impulse rectifier is proposed in [39]. It is based on the structural analysis of electrical drive systems whose structural model has been established. The structural model under various fault conditions is evaluated and optimized. In [11], an FI strategy for safety testing and FD verification is presented in the electrical drive control systems. By simulating fault scenarios, the influence of fault signals on the system is analyzed from the perspective of mechanism. The FI strategy based on signal adjustment is adopted to pave the way for subsequent data-driven FD experiments.

As a common abnormality of the electrical drive systems, the incipient fault has always been an international problem that needs to be solved urgently. Since it is hard to identify the existence of incipient faults from the perspective of model-based diagnosis method, data-driven fault diagnosis method has been widely concerned, and a large number of research results are published [40,41]. Aiming at the crowding problem caused by incipient fault, a fault detection and diagnosis method is proposed based on probability-relevant PCA [25]. The effectiveness of the proposed method is verified in incipient FD by using HIL platform. In [42], considering the heterogeneity of sensor distribution, a multi-blocks system monitoring scheme is proposed. In the experimental verification, HIL is used to compare the performance on the different comprehensive evaluation indexes. For multimode fault detection, a just-in-time-learning aided CCA is proposed in [31] by using the system modal change of HIL platform, which overcomes the traditional method in the single working condition process.

In summary, the HIL platform provides an integration of environment of FD and FD, which improves the possibility for the practical application of advanced theoretical methods.

5. Conclusions

In this paper, an HIL-based platform has been developed for simulating the realistic faults in electrical drive systems of high-speed trains. Then FD algorithms have been tested to detect the obtained faults, whose performance can be evaluated via comprehensive indices. This platform is an integrated design that covers both FI and FD techniques, providing engineers or researchers with a reliable simulation system. The design platform is developed based on an electrical drive system of CRH2-type trains. The abundant attempts of FD have illustrated its effectiveness and feasibility. In the coming ten years, the authors believe that, based on this tutorial, there will be lots of follow-up publications and applications. It can not only analyze the electrical fault of the electric drive system but also identify the mechanical fault.

Author Contributions: Formal analysis, X.Q.; Funding acquisition, C.C.; Methodology, X.Y.; Project administration, H.C.; Software, K.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61903047 and U20A20186, in part by the Department of Science and Technology of Jilin Province under Grant 20200401127GX.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Givoni, M. Development and impact of the modern high-speed train: A review. *Transp. Rev.* **2006**, *26*, 593–611. [[CrossRef](#)]
2. Zhang, S. *Fundamental Application Theory and Engineering Technology for Railway High-Speed Trains*; Science Press: Beijing, China, 2007.
3. Romanenko, A.; Muetze, A.; Ahola, J. Incipient bearing damage monitoring of 940-h variable speed drive system operation. *IEEE Trans. Energy Convers.* **2017**, *32*, 99–110. [[CrossRef](#)]
4. Chen, Z.; Haynes, K.E. *Chinese Railways in the Era of High-Speed*; Emerald Group Publishing Limited: Bingley, West Yorkshire, UK, 2015.
5. Chen, H.; Jiang, B. A review of fault detection and diagnosis for the traction system in high-speed trains. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 450–465. [[CrossRef](#)]
6. Chen, H.; Jiang, B.; Ding, S.X.; Huang, B. Data-driven fault diagnosis for traction systems in high-speed trains: A survey, challenges, and perspectives. *IEEE Trans. Intell. Transp. Syst.* **2020**. [[CrossRef](#)]
7. Garramiola, F.; Poza, J.; Madina, P.; del Olmo, J.; Ugalde, G. A hybrid sensor fault diagnosis for maintenance in railway traction drives. *Sensors* **2020**, *20*, 962. [[CrossRef](#)]
8. Xu, Z.; Wang, W.; Sun, Y. Performance degradation monitoring for onboard speed sensors of trains. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 1287–1297. [[CrossRef](#)]
9. Feng, D.; Lin, S.; Yang, Q.; Lin, X.; He, Z.; Li, W. Reliability evaluation for traction power supply system of high-speed railway considering relay protection. *IEEE Trans. Transport. Electrification.* **2019**, *5*, 285–298. [[CrossRef](#)]
10. Wu, Y.; Jiang, B.; Wang, Y. Incipient winding fault detection and diagnosis for squirrel-cage induction motors equipped on CRH trains. *ISA Trans.* **2020**, *99*, 488–495. [[CrossRef](#)]
11. Yang, C.; Yang, C.; Peng, T.; Yang, X.; Gui, W. A fault-injection strategy for traction drive control systems. *IEEE Trans. Ind. Electron.* **2017**, *64*, 5719–5727. [[CrossRef](#)]
12. Yang, X.; Yang, C.; Peng, T.; Chen, Z.; Liu, B.; Gui, W. Hardware-in-the-loop fault injection for traction control system. *IEEE J. Emerg. Sel. Top. Power Electron.* **2018**, *6*, 696–706. [[CrossRef](#)]
13. Arlat, J.; Aguera, M.; Amat, L.; Crouzet, Y.; Fabre, J.C.; Laprie, J.C.; Martins, E.; Powell, D. Fault injection for dependability validation: A methodology and some applications. *IEEE Trans. Softw. Eng.* **1990**, *16*, 166–182. [[CrossRef](#)]
14. Benso, A.; Prinetto, P. *Fault Injection Techniques and Tools for Embedded Systems Reliability Evaluation*; Springer: New York, NY, USA, 2003.

15. Andrés, D.D.; Ruiz, J.C.; Gil, D.; Gil, P. Fault emulation for dependability evaluation of VLSI systems. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2008**, *16*, 422–431. [[CrossRef](#)]
16. Lu, A.; Yang, G. False data injection attacks against state estimation in the presence of sensor failures. *Inform. Sci.* **2020**, *508*, 92–104. [[CrossRef](#)]
17. Hasanzadeh, A.; Edrington, C.S.; Stroupe, N.; Bevis, T. Real-time emulation of a high-speed microturbine permanent-magnet synchronous generator using multiplatform hardware-in-the-loop realization. *IEEE Trans. Ind. Electron.* **2014**, *61*, 3109–3118. [[CrossRef](#)]
18. Gou, B.; Ge, X.; Wang, S.; Feng, X.; Kuo, J.B.; Habetler, T.G. An open-switch fault diagnosis method for single-phase PWM rectifier using a model-based approach in high-speed railway electrical traction drive system. *IEEE Trans. Power Electron.* **2016**, *31*, 3816–3826. [[CrossRef](#)]
19. Wu, Y.; Jiang, B.; Lu, N. A Descriptor System Approach for Estimation of Incipient Faults With Application to High-Speed Railway Traction Devices. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *49*, 2108–2118. [[CrossRef](#)]
20. Estima, J.O.; Cardoso, A.J.M. A new algorithm for real-time multiple open-circuit fault diagnosis in voltage-fed PWM motor drives by the reference current errors. *IEEE Trans. Ind. Electron.* **2012**, *60*, 3496–3505. [[CrossRef](#)]
21. Henao, H.; Kia, S.H.; Capolino, G.A. Torsional-vibration assessment and gear-fault diagnosis in railway traction system. *IEEE Trans. Ind. Electron.* **2011**, *58*, 1707–1717. [[CrossRef](#)]
22. Cabal-Yepez, E.; Garcia-Ramirez, A.G.; Romero-Troncoso, R.J.; Garcia-Perez, A.; Osornio-Rios, R.A. Reconfigurable monitoring system for time-frequency analysis on industrial equipment through STFT and SWT. *IEEE Trans. Ind. Informat.* **2013**, *9*, 760–771. [[CrossRef](#)]
23. Bennett, S.M.; Patton, R.J.; Daley, S. Sensor fault-tolerant control of a rail traction drive. *Control Eng. Pract.* **1999**, *7*, 217–225. [[CrossRef](#)]
24. Chen, H.; Jiang, B.; Lu, N.; Chen, W. *Data-Driven Detection and Diagnosis of Faults in Traction Systems of High-Speed Trains*; Springer Nature: Cham, Switzerland, 2020.
25. Chen, H.; Jiang, B.; Ding, S.X.; Lu, N.; Chen, W. Probability-relevant incipient fault detection and diagnosis methodology with applications to electric drive systems. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 2766–2773. [[CrossRef](#)]
26. Zhou, D.; Zhao, Y.; Wang, Z.; He, X.; Gao, M. Review on diagnosis techniques for intermittent faults in dynamic systems. *IEEE Trans. Ind. Electron.* **2020**, *67*, 2337–2347. [[CrossRef](#)]
27. Chen, H.; Jiang, B.; Ding, S.X. A broad learning aided data-driven framework of fast fault diagnosis for high-speed trains. *IEEE Intell. Transp. Syst. Mag.* **2019**, *13*, 83–88. [[CrossRef](#)]
28. Chen, Z.; Ding, S.X.; Peng, T.; Yang, C.; Gui, W. Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms. *IEEE Trans. Ind. Electron.* **2018**, *65*, 1559–1567. [[CrossRef](#)]
29. Chen, H.; Jiang, B.; Zhang, T.; Lu, N. Data-driven and deep learning-based detection and diagnosis of incipient faults with application to electrical traction systems. *Neurocomputing* **2020**, *369*, 429–437. [[CrossRef](#)]
30. Hyvärinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
31. Chen, Z.; Liu, C.; Ding, S.X.; Peng, T.; Yang, C.; Gui, W.; Shardt, Y.A.W. A just-in-time-learning aided canonical correlation analysis method for multimode process monitoring and fault detection. *IEEE Trans. Ind. Electron.* **2020**, *68*, 5259–5270. [[CrossRef](#)]
32. Anderson, T. *An Introduction to Multivariate Statistical Analysis*, 2nd ed.; Wiley: New York, NY, USA, 1984.
33. Luo, H.; Yang, X.; Krueger, M.; Ding, S.X.; Peng, K. A plug-and-play monitoring and control architecture for disturbance compensation in rolling mills. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 200–210. [[CrossRef](#)]
34. Chen, H.; Jiang, B.; Chen, W.; Li, Z. Edge computing aided framework of fault detection for traction control systems in high-speed trains. *IEEE Trans. Veh. Technol.* **2020**, *69*, 1309–1318. [[CrossRef](#)]
35. Ding, S.X. *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*; Springer: London, UK, 2014.
36. Yang, C.; Peng, T.; Tao, H.; Yang, C.; Gui, W. Review of recent research on fault injection for high-speed train information control systems. *Sci. Inf. Sci.* **2020**, *50*, 465–482.
37. Guo, T.; Sang, J.; Chen, M.; Zhang, J.; Tai, X.; Zhou, D. A fault detection method for a braking system of high-speed trains. *Sci. Inf. Sci.* **2020**, *50*, 483–495. [[CrossRef](#)]
38. Liu, Q.; Zhan, Z.; Wang, S.; Liu, Y.; Fang, T. Data-driven multimodal operation monitoring and fault diagnosis of high-speed train bearings. *Sci. Inf. Sci.* **2020**, *50*, 527–539.
39. Li, X.; Xu, J.; Chen, Z.; Xu, S.; Liu, K. Real-Time Fault Diagnosis of Pulse Rectifier in Traction System Based on Structural Model. *IEEE Trans. Intell. Transp. Syst.* **2020**. [[CrossRef](#)]
40. Chen, H.; Chai, Z.; Jiang, B.; Huang, B. Data-Driven Fault Detection for Dynamic Systems with Performance Degradation: A Unified Transfer Learning Framework. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12. [[CrossRef](#)]
41. Cheng, C.; Qiao, X.; Luo, H.; Wang, G.; Teng, W.; Zhang, B. Data-Driven Incipient Fault Detection and Diagnosis for the Running Gear in High-Speed Trains. *IEEE Trans. Veh. Technol.* **2020**, *69*, 9566–9576. [[CrossRef](#)]
42. Cheng, C.; Qiao, X.; Zhang, B.; Luo, H.; Zhou, Y.; Chen, H. Multiblock Dynamic Slow Feature Analysis-Based System Monitoring for Electrical Drives of High-Speed Trains. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–10.