Data-Driven Fault Diagnosis for Traction Systems in High-Speed Trains: A Survey, Challenges, and Perspectives

Hongtian Chen, *Member, IEEE*, Bin Jiang, *Fellow, IEEE*, Steven X. Ding, and Biao Huang, *Fellow, IEEE*

***Abstract*—Recently, to ensure the reliability and safety of high-speed trains, detection and diagnosis of faults (FDD) in traction systems have become an active issue in the transportation area over the past two decades. Among these FDD methods, data-driven designs, that can be directly implemented without a logical or mathematical description of traction systems, have received special attention because of their overwhelming advantages. Based on the existing data-driven FDD methods for traction systems in high-speed trains, the first objective of this paper is to systematically review and categorize most of the mainstream methods. By analyzing the characteristic of observations from sensors equipped in traction systems, great challenges which may prevent successful FDD implementations on practical high-speed trains are then summarized in detail. Benefiting from theoretical developments of data-driven FDD strategies, instructive perspectives on this topic are further elaborately conceived by the integration of model-based FDD issues, system identification techniques, and new machine learning tools, which provide several promising solutions to FDD strategies for traction systems in high-speed trains.**

***Index Terms*—Data-driven, fault detection and diagnosis (FDD), traction systems, high-speed trains.**

I. Introduction

AMONG today's intelligent transportation means, the high-speed train has been one of the most desirable tools thanks to its variously salient advantages such as high speed and low energy consumption [1]–[10]. Although the advanced high-speed trains have begun to benefit society and economy all over the world, the associated negative effects are unavoidably induced because of the lack of effective deployment, rational institutional supervision, reliable engineering design, etc. [7], [11]–[14]. This leads to that, system reliability and operational safety of high-speed trains will be affected [15]–[21], and thus faults/failures even catastrophes may appear unexpectedly [22]–[31]. Therefore, the fault detection and diagnosis (FDD) topic has become particularly important for high-speed trains.

Regarded as the "heart" of high-speed trains, traction systems provide the tractive effort for supporting continuous operations of trains, and their safety and reliability are therefore crucial [32]. However, as the attended time of high-speed trains increases, irreversible scenarios such as performance degradation and component aging will arise in traction systems [15]. These intractable cases must give rise to unpermitted deviations of characteristic properties or parameters of traction systems from their normal conditions, and can be uniformly defined as "faults" or "failures" [11]. As summarized in [33], these faults can be classified according to component locations, from traction transformers to wheel gears, where faults occur.

A. Current Strategies

Generally, approaches to detect and diagnose faults in traction systems of high-speed trains mainly depend on [34]: i) manual check/inspection in off-line means; and ii) preset alarm mechanism in online means. In 2017, National Railway Administration of the People's Republic of China published "Rules for Operation and Maintenance of Electric Multiple Units" in which, according to the operational time and distance, manual check with the help of large-scale equipments generally covers five grades and its implementation must be done at the fixed maintenance stations. Benefitting from merits such as the simplicity and directness, the univariate control charts gain popularity for the preset alarm mechanism adopted in traction protection devices, where to judge if there is a fault or not can be easily achieved in real time [1].

With the significantly growing degree of automation and intelligence, high-speed trains have become more complicated than their original appearances such as the "Shinkansen" developed by Japan [35]. This truth poses that successful detection, diagnosis, and location of faults may not be guaranteed if only the two kinds of techniques aforementioned continue being adopted. Concretely, the majority of activate FDD methods for traction systems in high-speed trains can be

Manuscript received April 14, 2020; revised July 30, 2020; accepted September 30, 2020. Date of publication October 22, 2020; date of current version March 9, 2022. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant NC2020002 and Grant NP2020103, and in part by the National Natural Science Foundation of China under Grant 62020106003. The Associate Editor for this article was Y. Lv. *(Corresponding author: Hongtian Chen.)*

Hongtian Chen and Biao Huang are with the Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada (e-mail: chtbaylor@163.com; biao.huang@ualberta.ca).

Bin Jiang is with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China, and also with the Jiangsu Key Laboratory of Internet of Things and Control Technologies, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: binjiang@nuaa.edu.cn).

Steven X. Ding is with the Institute for Automatic Control and Complex Systems (AKS), University of Duisburg-Essen, 47057 Duisburg, Germany (e-mail: steven.ding@uni-due.de).

Digital Object Identifier 10.1109/TITS.2020.3029946

generally classified into three categories [15]: signal analysis-based, model-based, and data-driven methods. There is no doubt that data-driven FDD methods will show superior advantages when an accurate mathematical model or expert knowledge about traction systems is absent.

In order to control and monitor traction systems in real time, there are large amounts of sensors equipped in high-speed trains; for example, more than 1000 sensors are mounted on a CRH3-type vehicle [32]. Thanks to the low design efforts and simple forms of data-driven FDD methods, they have been widely investigated owing to their predominant ability to tackle these vastly high-frequency measurements from sensors [11]. To be more specific, data-driven FDD methods can avoid establishing any mathematical modeling of high-speed trains based on first principles, but proceed to directly extract the latent information from historical data sets [15].

B. Objective and Structure

Greatly motivated by these observations, it becomes essential to comprehensively review these advanced and updated FDD methods to improve their applicability, capacity, as well as efficiency. The main contributions of this work can be summarized as follows.

- The first emphasis of this paper is on the system modeling based on both the first principles and signal form. It is fundamental to understand the basics of data-driven methods, to summarize challenges, and to detail perspectives in this FDD topic.
- The second aim is to carry out a systematic review of the emerging research work, the so-called data-driven FDD methods, for traction systems in the past ten years, providing researchers and practitioners with a thorough grasp from both theoretical and application aspects.
- The third attention is, by sufficiently mining inherent natures of signals measured from traction systems, to enumerate important challenges being faced in practice.
- The final attempt is, strongly motivated by practical applications and studied from the up-to-date data analysis methods, to delineate research opportunities on data-driven FDD methods for traction systems.

It should be mentioned that, although time-frequency analysis-based FDD methods also depend on observations from traction systems, the main concerns of this paper will be the focuses on ***multivariate statistical analysis***-based, other ***machine learning***-based, and ***system identification***-based FDD methods for traction systems in high-speed trains.

The rest of this paper is organized as follows. Section II begins with the prerequisite of traction systems, followed by the description of traction systems in multiple forms. Section III details data-driven FDD methods, where test statistics and basic techniques are systematically reviewed. Based on the inherent natures of signals measured from high-speed trains, Section IV is dedicated to discovering the current challenging topics of data-driven FDD techniques in their designs and implementation phases. Along with these difficulties in practical application, Section V elaborately makes an exhaustive trend prediction on the potential FDD methods for traction systems. The last Section VI summarizes this study.

II. Modeling of Traction Systems

Naturally, modeling or description of traction systems of high-speed trains is the first step in FDD procedures and is hence of great importance. In this section, traction systems of high-speed trains are firstly introduced. Then, four general mathematical models of traction systems, as well as their forms with different faults, are formulated.

A. Traction Systems of High-Speed Trains

Although there is a variety of series of high-speed trains developed by countries such as Japan, France, Germany, China, and so forth, these latest traction systems operate in a similar manner and mechanism [1]. To be precise, traction systems of China Railway High-Speed (CRH) series can be regarded as the seamless integration of technologies from Japan, Germany, and France. Taking the CRH2-type high-speed train as an example, traction systems can transform the electric energy from 25 kV AC network into mechanical energy and thereby drive the whole train [4]. In general, a whole high-speed train has multiple traction systems (i.e., electric multiple units) supporting high-speed operations [7]. Moreover, it is able to operate at more than 150 km/h [15]. For example, CRH2A and CRH2C-type high-speed trains are respectively with the train set configurations "4M/4T" and "6M/2T" [7], where "M" means the motor coach and "T" is the trailer coach.

As shown in Fig. 1, one traction system of the high-speed train consists of a traction transformer, a traction rectifier, a DC-link, a traction inverter, four traction motors, and a traction control unit (TCU) [36] where "PI" is the abbreviation of proportional-integral. Also, data acquisition equipment is a key part, recording the operation condition via sensor measurements [1]. In addition, four types of sensors (i.e., voltage, current, speed, and temperature sensors [4], [11], [37]) are equipped in the traction system where voltage, current, and speed sensors are commonly used for both the control and FDD purposes [17].

Thanks to the widespread deployment of sensors together with their information exchange via data bus [4], the data-driven FDD methods are gaining popularity. To start reviewing these FDD methods from the technological viewpoint, a solidly modeling of traction systems is necessary, as introduced in the following subsections.

B. Description Based on First Principles

Schematically, the structure from power supply to traction motors, together with the operation mechanism, of CRH2 is shown in Fig. 1, where pale green parts represent different components [38].

Based on the first principles, the continuous time-invariant state-space model (Model I) of the traction system (from traction inverters, traction motors, to TCUs) in Fig. 1 can be

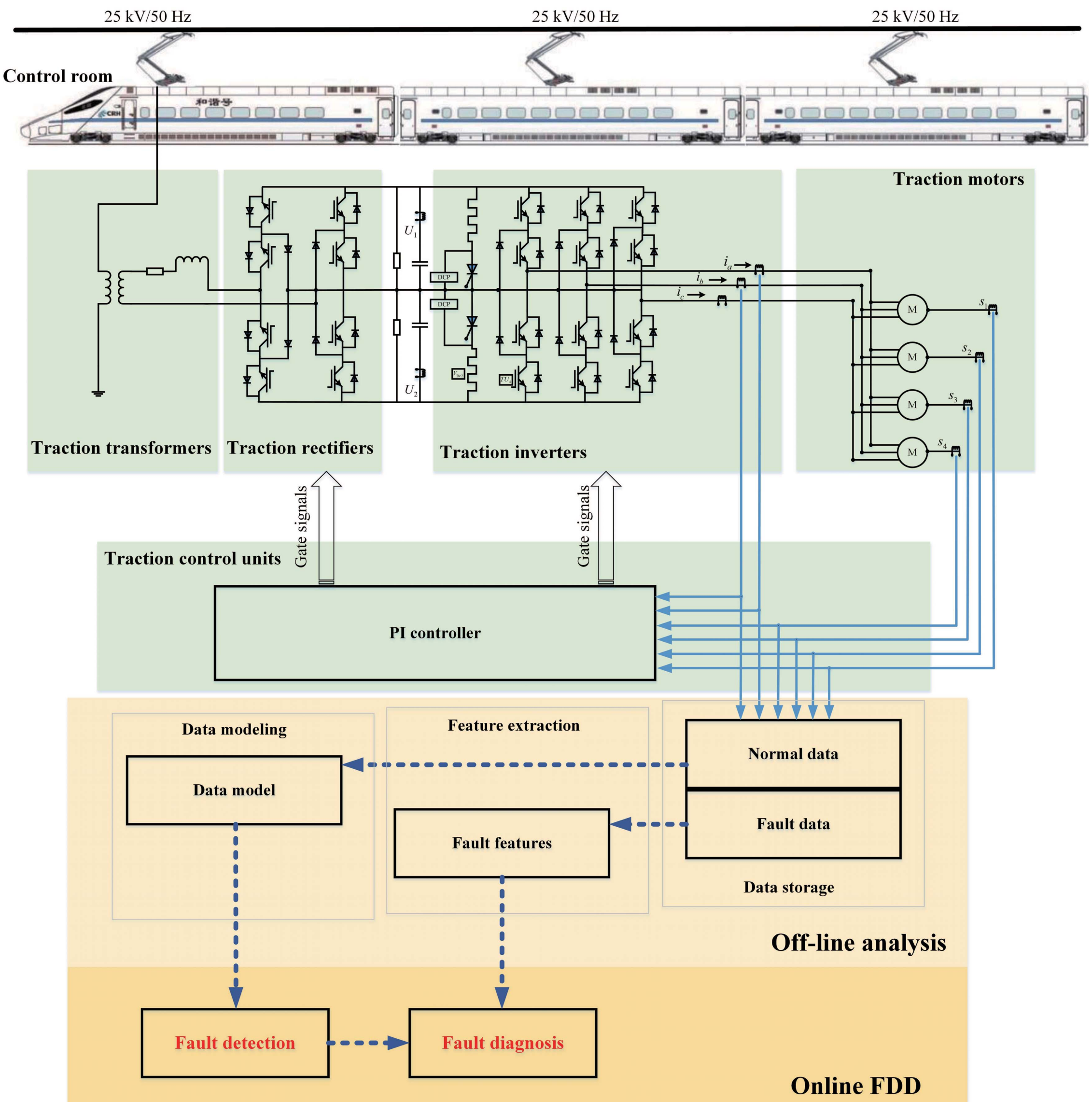

Fig. 1. The structure of traction systems in high-speed trains.

formulated as follows [39]

$$\begin{aligned} \dot{x} &= Gx + Hu \\ y &= Cx \end{aligned} \tag{1}$$

where

$$G = \begin{bmatrix} \frac{-R_r L_m^2 - R_s L_r^2}{\sigma L_s L_r^2} & 0 & \frac{R_r L_m}{\sigma L_s L_r^2} & \frac{L_m \omega_r}{\sigma L_s} \\ 0 & \frac{-R_r L_m^2 - R_s L_r^2}{\sigma L_s L_r^2} & \frac{-L_m \omega_r}{\sigma L_s} & \frac{R_r L_m}{\sigma L_s L_r^2} \\ \frac{L_m R_r}{L_r} & 0 & -\frac{R_r}{L_r} & -\omega_r \\ 0 & \frac{L_m R_r}{L_r} & \omega_r & -\frac{R_r}{L_r} \end{bmatrix},$$

$$H = \begin{bmatrix} \frac{1}{\sigma L_s} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma L_s} & 0 & 0 \end{bmatrix}^T, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

and

$$x = [i_{s\alpha}\ i_{s\beta}\ \psi_{r\alpha}\ \psi_{r\beta}]^T,$$
$$u = [u_{s\alpha}\ u_{s\beta}]^T, \quad y = [i_{s\alpha}\ i_{s\beta}]^T.$$

Parameters $i_{s\alpha}$ and $i_{s\beta}$ respectively represent stator currents; $\psi_{r\alpha}$ and $\psi_{r\beta}$ are the respective rotor fluxes; $u_{s\alpha}$ and $u_{s\beta}$ denote stator voltages, respectively; and all of six parameters above are less of physical meaning and formed in the $\alpha - \beta$ coordinate. In addition, ω_r is the motor rotor speed, and σ is the coefficient of magnetic leakage which can be obtained by $\sigma = (L_s L_r - L_m^2)/(L_s L_r)$. Other parameters, together with their values in CRH2, are given in Table I.

Based on the control theory such as Laplace transform [40], the discrete time-variant Model II, which is equivalent to Model I in (1), can be described by [41]

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \tag{2}$$

where $x \in R^n$, $u \in R^l$, and $y \in R^m$.

Both Model I in (1) and Model II in (2) are popular in model-based FDD methods for traction systems of high-speed trains [11], [39], [42], [43]. The main purposes of introductions of two above models include: 1) to make a distinct comparison between the existing model-based and data-driven FDD methods for high-speed trains; 2) to speculate future FDD investigations for highly dynamic high-speed trains; 3) to establish a bridge between model-based and data-driven FDD methods.

Consider noises and faults in traction systems of high-speed trains, the nominal Model II can be further modeled by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + E_f f(k) + w(k) \\ y(k) &= Cx(k) + F_f f(k) + v(k) \end{aligned} \tag{3}$$

where $w \in R^n$ and $v \in R^m$ are noise sequences and are normally distributed [44] if no additional condition is given. It becomes evident that the fault matrices, E_f and F_f with compatible dimensions, indicate the place where any fault occurs in traction systems, and f stands for the fault amplitude.

C. Description in Signal Forms

Comparing with Model I given in (1) and Model II described by (2) or (3), a more precise and straightforward mathematical model without improper assumptions is preferred. A direct description of traction systems, defined as Model III, in practice is to serve as data-driven FDD methods, which can be schematically described in signal forms by [17]

$$z(k) = \left[u^T(k) \quad y^T(k)\right]^T \in R^{l+m} \tag{4}$$

where z involves all measured signals in traction systems.

$$z_f(k) = z(k) + f(k) \in R^{l+m} \tag{5}$$

It is obvious that, the formulation of Model III presented above consisting of multiple measurement signals is done without any assumptions and also avoids impracticable and sophisticated modeling of traction systems in high-speed trains. In fact, various characteristics or features can be furthermore exploited based on a large amount of measured data from high-speed trains. For example, depending on Model III, to describe a traction system of high-speed trains is achieved in [45] via the probability distribution of measured signals.

D. Description by Input and Output Data

In order to enlighten the bridge and difference between the state space and signal-based models, we review a specific description with stacked input and output data in this subsection.

In order to consider the dynamic behavior of traction systems hidden in input and output data which is particularly important, several stacked matrices have been generally used to build the link between Model II in (2) and Model III in (4) [46]. Denote $\omega(k)$ as a vector to stack vectors and matrices in such form [47]:

$$\begin{aligned} \omega_s(k) &= \left[\omega^T(k-s) \ \dots \ \omega^T(k)\right]^T \in R^{(s+1)\varpi} \\ \Omega_k &= [\omega(k) \ \dots \ \omega(k+N-1)] \in R^{\varpi \times N} \\ \Omega_{k,s} &= \left[\Omega_{k-s}^T \ \dots \ \Omega_k^T\right]^T \in R^{(s+1)\varpi \times N} \end{aligned} \tag{6}$$

where $\omega(k) \in R^{\varpi}$, and s, N mean some large integers. According to definitions in (6), the Model IV can be therefore depicted by [48]

$$Y_{k,s} = \Gamma_s X_{k-s} + H_{u,s} U_{k,s} + H_{w,s} W_{k,s} + V_{k,s} \tag{7}$$

which combines Model II and Model III, where $Y_{k,s} \in R^{(s+1)m \times N}$, $H_{u,s} \in R^{(s+1)m \times (s+1)l}$, $H_{w,s} \in R^{(s+1)m \times (s+1)n}$, and

$$\Gamma_s = \left[C^T \ \dots \ (CA^s)^T\right]^T \in R^{(s+1)m \times n}, \tag{8}$$

$$H_{u,s} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ CB & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ CA^{s-1}B & \dots & CB & 0 \end{bmatrix}, \tag{9}$$

$$H_{w,s} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ C & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ CA^{s-1} & \dots & C & 0 \end{bmatrix}. \tag{10}$$

What makes Model IV attractive is its ability to consider the dynamic performance of traction systems when only the input and output data of traction systems are available.

Except for the aforementioned models, data-driven strategies are also used for establishing train control models. For example, [49] proposes three modeling algorithms with the aid of large-scalely historical data, where one linear model is similar to Model II in (2) and two nonlinear models are developed according to measurements of Model III in (4). By the integration of expert knowledge from multiple experienced train drivers, data-driven models are recently developed in [50] for high-speed trains.

III. Basic Data-Driven FDD Approaches for Traction Systems of High-Speed Trains

In this section, basic data-driven FDD methods and their variants for traction systems of high-speed trains are reviewed, also with a premier emphasis on the hypothesis test used for fault detection. Up to now, most of these FDD schemes belong to static methods, whose schematic description is sketched in Fig. 2. Their basics are theoretically and practically instrumental for FDD applications to the traction systems.

A. Hypothesis Test and Test Statistics

Considering a measurement z from traction systems as given in Model III, it can be projected into a new space via linear or nonlinear projections:

$$g : z \longmapsto \omega \in R^{\varpi} \tag{11}$$

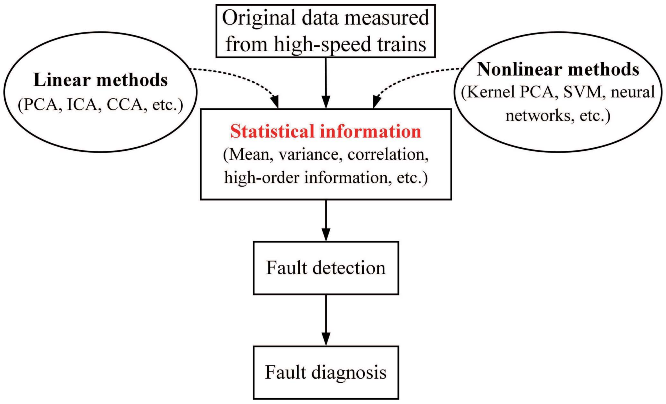

Fig. 2. A schematic description of data-driven static FDD methods for traction systems in high-speed trains.

where g is the operating rules that can describe all techniques adopted in data-driven FDD methods. The measurement z can be labeled as z_n and z_f based on the real work conditions of traction systems in high-speed trains, where n and f are the abbreviations of "normal" and "faulty". According to (11), we have $g : z_n \longmapsto \omega_n$ and $g : z_f \longmapsto \omega_f$. In addition, we can define two regions, $\mathcal{S}_n$ and $\mathcal{S}_f$, which are namely the control and rejection regions, where $\mathcal{S}_n \bigcap \mathcal{S}_f = 0$ and $\mathcal{S}_n \bigcup \mathcal{S}_f = R^{\varpi}$.

Given z measured from traction systems under an unknown condition, the hypothesis test to make a reliable fault decision can be uniformly formulated as follows:

$$\begin{cases} H_0 : \omega \in \mathcal{S}_n, & \text{fault-free} \\ H_1 : \omega \in \mathcal{S}_f, & \text{faulty} \end{cases} \tag{12}$$

where H_0 and H_1 are the null hypothesis and alternative hypothesis, respectively.

To evaluate the performance of fault detection (FD) systems for high-speed trains, two commonly used indices depending on (12) are the missing alarm rate (MAR) and false alarm ratio (FAR), as defined below

$$\begin{aligned} &\text{FAR: P}(H_1|\omega_n) \\ &\text{MAR: P}(H_0|\omega_f) \end{aligned} \tag{13}$$

where $\mathrm{P}(\cdot)$ means the probability. Naturally, a proper boundary to faithfully distinguish $\mathcal{S}_n$ and $\mathcal{S}_f$ is preferred, based on which a satisfactory tradeoff between FAR and MAR could realize.

In practical FDD tasks for traction systems, a test statistic J is usually defined on z in Model III, or on ω in (11), or on residuals generated from Model II and VI. The commonly used terms J are with different choices dependent upon purposes. Some of well-known Js include T^2 test statistics [47], the squared prediction error (SPE) [51], Kullback-Leibler divergence [52], mean value [47], 2-Norm [53], trace [54], entropy [55], Hellinger distance [56], along with some combined indexes [57]–[60].

Here, we introduce two common test statistics adopted in the FDD applications to traction systems. For the obtained ω, T^2 test statistic is defined in the form such that

$$T^2(\omega) = \omega^T \Sigma_\omega^{-1} \omega \tag{14}$$

where Σ_ω is the covariance matrix of ω; and another is the SPE defined to be

$$\mathrm{SPE}(\omega) = \omega^T \omega. \tag{15}$$

For example, T^2 used in (12) is that

$$\begin{cases} H_0 : T^2(\omega) \le J_{th,T^2}, & \text{fault-free} \\ H_1 : T^2(\omega) > J_{th,T^2}, & \text{faulty} \end{cases} \tag{16}$$

where J_{th,T^2} is the threshold of T^2, responsible for setting the region or boundary between H_0 and H_1. In this scenario, to judge whether there is a fault appearing in traction systems or not could be theoretically transformed into the following optimization problem:

$$\begin{aligned} &\max\ \mathrm{P}(T^2(\omega) > J_{th,T^2}|\omega \in \mathcal{S}_f) \\ &\text{s.t. } \mathrm{P}(T^2(\omega) > J_{th,T^2}|\omega \in \mathcal{S}_n) = \alpha \end{aligned} \tag{17}$$

where α is the significance level equal to FAR numerically. More details on the solution of (17) can be found in [61].

It should be especially mentioned that, the choice of thresholds used for monitoring operating conditions of traction systems in high-speed trains mainly relies on the expert knowledge via abundant experiments (the so-called trial and error), where the maximum of J is a critical test statistic and the filtering technique is usually involved in practice [1], [15], [34], [62]. These manners tip us off to the fact that the flood of FARs in practical FDD applications to high-speed trains is firmly prohibited, even if they are loss of physical and mathematical basis in some sense [48].

Extensive literature can be found in [11], [15], where T^2 and SPE are the most commonly used fault indicators [63]–[68] to complete FD tasks for traction systems. For discovering additive faults in traction systems, T^2 has been mathematically proved, following the Neyman Pearson lemma, to be of the best fault detectability [44]. For other scenarios, the test statistics, such as SPE [53] or Kullback-Leibler divergence [69], would show superior FD performance and therefore are gradually utilized when performing FD tasks for traction systems in high-speed trains. Besides, the generalized likelihood ratio is also attractive in practice [70], [71].

B. Principal Component Analysis

Since the early studies about principal component analysis (PCA) by Person in 1901 [72] and Hotelling in 1933 [73], it is widely used for reducing the dimensionality of original data while preserving as much of the variation information as possible in a lower space [74]. Nowadays, PCA and its variants-based FDD solutions have benefited many areas such as chemical engineering [75], electrical engineering [76], and intelligent transportation [15] significantly.

When high-speed trains operate under a steady condition, z given in Model III described by (4) can be viewed as a set of random signals following a certain but unknown distribution. Then the objective function of PCA regarding the normalized z can be defined as

$$\begin{aligned} &\arg\ \max_p\ p^T \Sigma_z p \\ &\text{s.t. } p^T p = 1 \end{aligned} \tag{18}$$

where Σ_z is the covariance matrix of z, and p is the loading vector. By introducing a Lagrange multiplier and eigen-decomposition [75], it is straightforward from (18) that

$$P = [P_p \;\; P_r] = [p_1 \;\ldots\; p_{l+m}] \in R^{(l+m)\times(l+m)}$$
$$\Lambda = [\Lambda_p \;\; \Lambda_r] = \mathrm{diag}(\sigma_1^2 \;\ldots\; \sigma_{l+m}^2) \in R^{(l+m)\times(l+m)} \quad (19)$$

where p_i and σ_i^2 are the eigenvector and eigenvalue of Σ_z, respectively; P_p and P_r are the loading matrices in principal and residual subspaces, respectively; and $P_p \in R^{(l+m)\times n_p}$, $P_r \in R^{(l+m)\times(l+m-n_p)}$, $\Lambda_p \in R^{n_p\times n_p}$, $\Lambda_r \in R^{(l+m-n_p)\times(l+m-n_p)}$ for the number of principal components n_p. In fact, P and Λ can also be obtained via performing the singular value decomposition on Σ_z [57]. Define the score matrix such that $T = [t_1 \;\ldots\; t_{l+m}]$, where $t_i = P^T z$. The numeral relationship between variation (or called uncertain) information and t_i holds:

$$\sigma_i^2 = \frac{1}{N_z - 1} t_i^T t_i = \mathrm{var}\{t_i\} \quad (20)$$

where N_z is the number of z.

After PCA-based dimensional reduction and modeling, operation variations of traction systems will be mainly reflected in the principal component subspace [64], i.e.,

$$\hat{z} = P_p t = P_p P_p^T z. \quad (21)$$

Accordingly, the residual subspace depicts non-systematic information of traction systems, which is stated as follows

$$\tilde{z} = z - \hat{z} = (I - P_p P_p^T) z. \quad (22)$$

Because of the orthogonality of two subspaces obtained above, it holds that [77]

$$\hat{z}^T \tilde{z} = 0. \quad (23)$$

Considering a fault f appears in traction systems of high-speed trains as given in (5), anomalous behaviors will have influence on $\hat{z}$ in (21) or $\tilde{z}$ in (22). Regard $\hat{z}$ and $\tilde{z}$ as ω in (11), on which both T^2 and SPE can be defined to monitor the operation conditions of traction systems.

Thanks to its salient strengths such as dimension reduction and feature extraction, PCA has become one of the most fruitful methods used for the FDD applications in traction systems of high-speed trains. Early adoption of PCA is reported in [78] to detect stator faults in general traction motors using three-phase currents. After that PCA, together with its variants, has been abundantly developed for the FDD purposes [79].

For example, combining with random forest, a kernel PCA is proposed in [80] to detect and diagnose faults in traction transformers of high-speed trains. To discover incipient abnormalities appearing in insulated-gate bipolar transistors (IGBTs), a multi-mode kernel PCA is thus developed in [36] for high-speed trains. Recently, a modified kernel PCA is also proposed to detect composite faults by considering and combining with time-frequency features [81].

It is a fact that, the practical characteristics of traction systems render the FDD application of traditional PCA unreliable and impracticable [11]. For example, fault decisions including detection and diagnosis of faults are often needed online or in nearly real-time; it is not able to complete online FDD for high-speed trains using kernel PCA-related methods because the sampling frequency of sensors is very high, as presented in Section II. Therefore, most of the extended FDD methods based on the modified PCA are linear for seeking low computation loads. For detecting and estimating incipient faults in traction systems, the Kullback-Leibler divergence is respectively introduced into linear PCA framework in [82] and [83], where the improved computation efficiency is achieved by using two coordinate transformations when the probability density functions of signals are estimated. Differently, [84] proposes a fast algorithm based on PCA and Gaussian transform to perform FDD tasks. Most recently, [19] designs a fault-relevant PCA method for traction systems of high-speed trains, which provides a better FD performance by modeling multiple faults.

C. Independent Component Analysis

In order to deal with the measured non-Gaussian signals, independent component analysis (ICA), regarded as a typical modification of PCA concerning non-Gaussian features [85], is introduced to detect and diagnose faults in traction systems [86]. It can search a linear combination with higher-order information hidden in the signals measured from high-speed trains [87]. It is remarkable that, ICA is responsible for extracting higher-order statistics by extracting independent components from actual measurements, and then recovering original signals as much as possible [57].

As mentioned in [11], [15], [63], [86], traction systems of high-speed trains operate with obvious non-Gaussian characteristics. The non-Gaussian information thus plays a vital role in FDD applications (such as system modeling and threshold setting) for traction systems in high-speed trains.

Considering the measurement variable z in Model III, ICA tries to extract the hidden statistically independent components via the following objective function:

$$\hat{s}(k) = W z(k)$$
$$\text{s.t. max} \;\; \mathcal{M}(\hat{s}(k)) \quad (24)$$

where $\hat{s}$ is the estimation of independent components; $\mathcal{M}$ is the non-Gaussian measurement function which can be mathematically defined by maximum likelihood estimation, minimization of mutual information, tensors, etc. [87]; W is the so-called demixing matrix that ICA is to find; and $\hat{s} \in R^d$, $W \in R^{d\times(l+m)}$, $d < l+m$. Among different $\mathcal{M}$s, there are existing connections and bridges in fact which can be easily checked and found using the likelihood.

By the use of PCA given in (19) for removing the effect of first- and second-order statistics, the whitening-transformation, $Q \in R^{(l+m)\times(l+m)}$, can be computed as [88]

$$Q = \Lambda^{-1/2} P^T \quad (25)$$

which delivers a linear projection such that $x(k) = Qz(k)$ satisfying $\mathrm{var}(x(k)) = I$ [11].

According to $\mathcal{M}$ defined in (24), an orthogonal matrix B can be obtained and then used to estimate the latent independent components, i.e.,

$$\hat{s}(k) = B^T x(k) = W z(k) \tag{26}$$

where

$$W = B^T Q \tag{27}$$

which forms the solution of ICA.

Here, W can also be regarded as the operating rule g defined in (11). If there is a fault appearing in traction systems of high-speed trains, T^2 and SPE can be defined on both $\hat{s}$ or residuals of $z(k)$ to perform FDD tasks [89].

The unique feature of ICA is that it performs well for non-Gaussian signals of traction systems via extracting higher-order statistical information. This has naturally catalyzed ICA-based FDD applications to high-speed trains.

Motivated by the study in [89], an improved ICA method is investigated in [86] to detect incipient faults in traction systems in real-time. Meanwhile, the ICA-based strategies are then used to perform FDD tasks for the traction transformers in [90] and [91]. Combing with artificial neural networks, a neural ICA is designed for FDD of traction gearbox in [92]. By establishing a set of sub-models, a modified ICA is designed in [93] to detect and isolate sensor faults in traction systems. In addition, some extended applications of ICA for high-speed trains can also be found such as digital modulation identification under impulsive noise conditions [94].

D. Canonical Correlation Analysis

Canonical correlation analysis (CCA) is originally proposed to measure the linear relationship between two sets of variables [95]. Regarded as a fundamental technique, it is extensively applied to identification, filtering, and control for both static and dynamic systems by the use of the linear or nonlinear relation among past, present, and future samples [96]. Its computation primarily involves one of the most numerically stable singular value decomposition instead of the Kalman filter used in other system identification techniques. It is a typical multivariate statistical analysis method, based on which the FDD methods [47] together with their applications in high-speed trains [63], [97], [98] are recently developed.

The CCA tries to find basis vectors, defined as $w_u \in R^l$ and $w_y \in R^m$, such that onto which the correlation between projections of u, y is mutually maximized [99]. Consider the mean-normalized z in (4) under steady conditions, the objective function of CCA can hence be defined in the following general form [100]:

$$\begin{aligned} \arg \max_{w_u, w_y} \quad & w_u^T \Sigma_{u,y} w_y \\ \text{s.t.} \quad & w_u^T \Sigma_u w_u = 1 \\ & w_y^T \Sigma_y w_y = 1 \end{aligned} \tag{28}$$

where Σ_u, Σ_y, and $\Sigma_{u,y}$ are the covariance matrices of u, y, and between u and y, respectively. By introducing two Lagrange multipliers such as λ_u and λ_y, the maximization problem in (28) with two equality constraints becomes

$$\begin{aligned} \mathcal{L}(w_u, w_y, \lambda_u, \lambda_y) = w_u^T \Sigma_{u,y} w_y - \frac{\lambda_u}{2}(w_u^T \Sigma_u w_u - 1) \\ - \frac{\lambda_y}{2}(w_y^T \Sigma_y w_y - 1). \end{aligned} \tag{29}$$

In respect to w_u and w_y, taking partial derivatives yields

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w_u} &= \Sigma_{u,y} w_y - \lambda_u \Sigma_u w_u, \\ \frac{\partial \mathcal{L}}{\partial w_y} &= \Sigma_{y,u} w_u - \lambda_y \Sigma_y w_y. \end{aligned} \tag{30}$$

Let $\partial \mathcal{L}/\partial w_u = 0$ and $\partial \mathcal{L}/\partial w_y = 0$. It implies that w_u, w_y used in (30) are the solution of CCA defined in (28) and $\lambda_u - \lambda_y = 0$. Without loss of generality, we can assume Σ_y is invertible and $\lambda = \lambda_u = \lambda_y$. Simple derivations result in

$$\Sigma_{u,y} \Sigma_y^{-1} \Sigma_{y,u} w_u = \lambda^2 \Sigma_u w_u \tag{31}$$

and

$$w_y = \Sigma_y^{-1} \Sigma_{y,u} w_u / \lambda. \tag{32}$$

By solving a generalized eigenvalue problem left in (31), λ together with w_u and w_y can be therefore obtained, where w_u and w_y form the co-ordinate system maximizing the correlation between $w_u^T u$ and $w_y^T y$. Accordingly, considering that W_u, W_y, and Σ consist of w_u, w_y, and λ, the projection g in (11) can be described in the following forms:

$$\begin{aligned} g_1 : \omega_1(k) &= W_u^T [I \quad 0] z(k) - \Sigma W_y^T [0 \quad I] z(k) \\ \text{and } g_2 : \omega_2(k) &= W_y^T [0 \quad I] z(k) - \Sigma^T W_u^T [I \quad 0] z(k) \end{aligned} \tag{33}$$

where Σ may not be the square matrix, and $\dim(\Sigma) = \min(l, m)$. As described in (5) that an unexpected fault $f(k)$ appears in traction systems of high-speed trains from the k-th sampling time, both T^2 and SPE can be applied for online FDD tasks, based on which abnormal derivations in $\omega_1(k)$ or $\omega_2(k)$ caused by $f(k)$ can be detected [63].

The attractiveness of CCA-based FDD applications to traction systems of high-speed trains lies in, by reducing the uncertain information, the improved fault detectability [101], [102]. Besides, CCA has proved its superior FD performance for detecting sensor faults that do not propagate in closed systems [63]. It is hence worth further investigations on CCA-based FDD for traction systems of high-speed trains, where extra bonuses could be delivered by consideration of correlations among different samplings or different variables of traction systems in high-speed trains [11].

To perform FD tasks for traction systems, a generalized CCA is developed in [63] aiming at achieving maximized fault detectability for non-Gaussian trains, where optimal threshold setting is obtained by a randomized algorithm. To achieve better FD performance for traction systems, the moving average technique is introduced into the CCA framework to make new residuals more sensitive to unexpected faults [97]. Most recently, CCA is adopted to formulate a modified BLS in [103], where fast FDD for traction systems of high-speed trains can be implemented with considerable scalability. In [104], by consideration of the correlation between different measurement variables, CCA is thus used to detect faults in traction circuits of high-speed trains, and the extracted fault directions are then used for fault isolation. Besides that, to the ensure safety of high-speed trains, [105] develops a data-driven monitoring method by the use of autocorrelation of measurable signals. Afterwards, the CCA-based dynamic FD approach is developed, according to Model IV given in (7), in [98] to identify key matrices based on which residual signals are therefore effective for dynamic traction systems.

E. Other Machine Learning Techniques

Expect for the mentioned methods above, other techniques, including random forest [80], manifold learning [106], k-nearest neighbors [107], partial least squares [66], neighborhood preserving embedding [108], and so on [109], are also popular subfields of machine learning, and have been demonstrated their successful FDD applications to traction systems of high-speed trains, when data acquired from traction systems under both health and faulty conditions are necessarily available.

Reinspecting on all data-driven FDD methods for traction systems of high-speed trains, they can be broadly partitioned into two general categories, i.e., unsupervised learning-based and supervised learning-based FDD methods [15]. Specifically speaking, the former completes data modeling (or called feature extraction) via discovering hidden structures from measurements without labels [110]; and then the test statistics are responsible for defining normal operating regions for traction systems [111]. In this sense, sufficient data under normal conditions of traction systems should be known beforehand [11]. In parallel, the latter deal with labeled measurements from traction systems, whose core is the so-called classification or regression techniques. Depending upon labels of data, these methods are adopted to discover common features for the same conditions including normal and fault scenarios, and then directly perform FDD tasks for traction systems of high-speed trains [15].

The most popular unsupervised learning ones have already been reviewed in previous content. The following focus will be on the supervised learning-based FDD methods together with their applications to traction systems of high-speed trains. Denoting $z_f(k^\kappa)$ as data belonging to fault statuses, the off-line data sets corresponding to (5) can be described as

$$\left\{z_n(k^0), z_f(k^1), \ldots, z_f(k^\kappa), \ldots, z_f(k^K)\right\} \tag{34}$$

where $\kappa = 1, \ldots, K$, $k^i = 1, \ldots, N^i$, and $i = 0, 1, \ldots, K$. Here the subscript “0” refers to normal conditions, and other subscripts refer to different fault conditions for traction systems of high-speed trains.

Similarly, the operating rule g can be defined on the whole data space (including z_n and z_f) to obtain new features/projections, namely ω_n and ω_f. Being consistent with (12), we can redefine $\mathcal{S}_n$ as $\mathcal{S}_0$, and then divide $\mathcal{S}_f$ into K sub-regions accordingly for describing all fault scenarios of traction systems, i.e., $\mathcal{S}_1, \ldots, \mathcal{S}_K$. Obviously, $\mathcal{S} = \mathcal{S}_0 \bigcup \mathcal{S}_\kappa = \bigcup_{i=0}^{K} \mathcal{S}_i$ covers all operation regions.

Therefore, the false diagnosis ratio (FDR) for the i-th case, which is similar to the definition of FAR, is determined by

$$\mathrm{FDR}_i : \quad \frac{\text{No. } (\omega \in \mathcal{S}_{\bar{i}} | \omega \in \mathcal{S}_i)}{N^i} \tag{35}$$

where $\bar{i} = 0, 1, \ldots K$ and $\bar{i} \neq i$. It can hence be asserted that, by minimizing the FDR_i as much as possible, the supervised learning methods are then developed to achieve successful FDD tasks for traction systems of high-speed trains.

Ideally, the supervised learning method that leads to minimizing all FDRs is the best. Unfortunately, this objective is highly challenging. In practical FDD applications to traction systems of high-speed trains, it is reasonable to accept a small number of false diagnosis results. Therefore, the general objective that all supervised learning methods pursue is

$$\begin{aligned} \min \quad & \mathrm{FDR}_i \\ \text{s.t.} \quad & \sum_{\bar{i}=0}^{K} \mathrm{FDR}_{\bar{i}} \leq \beta \end{aligned} \tag{36}$$

where β is a small constant related to FDR. More specifically, these supervised machine learning methods try to find g such that the objective defined in (36) can be achieved. Here, g could be the projections in both fixed and various forms, and could also be the structures such as a set of linked neurons in deep learning methods.

For example, g in the support vector machine (SVM) method is a fixed projection that maps the original data into high-dimensional space where the obtained hyperplane solves (36). By analyzing vibration signals of high-speed trains, a modified SVM is proposed in [112] based on the correlation hidden in the time-series data. To test the performance of gearbox, an SVM-based classifier is designed in [113] based on the selected principal components. Cooperating with empirical mode decomposition, a least-squares SVM is proposed for detecting sensor faults in traction systems of high-speed trains [114]. A similar approach is also recently found based on SVM in the time-frequency domain for the online FDD application to traction systems of high-speed trains [115]. For isolating faults in traction systems, an efficient strategy based on SVM is developed in [19] after successful FD procedures. By consideration of the unbalanced data problems of high-speed trains, [116] designs a modified SVM for reliable FDD purposes. To tackle the FDD issue for dynamic traction systems with unbalanced data, a multi-classification SVM is studied in [98] by adding constraints on penalty parameters.

In the Bayesian inference, g is the manner of transforming original data into their probability or distribution information [117], [118]. In light of this information, Bayes' theorem details the predictive possibilities about FDD results for traction systems of high-speed trains [11]. Absolutely, the goal also satisfies (36). After successful detection of faults in traction systems of high-speed trains, a supervised Bayesian classifier is designed in [17] to distinguish different fault statues. By establishing probability matrices from test statistics, [119] proposes a Bayesian method for diagnosing incipient faults in traction systems, where the strong generalization capacity is achieved by weakening effects caused by different fault amplitudes. In [120], successful FDD for traction systems is conducted with a Bayesian network. Recently, a Bayesian network is developed for monitoring of high-speed trains with consideration of noise and missing data [121].

Besides, an attractive research interest could be observed based on recent literature, in which FDD tasks for traction systems of high-speed trains are intensively investigated by using (artificial) neural networks [15], such as deep learning [122] and broad learning [123]. Both of them can be called an "end-to-end" manner, which can learn the hidden representations automatically between raw measurements and FDD results [124]. For this kind of methods, g is the hyper parameter (which can be regarded as neurons) driving the whole network to complete FDD tasks for traction systems of high-speed trains according to (36). Based on the convolutional neural network, [125] is dedicated to the FDD applications in high-speed trains when imbalanced data problem is present. On the basis of the restricted Boltzmann machine, a deep belief network is designed in [126] to achieve FDD tasks for on-board equipment of high-speed trains. In [127], a residual learning scheme is proposed, via processing vibration signals, for diagnosing faults in rotating traction motors. To identify the root cause of faults appearing in high-speed trains, an alternative algorithm based on neural networks is proposed in [128] by sufficient use of unstructured data. To the best of our knowledge, to achieve FDD purposes for traction systems of high-speed trains, the update of g usually comes at the cost of high computation loads and storage overhead. To mitigate this situation, a fast FDD framework using broad learning is well designed in [103] where accurate FDD results for traction systems of high-speed trains can be realized in a real-time fashion because of its considerable computation efficiency.

IV. CHALLENGES

Different from data-driven FDD techniques in other fields, traction systems of high-speed trains have their natures which must be taken into account in the practical design of data-driven FDD schemes. These natures pose considerable challenges and barriers that should be considered and be further eliminated. As shown in Fig. 3, these challenges become tricky in practice, especially for dynamic traction systems. This section will comprehensively present these challenges according to the features of measurements from traction systems, based on which we think that both theoretical and practical researchers will benefit.

A. Non-Gaussianity

As especially pointed out in [15], [19] and [63], the data measured from traction systems are non-Gaussian distributed. For example, three-phase currents, which are inputs of traction motors, are the sinusoidal signals [129]. This fact leads to periodic variations that are related to the preset space vector pulse width modulation, although the train operates in one fixed point [11]. Besides, signals in the rotatory machines such as traction systems are usually with obvious periodicity.

Therefore, in the data-driven FDD designs together with their application in high-speed trains, two phases that should

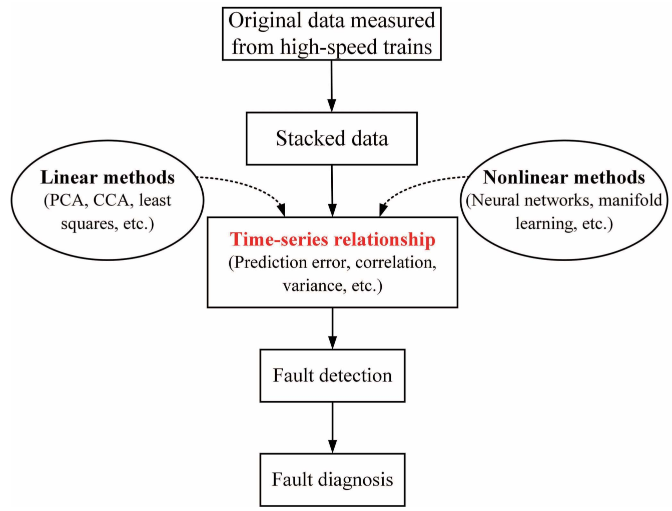

Fig. 3. A schematic description of data-driven dynamic FDD methods for traction systems in high-speed trains.

be paid special attention to are: 1) the accurate modeling/extraction of non-Gaussianity hidden in original measurements; and 2) the proper choice of threshold.

The underlying theoretical assumption behind the direct use of traditional PCA [57] and CCA [47], which is not explicitly mentioned, is that all measured variables obey Gaussian distributions. In fact, the Gaussian assumption on z in Model III for traction systems is far from being realistic [130]. For this case, traditional PCA, as well as traditional CCA, is arguably not an optimal solution used for FDD applications in traction systems of high-speed trains.

Meanwhile, this non-Gaussianity is also reflected in test statistics used for judging whether there is a fault appearing in traction systems or not [11]. It naturally increases difficulties in choosing proper thresholds in FD phases [63]. For example, kernel density estimation is the commonly used technique [64] when the practitioner determines thresholds of test statistics (or the boundary of $\mathcal{S}_0$).

B. Nonlinearity

The nonlinearity of traction systems in high-speed trains is reflected in two aspects. One is the nonlinearity among the past, current, and further operation states of traction systems [131], i.e., the nonlinear relationship among $x(k-1)$, $x(k)$, and $x(k+1)$ given in Model II. As summarized in [64] and [132], for a general traction system, the linearity of a magnetic circuit, the neglect of core loss, symmetrical assumptions on stator currents, etc., are usually used in establishing Model II in (2). Describing a traditional train will involve 84 mathematical equations at least [133], and this work is basically done only if multiply assumptions of linearity exist [11]. For the modern high-speed trains, establishing their mathematical expressions based on the first principles must be of more difficulties and needs more assumptions on linearities. In addition, another aspect is the nonlinear relationship among different variables [15], i.e., the nonlinearity among $u(k)$ and $y(k)$ given in Model III. For example, without loss of generality, the alternating opening and closing IGBTs in traction systems cause the nonlinearity among the intermediate voltage and three-phase output currents [36].

To achieve successful FDD tasks for traction systems of high-speed trains, two nonlinear relationships [11], among system states ($x(k-1)$, $x(k)$, and $x(k+1)$) and among variables ($u(k)$ and $y(k)$), are of the main interest. This is because the more accurate models mean the fewer uncertainties, leading to more satisfactory FDD performance. It is a loss of meaning if traditional multivariate analysis techniques are directly used without any modifications. Therefore, this unsettled challenge should deserve an increasing focus of attention.

C. Dynamic Operation

Firstly, the high-speed trains work under six conditions during conventional operations: starting, speeding, traction, coasting, braking, and stopping [133]. Secondly, trains work in different traction speeds in order to cater to real-time circumstances [1]. For example, the drivers will reduce speeds before trains go through the tunnel. Last but not least one is the externally time-varying disturbances [15] such as the unpredictable winds [134]. Therefore, the dynamic operation of traction systems can be easily found and checked.

It can be readily observed from the existing publications that most of the advanced data-driven FDD methods for traction systems are static. These methods work well when traction systems of high-speed trains operate around a fixed point [15]. When the high-speed trains operate in practice, the measurements collected from traction systems are highly autocorrelated and also timely dependent [48]. It is therefore inappropriate for direct extensions of these methods to online FDD of traction systems.

For example, the popular contribution plot-based diagnosis of faults [75] in traction systems is out of the question, since the well-designed control strategy (i.e., double-closed loops including the current loop and speed loop) in TCU makes the whole system suffer from smearing [15]. In fact, the dynamic FDD designs are desirable to tackle this challenge [46], [135]. However, to obtain (2) via the first principle, a deep understanding of the operational mechanism together with a highly sophisticated modeling is indispensable or is not feasible. Hence, more attention should be devoted to the dynamic operation when engineers design online FDD algorithms for traction systems of high-speed trains.

D. Big Data

Thanks to advanced sensor technology, a large number of sensors are mounted on high-speed trains for various purposes, such as real-time control and monitoring as well as follow-up maintenances. However, the accompanying problem is the so-called big data in storage, management, processing, etc. [48]. More specifically, the big data problem [136], which poses a great challenge on FDD tasks for traction systems, arises because of the following four factors: long operations [7], large-scale measurement points [1], high sampling frequency [11], and different data sources [137].

For example, there are more than 1000 sensors mounted on CRH380-type high-speed trains [4]; and the number of sensors may be more than 2000 in one train group with 16 carriages [137]. Such plenty of distributed sensors make data come from different levels [66] (i.e., device level, vehicle level, and train level), resulting in the mixture of structured and unstructured data [138]. In addition, the sampling time of sensors equipped in traction systems is 4×10^{-4}s [1]. It is also a fact that most of trains have long operations per day. Due to these reasons, the big data problem in FDD applications to high-speed trains, which is basically a double-edged sword [139], has received increasing attention [107], [140].

Despite the current advances such as edge computing [48] and map-reduce technique [141], real-time monitoring, diagnosis, together with prediction of faults in high-speed trains have not been fully studied, and of great interests in practice.

E. Heterogeneity

Coming from practical FDD applications in traction systems of high-speed trains, the survey paper [15] points out the "heterogeneity" problem for the first time. The heterogeneity refers to dissimilarities in data from one type- to another type-trains, from experimental setups to practical high-speed trains, from different time-in-service trains, from one vehicle to the other in a train group, etc [142]. Naturally, it is different from noises and disturbances affecting high-speed trains because the heterogeneity makes the data model migrate unexpectedly from one normal case to another normal one.

Most of the existing FDD methods for high-speed trains could not be capable of automatic update with such heterogeneity, which considerably limits the effectiveness and flexibility of these FDD methods. Therefore, an uniform data-based model together with, based on which, the designed FDD method are preferable and valuable by considering the variability/teterogeneity as especially mentioned, although this validation will be a lengthy endeavor [11].

F. Real-Time Ability

It is our ultimate goal that the developed FDD approaches can work for traction systems of high-speed in a real-time fashion because fast and effective decisions including detection and diagnosis of faults are often needed online or in near real-time [11]. For example, for the high-sample-frequency sensors mounted on traction systems, the FDD system, which needs to make a decision less than one sampling time (i.e., 4×10^{-4}s), should be of sufficiently high computation efficiency [34].

Of course, the most intuitive way to tackle this challenge is increasing the time internals of both off-line and online data from traction systems. However, the increased time internals should satisfy the Nyquist sampling theorem such that the sparse data can reflect or reconstruct the original data as much as possible. Other ways are to improve the computation efficiency of FDD methods for traction systems. This is the main reason why kernel-based FDD methods are impracticable for performing FDD tasks for traction systems of high-speed trains. Moreover, a similar problem also arises in dynamic FDD methods. As given in (6) and (7), a large s contributes to satisfactory identification accuracy as well as improved FDD performance, but simultaneously resulting in considerable computation loads. Therefore the real-time ability should be a premise for the advanced FDD methods in this field [15].

Based on the thoroughly-analyzed and well-described challenging issues above, perspectives of FDD applications in taction systems of high-speed trains will be unfolded in the subsequent section.

V. Perspectives

Despite the current progress of these advanced methods made so far, data-driven FDD designs for traction systems as well as for the whole trains are still in their embryonic forms. Therefore, an exhaustive trend prediction on the potential FDD methods is perceptively made in this section. We truly believe that these perspectives will provide engineers with some instructive and valuable guidance.

A. Model Migration Against Variabilities

The presence of considerable heterogeneity necessitates investigations of some migration methods [143]. A special objective of model migration considered in FDD studies for traction systems of high-speed trains is to uncover the common features/knowledge that can benefit each different and individual train. From this viewpoint, there are at least three promising directions, as summarized follows.

1) Transfer learning (or called knowledge transfer [144]): The research on transfer learning-based FDD applications in traction systems of high-speed trains is motivated by the fact that the engineers can intelligently find, learn, summarize the knowledge hidden in previously collected data sets, and then apply these knowledge to solve FDD problems in different scenarios with satisfactory performance. This technique can work well for the FDD tasks of traction systems with variabilities, because it retains and reuses the previously learned knowledge in our expected manner [145].
2) Robustness: In the control field including the FDD subfield, robustness is a property against or tolerating perturbations that might affect system performance [119]. If the heterogeneity of traction systems can be studied in-depth, it is expected that FDD methods with robustness could work sufficiently well against variabilities [146].
3) General data models: Focusing on the universal behaviors of traction systems, it is desirable to develop a general data model that can cover a wide spectrum of multiple individuals, and that can disregard the dissimilarities at the same time [147]. The most intuitive (but not optimal) idea is to assign weighting factors for both similarity and dissimilarity.

Successful designs of anyone above will greatly expend the feasible horizon of FDD methods for traction systems.

B. Artificial Intelligence-Based Methods

Artificial intelligence-aided FDD methods can carry on the FDD tasks for traction systems in a human-like way. This therefore leads to the coming wave of popularity in artificial intelligence-based FDD methods for traction systems [11], especially the neural networks-based strategies [97], [103], [140], [148], [149]. To be precise, artificial intelligence is a more generalized category than machine learning [150]. By the use of qualitative knowledge representations, the belief rule base can be classed as artificial intelligence, rather than machine learning. Its advantages of FDD applications in high-speed trains are shown in [151].

Taking neural networks as an example, we will detail research opportunities in light of their advantages which are evidenced by the existing FDD studies.

Critically neural networks are well suited for handling nonlinear and non-Gaussian signals of traction systems because of their superior abilities in dealing with nonlinearity and non-Gaussianity. Based on these accurate relationships among different variables in (4), data-driven modeling, based on which the FDD applications in traction systems of high-speed trains, will become easier and more reliable.

Meanwhile, the focus of neural networks can be on the relationships among different states of traction systems in high-speed trains. By this way to discover dynamic behaviors of traction systems, neural networks will be of great applicabilities in the FDD studies. For example, by the only use of I/O data from traction systems, neural networks can generate a projection from $U_{k,s}$ to $Y_{k,s}$ given in (7), based on which the residual generator can be designed for the FDD application.

These possibly valuable directions on data-driven FDD methods for high-speed trains deserve attention, which is firmly based on the premise that both computation ability and interpretability of neural networks can be improved.

C. System Identification-Based Methods

Let the considered traction system of high-speed trains under normal operation be specified by the Model II given in (2), the precondition of designing residual generators (used for FDD purposes) suitable for dynamic traction systems is that system matrices are known. Rather than establishing Model II by first principles to determine parameters in (2), system identification can provide an alternative solution [152] by the use of stacked vectors and matrices given in (6). For example, motivated by [46], [152], [153], the latest research [11] and [48] respectively mention and develop system identification-based FDD strategies for dynamic traction systems of high-speed trains. These updated results are a powerful indication that system identification to obtain full parameters or partial/key parameters are more powerful than some static approaches when high-speed trains are on active service.

It is an indisputable fact that to acquire the causality based on the knowledge or experiences from engineers of high-speed trains is an impossible task or a troublesome one at least. For the complex traction systems as well as the whole high-speed trains, identification of system matrices from stacked variables/data to reduce the requirement on causality is hence particularly important and is easier to be achieved. Successful realization means that the propagation path of faults could be tracked and the interactive influence caused by faults could be discovered. Besides, an additional advantage of system identification is that it can be performed without caring too much about the non-Gaussian problem of measured signals in traction systems.

D. Non-Gaussian and Nonlinear Methods

As presented in Section IV, measured signals of traction systems in high-speed trains are characterized by the considerable non-Gaussianity and nonlinearity [11], on which the focus will be helpful for modeling traction systems and hence for improving FDD performance.

To a certain extent, the methods (such as SVM and ICA) will be of great advantage for dealing with non-Gaussianity of traction systems. Instead of these solutions, there is another attempt that transforms original signals of traction systems into features approximately obeying Gaussian distributions, where two coordinate transformations are adopted in the data preprocessing stage [82], [83]. Besides that, some proper data transformations making data modeling more efficient are appreciated in practice. For example, [154] readily establishes the relationship between Gaussian and non-Gaussian signals.

Study on nonlinearity among both signals and time-series measurements plays an important role in the improvements of FDD performance for traction systems of high-speed trains. Recalling that kernel-aided data-driven schemes must result in the dimension explosion problem when the number of training data is sufficient. Hence, the best remedy for handling the nonlinearity problem is to find alternative tools or to modify the linear data-driven methods.

In accordance with both the non-Gaussian and nonlinear characteristics of measured signals, many efforts should also be spent on redefining some new test statistics as well as their proper boundaries [15].

E. Fault-Tolerant Techniques

Closely associated with the safety and reliability of high-speed trains, the fault-tolerant control involves, after successful detection and diagnosis of a fault, reconfiguring the controller structure or accommodating the controller parameters in TCU so that the safe operation of high-speed trains can be guaranteed. In the designs of such kinds of fault-tolerant control methods, the stability and safety of traction systems in high-speed trains are essential requirements [155]. In fact, the PI controller widely used in current high-speed trains has the fault-tolerant ability, which can eliminate performance deterioration to some degree caused by faults with small amplitudes. To further improve the safety of high-speed trains such as to keep the degraded control performance within the acceptable level and meanwhile to preserve stability conditions, an effective fault-tolerant control strategy for traction system is necessary.

Of course, the classical PI controller will not be replaced up to now because its feasibility and practicability are sufficiently demonstrated on currently developed trains. Besides, the replacement of current PI controllers leads to great economic costs. It is a natural and engineering way to perform the fault-tolerant control task based on the current PI controllers. For example, by the use of collected data from traction systems, the plug-and-play technique-aided fault-tolerant methods can be developed, which optimizes the operation performance of trains in the presence of faults but leaves the pre-designed PI controllers untouched. Study [156] shows the plug-and-play technique to be potentially more efficient and feasible than their conventional fault-tolerant control counterparts. These similar strategies, which are developed based on original controllers, will be of great interest in future research.

F. Parallel Monitoring

In one region or one country, there may be lots of high-speed trains in service simultaneously. For example, more than 5000 trains are on active service in China every day. It makes the centralized decisions (including detection and diagnosis of faults) online or in near real-time via one ground control center impossible. The parallel monitoring can assign these FDD tasks into every train. By this means plenty of data can be processed simultaneously to achieve real-time monitoring of all high-speed trains [48].

Thanks to the intercommunications between control centers and rolling stocks, parallel monitoring of multiple traction systems as well as trains becomes possible, under this framework the control theory (together with system identification) will reshine because of its abilities in dealing with dynamic behaviors and disturbances of traction systems. From this view of point, parallel monitoring, together with parallel diagnosis, can be regarded as a new computing paradigm, which can embrace various FDD methods [157].

Up to now, its infancy can be observed from the existing publications. For example, the study dedicated to parallel processing algorithms can be found in [158]; and the edge computing-aided FDD scheme is predicted in [48]. It seems that the edge computing technique may be more effective than, for example, the traditional cloud computing. It is expected that many valuable investigations will emerge in the future.

G. Other Issues

Apart from the aforementioned perspectives, there are other open problems deserving more investigations in depth, which are helpful for ensuring and improving the safety of traction systems as well as high-speed trains. Owing to space constraints, these studies are briefly summarized as follows.

1) Sensor allocation: Effective sensor allocation makes availability of effective and sufficient information about traction systems that can be captured for monitoring traction systems as well as the railway [159].
2) Semi-supervised learning-based FDD strategies: These schemes are attractive [160]–[162] for FDD applications to traction systems when the acquisition of explicit labels of data sets is achieved at the high cost of time and human efforts or is very difficult to achieve [163].
3) Reinforcement learning-based FDD strategies: By referring to an agent and receiving observations and reward, reinforcement learning makes FDD decisions for traction systems based on sequences of process, which can maximize the FDD objectives via a trial-and-error learning manner [164].
4) Forecast abnormal events: It forecasts events which will endanger the traction systems as time goes on, and can

predict the remaining useful life of key components (such as IGBTs [165]) in traction systems. It belongs to the prophetic manner to ensure the safety of high-speed trains [166].

5) Predictive and proactive maintenance: The maintenance, a broader behavior comparing with the FDD tasks, is an important part to ensure the reliability of high-speed trains. Predictive and proactive maintenance will be witnessed with the popularity in the coming years because of its doubled inspections [167].
6) Distributed FDD strategies: There are connections between individual carriages including motor coaches and trailer coaches, compelling the whole train to operate at the same speed [7]. One high railway rolling stock can be seen as an integrated system consisting of multiple carriages whose information is shared and energy is interactive. Distributed schemes will improve the FDD performance in the case that information is shared with each other and then is utilized.
7) Computer vision-based FDD schemes: Regarded as a typical kind of smart sensors, digital cameras have been widely mounted on intelligent systems, providing system image information that is distinct from traditional measurement data [168]. This complementary information greatly promotes vision-based FDD strategies for high-speed trains [169]. This kind of data-driven FDD strategies, which involves interdisciplinary tasks, is different from the methods involved in this work.

VI. Conclusion

Beginning with an emphasis on data modeling, this paper presents a systematic review, challenges, and perspectives of data-driven FDD methods for traction systems in high-speed trains. Along with these notable FDD research, an explosive growth of data-driven FDD methods would be witnessed to provide exceptional capabilities for performing FDD tasks in the coming years. To this end, several important remarks, drawn from practical and theoretical FDD research for traction systems of high-speed trains, are listed as follows:

- The existence of data with poor quality, such as outliers and missing data, is inevitable in practice. It reduces the accuracy and effectiveness of data models established. Therefore, the first step should check the data quality, and then carry out data cleaning if necessary.
- The FDD problem, as well as failure prediction, cannot be simply regarded as the so-called classification problem, especially for traction systems with dynamic behaviors.
- The essential of static FDD methods lies in discovering and making use of the certain statistical characteristics (such as the correlation) that signals follow, when traction systems work at one fixed operation condition.
- The essential step of dynamic FDD strategies is, by the use of a massive amount of data collected from traction systems of high-speed trains, to mine/identify the relationship between inputs and outputs, and based on which to design residuals independent of various inputs.

Appendix A
Parameters of Traction Systems

TABLE I
Parameters of Traction Systems in CRH2 [11], [170]

No.	Quantity	Symbol	Value (Unit)
1	inductance in stator side	L_s	0.0281 (H)
2	inductance in rotor side	L_r	0.0280 (H)
3	mutual inductance of motor	L_m	0.0268 (H)
4	resistance in rotor side	R_r	0.1267 (Ω)
1	sampling time	T_s	4×10^{-4} (s)
2	pole pairs	p	2 (1)
3	resistance in stator side	Rs	0.228 (Ω)
4	resistance in rotor side	Rr	0.1267 (Ω)
5	mutual inductance of motor	Lm	0.0268 (H)
6	leakage inductance in stator side	Lls	0.0013 (H)
7	leakage inductance in rotor side	Llr	0.0013 (H)
8	inductance in stator side	Ls	0.0281 (H)
9	inductance in rotor side	Lr	0.0280 (H)
10	intermediate voltage	U_d	3200 (V)
11	capacitor of DC-link	C_d	8×10^{-3} (F)
12	filter inductance	L	0.42×10^{-3} (H)
13	filter capacitor	C	6×10^{-3} (F)

Acknowledgment

The authors are sincerely grateful to the Associate Editor and anonymous reviewers for their insightful and constructive comments that helped us improve this work greatly. In addition, they would also like to thank CRRC Zhuzhou Institute Co., Ltd., for the experimental platform of high-speed trains; Prof. Donghua Zhou in Tsinghua University for his contributions to intermittent FDD work in braking systems; Prof. S. Joe Qin in the City University of Hong Kong for his contributions to big data-based FDD work in traction systems; Prof. Chunhua Yang in Central South University for her contributions to fault injections in traction systems; Prof. Baigen Cai in Beijing Jiaotong University for his contributions to FDD work in train control systems; and especially, the authors thank for their significant help with this review.

References

[1] S. Zhang, *Fundamental Application Theory and Engineering Technology for Railway High-speed Trains*. Beijing, China: Science Press, 2007.
[2] P. H. Bowers, "Railway reform in Germany," *J. Transp. Econ. Policy*, vol. 30, no. 1, pp. 95–102, Jan. 1996.
[3] F.-Y. Wang, "Artificial intelligence and intelligent transportation: Driving into the 3rd axial age with ITS," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 4, pp. 6–9, Oct. 2017.
[4] J. Feng, J. Xu, W. Liao, and Y. Liu, "Review on the traction system sensor technology of a rail transit train," *Sensors*, vol. 17, no. 6, pp. 1–16, Jun. 2017.
[5] J. P. Arduin and J. Ni, "French TGV network development," *Jpn. Railway Transp. Rev.*, vol. 40, no. 3, pp. 22–28, Mar. 2005.
[6] L. Chen, X. Hu, W. Tian, H. Wang, D. Cao, and F.-Y. Wang, "Parallel planning: A new motion planning framework for autonomous driving," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 1, pp. 236–246, Jan. 2019.
[7] Z. Chen and K. E. Haynes, *China Railways Era High-Speed*. Bingley, U.K.: Emerald Group, 2015.
[8] S. Gao, Y. Hou, H. Dong, S. Stichel, and B. Ning, "High-speed trains automatic operation with protection constraints: A resilient nonlinear gain-based feedback control approach," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 4, pp. 992–999, Jul. 2019.

[9] B. Ning, T. Tang, Z. Gao, F. Yan, F. Y. Wang, and D. Zeng, "Intelligent railway systems in China," *IEEE Intell. Syst.*, vol. 21, no. 5, pp. 80–83, Sep./Oct. 2006.

[10] S. S. Chopade and P. K. Sharma, "High-speed trains," *Int. J. Modern Eng. Res.*, vol. 3, no. 2, pp. 1161–1166, Mar. 2013.

[11] H. Chen, B. Jiang, N. Lu, and W. Chen, *Data-Driven Detection Diagnosis Faults Traction System High-Speed Trains*. Cham, Switzerland: Springer, 2020.

[12] H. Chen, B. Jiang, S. X. Ding, N. Lu, and W. Chen, "Probability-relevant incipient fault detection and diagnosis methodology with applications to electric drive systems," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 6, pp. 2766–2773, Nov. 2019.

[13] J. Guzinski, H. Abu-Rub, M. Diguet, Z. Krzeminski, and A. Lewicki, "Speed and load torque observer application in high-speed train electric drive," *IEEE Trans. Ind. Electron.*, vol. 57, no. 2, pp. 565–574, Feb. 2010.

[14] J. Wang, J. Wang, C. Roberts, and L. Chen, "Parallel monitoring for the next generation of train control systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 330–338, Feb. 2015.

[15] H. Chen and B. Jiang, "A review of fault detection and diagnosis for the traction system in high-speed trains," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 2, pp. 450–465, Feb. 2020.

[16] Z. Liu, L. Wang, C. Li, and Z. Han, "A high-precision loose strands diagnosis approach for isoelectric line in high-speed railway," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 1067–1077, Mar. 2018.

[17] H. Chen, B. Jiang, N. Lu, and Z. Mao, "Deep PCA based real-time incipient fault detection and diagnosis methodology for electrical drive in high-speed trains," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 4819–4830, Jun. 2018.

[18] N. Qin, W. Jin, J. Huang, Z. Li, and J. Liu, "Feature extraction of high speed train bogie based on ensemble empirical mode decomposition and sample entropy," *J. Southwest Jiaotong Univ.*, vol. 1, pp. 27–32, Jan. 2014.

[19] H. Chen, B. Jiang, W. Chen, and H. Yi, "Data-driven detection and diagnosis of incipient faults in electrical drives of high-speed trains," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4716–4725, Jun. 2019.

[20] W. Bai, X. Yao, H. Dong, and X. Lin, "Mixed H_∞ fault detection filter design for the dynamics of high speed train," *Sci. China Inf. Sci.*, vol. 60, no. 4, pp. 1–3, Apr. 2017.

[21] Y. Song, Z. Liu, A. Rønnquist, P. Nåvik, and Z. Liu, "Contact wire irregularity stochastics and effect on high-speed railway pantograph-catenary tnteractions," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 8196–8206, Oct. 2020.

[22] D. Zhou, H. Ji, X. He, and J. Shang, "Fault detection and isolation of the brake cylinder system for electric multiple units," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 5, pp. 1744–1757, Sep. 2018.

[23] Z. Xu, W. Wang, and Y. Sun, "Performance degradation monitoring for onboard speed sensors of trains," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1287–1297, Sep. 2012.

[24] Y. Wu, B. Jiang, and N. Lu, "A descriptor system approach for estimation of incipient faults with application to high-speed railway traction devices," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 49, no. 10, pp. 2108–2118, Oct. 2019.

[25] C. Li, D. Chen, and L. Yang, "Research on fault detection of high-speed train bogie," in *Proc. 4th Int. Froum Decision Sci.*, Jan. 2017, pp. 253–260.

[26] A. Staino, Rc. Abou-Eid, A. Rojas, B. Fitzgerald, and B. Basu, "Health monitoring and aging assessment of HVAC refrigerant compressors in railway systems," in *Proc. Conf. Intell. Transp. Syst.*, Oct. 2019, pp. 2177–2182.

[27] C. Cheng, M. Li, W. Teng, and J. Chen, "Fault detection method of local outlier factor for high-speed train of running gear based on inter-variable variance," in *Proc. Int. Conf. Ind. Technol.*, Feb. 2019, pp. 1293–1298.

[28] L. Kou, Y. Qin, X. Zhao, and X. Chen, "A multi-dimension End-to-End CNN model for rotating devices fault diagnosis on high-speed train bogie," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2513–2524, Mar. 2020.

[29] Y. Fu, D. Huang, N. Qin, K. Liang, and Y. Yang, "High-speed railway bogie fault diagnosis using LSTM neural network," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 5848–5852.

[30] H. Henao, S. H. Kia, and G.-A. Capolino, "Torsional-vibration assessment and gear-fault diagnosis in railway traction system," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1707–1717, May 2011.

[31] C. Yang *et al.*, "Voltage difference residual-based open-circuit fault diagnosis approach for three-level converters in electric traction systems," *IEEE Trans. Power Electron.*, vol. 35, no. 3, pp. 3012–3028, Mar. 2020.

[32] B. Zhang, *CRH380BL-Type Electric Multiple Units*. Beijing, China: China Railway, 2014.

[33] C. Yang, C. Yang, T. Peng, X. Yang, and W. Gui, "A fault-injection strategy for traction drive control systems," *IEEE Trans. Ind. Electron.*, vol. 64, no. 7, pp. 5719–5727, Jul. 2017.

[34] B. Wang, *Operation Maintenance Electric Multiple Units*. Beijing, China: China Railway Publishing House, 2011.

[35] M. Givoni, "Development and impact of the modern high speed train: A review," *Transp. Rev.*, vol. 26, no. 5, pp. 593–611, Sep. 2006.

[36] H. Chen, B. Jiang, N. Lu, and Z. Mao, "Multi-mode kernel principal component analysis-based incipient fault detection for pulse width modulated inverter of China Railway High-speed 5," *Adv. Mech. Eng.*, vol. 9, no. 10, pp. 1–12, Oct. 2017.

[37] F. Garramiola, J. del Olmo, J. Poza, P. Madina, and G. Almandoz, "Integral sensor fault detection and isolation for railway traction drive," *Sensors*, vol. 18, no. 5, p. 1543, May 2018.

[38] L. Song, *Traction Control Technology for Railway High-speed Trains*. Beijing, China: China Railway Publishing House, 2009.

[39] B. Liu, T. Peng, L. Shi, Z.-Z. He, and C. Yang, "Multi fault diagnosis of traction motor current sensor based on state observer," in *Proc. IEEE Chin. Control Decision Conf.*, Yinchuan, China, May 2016, pp. 7058–7063.

[40] B. Friedland, *Control System Design: An Introduction to State-Space Methods*. North Chelmsford, MA, USA: Courier Corporation, 2005.

[41] H. Miranda, P. Cortes, J. I. Yuz, and J. Rodriguez, "Predictive torque control of induction machines based on state-space models," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 1916–1924, Jun. 2009.

[42] A. B. Youssef, S. K. El Khil, and I. Slama-Belkhodja, "State observer-based sensor fault detection and isolation, and fault tolerant control of a single-phase PWM rectifier for electric railway traction," *IEEE Trans. Power Electron.*, vol. 28, no. 12, pp. 5842–5853, Dec. 2013.

[43] Z. Mao, G. Tao, B. Jiang, and X.-G. Yan, "Adaptive compensation of traction system actuator failures for high-speed trains," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 2950–2963, Nov. 2017.

[44] S. X. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, Tools*, Berlin, Germany: Springer-Verlag, 2008.

[45] S. Yang and M. Wu, "Study on harmonic distribution characteristics and probability model of high speed EMU based on measured data," *J. China Rail Soc.*, vol. 32, no. 3, pp. 33–38, Mar. 2010.

[46] B. Huang and R. Kadali, *Dyn. Model., Predictive Control Perform. Monitoring: A Data-Driven Subspace Approach*, London, U.K.: Springer-Verlag, 2008.

[47] S. X. Ding, *Data-Driven Design Fault Diagnosis Fault-Tolerant Control System*, London, U.K.: Springer-Verlag, 2014.

[48] H. Chen, B. Jiang, W. Chen, and Z. Li, "Edge computing-aided framework of fault detection for traction control systems in high-speed trains," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1309–1318, Feb. 2020.

[49] J. Yin, S. Su, J. Xun, T. Tang, and R. Liu, "Data-driven approaches for modeling train control models: Comparison and case studies," *ISA Trans.*, vol. 98, pp. 349–363, Mar. 2020.

[50] C.-Y. Zhang, D. Chen, J. Yin, and L. Chen, "Data-driven train operation models based on data mining and driving experience for the diesel-electric locomotive," *Adv. Eng. Informat.*, vol. 30, no. 3, pp. 553–563, Aug. 2016.

[51] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, Aug. 1979.

[52] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.

[53] K. Zhang, Y. A. W. Shardt, Z. Chen, S. X. Ding, and K. Peng, "A brief survey of different statistics for detecting multiplicative faults in multivariate statistical process monitoring," in *Proc. 55th Conf. Decison Control, IEEE*, Dec. 2016, pp. 2152–2157.

[54] H. Chen, H. Yi, B. Jiang, K. Zhang, and Z. Chen, "Data-driven detection of hot spots in photovoltaic energy systems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 8, pp. 1731–1738, Aug. 2019.

[55] A. B. Yeh, D. K. J. Lin, and R. N. McGrath, "Multivariate control charts for monitoring covariance matrix: A review," *Qual. Technol. Quant. Manage.*, vol. 3, no. 4, pp. 415–436, Jan. 2006.

[56] H. Cramér, *Math. Methods Statistics*. Princeton, NJ, USA: Princeton Univ. Press, 1999.

[57] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6418–6428, Nov. 2014.

[58] Z. Wu, W. Jin, and N. Qin, "Fault feature analysis of high-speed train suspension system based on multivariate multi-scale sample entropy," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 3913–3918.

[59] G. A. Cherry and S. J. Qin, "Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis," *IEEE Trans. Semicond. Manuf.*, vol. 19, no. 2, pp. 72–159, May 2006.

[60] H. Ji and D. Zhou, "Incipient fault detection of the high-speed train air brake system with a combined index," *Control Eng. Pract.*, vol. 2020, Jul. 2020, Art. no. 104425.

[61] G. Casella and R. L. Berger, *Statiscis Inference*, Pacific Grove, CA, USA: Duxbury, 2002.

[62] H. Ji, K. Huang, and D. Zhou, "Incipient sensor fault isolation based on augmented mahalanobis distance," *Control Eng. Pract.*, vol. 86, pp. 144–154, May 2019.

[63] Z. Chen, S. X. Ding, T. Peng, C. Yang, and W. Gui, "Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1559–1567, Feb. 2018.

[64] H. Chen, B. Jiang, and N. Lu, "A multi-mode incipient sensor fault detection and diagnosis method for electrical traction systems," *Int. J. Control, Autom. Syst.*, vol. 16, no. 4, pp. 1783–1793, Aug. 2018.

[65] F. Garramiola, J. Poza, P. Madina, J. del Olmo, and G. Almandoz, "A review in fault diagnosis and health assessment for railway traction drives," *Appl. Sci.*, vol. 8, no. 12, p. 2475, Dec. 2018.

[66] Q. Liu, Q. Zhu, S. J. Qin, and Q. Xu, "A comparison study of data-driven projection to latent structures modeling and monitoring methods on high-speed train operation," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 6734–6739.

[67] G. Niu, L. Xiong, X. Qin, and M. Pecht, "Fault detection isolation and diagnosis of multi-axle speed sensors for high-speed trains," *Mech. Syst. Signal Process.*, vol. 131, pp. 183–198, Sep. 2019.

[68] T. Guo, D. Zhou, J. Zhang, M. Chen, and X. Tai, "Fault detection based on robust characteristic dimensionality reduction," *Control Eng. Pract.*, vol. 84, pp. 125–138, Mar. 2019.

[69] H. Chen, B. Jiang, N. Lu, and J. Wu, "Incipient fault detection method using Kullback-Leibler divergence for CRH5 and some remarks," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2018, pp. 881–885.

[70] E. Elbouchikhi, Y. Amirat, G. Feld, and M. Benbouzid, "Generalized likelihood ratio test based approach for stator-fault detection in a PWM inverter-fed induction motor drive," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6343–6353, Aug. 2019.

[71] F. Garramiola, J. Poza, P. Madina, J. del Olmo, and G. Ugalde, "A hybrid sensor fault diagnosis for maintenance in railway traction drives," *Sensors*, vol. 20, no. 4, p. 962, Feb. 2020.

[72] K. Pearson, "Principal components analysis," *London, Edinburgh, Dublin Phil. Mag. J. Sci.*, vol. 6, no. 2, p. 559, Jan. 1901.

[73] J. E. Jackson, *A User's Guide to Principal Compononts*. Hoboken, NJ, USA: Wiley, 1991.

[74] N. Vaswani, Y. Chi, and T. Bouwmans, "Rethinking PCA for modern data sets: Theory, algorithms, and applications," *Proc. IEEE*, vol. 106, no. 8, pp. 1274–1276, Aug. 2018.

[75] S. J. Qin, "Survey on data-driven industrial process monitoring and diagnosis," *Annu. Rev. Control*, vol. 36, no. 2, pp. 220–234, Dec. 2012.

[76] J. Wu, B. Jiang, H. Chen, and J. Liu, "Sensors information fusion system with fault detection based on multi-manifold regularization neighborhood preserving embedding," *Sensors*, vol. 19, no. 6, p. 1440, Mar. 2019.

[77] S. X. Ding *et al.*, "On the application of PCA technique to fault diagnosis," *Tsinghua Sci. Technol.*, vol. 15, no. 2, pp. 138–144, Apr. 2010.

[78] J. F. Martins, V. F. Pires, and A. J. Pires, "PCA-based on-line diagnosis of induction motor stator fault feed by PWM inverter," in *Proc. IEEE ISIE*, Jul. 2006, pp. 2401–2405.

[79] H. Chen, B. Jiang, N. Lu, and Z. Mao, "Data-based incipient actuator fault detection and diagnosis for three-phase PWM voltage source inverter," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 6443–6448.

[80] C. Dai, K. Huang, K. Hu, and Z. Liu, "Fault diagnosis approach of traction transformers in high-speed railway combining kernel principal component analysis with random forest," *IET Electr. Syst. Transp.*, vol. 6, no. 3, pp. 202–206, Sep. 2016.

[81] Z. Li, F. Chen, L. Wang, and B. Jiang, "Composite fault detection and diagnosis for IGBT and current sensor in CRH2 through modified EEMD and KPCA methods," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 5144–5149.

[82] H. Chen, B. Jiang, and N. Lu, "An improved incipient fault detection method based on Kullback-Leibler divergence," *ISA Trans.*, vol. 79, pp. 127–136, Aug. 2018.

[83] H. Chen, B. Jiang, and N. Lu, "Data-driven incipient sensor fault estimation with application in inverter of high-speed railway," *Math. Prob. Eng.*, vol. 2017, p. 8937356, Sep. 2017.

[84] A. Giantomassi, F. Ferracuti, S. Iarlori, G. Ippoliti, and S. Longhi, "Electric motor fault detection and diagnosis by kernel density estimation and Kullback–Leibler divergence based on stator current measurements," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1770–1780, Mar. 2015.

[85] G. Li and S. J. Qin, "Comparative study on monitoring schemes for non-Gaussian distributed processes," *J. Process Control*, vol. 67, pp. 69–82, Jul. 2018.

[86] H. Chen, B. Jiang, N. Lu, and W. Chen, "Real-time incipient fault detection for electrical traction systems of CRH2," *Neurocomputing*, vol. 306, pp. 119–129, Sep. 2018.

[87] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Hoboken, NJ, USA: Wiley, 2004.

[88] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process," *J. Process Control*, vol. 22, no. 9, pp. 1567–1581, Oct. 2012.

[89] J.-M. Lee, C. Yoo, and I.-B. Lee, "Statistical process monitoring with independent component analysis," *J. Process Control*, vol. 14, no. 5, pp. 467–485, Aug. 2004.

[90] Y. Wan, M. Tu, W. Yan, Y. Zhao, and E. Wang, "Fault diagnosis of high speed iron traction transformer for V/V type wiring based on fast independent component analysis," *Global J. Res. Eng.*, vol. 17, no. 8, pp. 1–5, Jan. 2018.

[91] M. Tu *et al.*, "Real-time diagnosis of high-speed rail traction transformer in different topologies," *J. Eng.*, vol. 2019, no. 16, pp. 2084–2088, Mar. 2019.

[92] J. Weidong, "Fault diagnosis of gearbox by FastICA and residual mutual information based feature extraction," in *Proc. Int. Conf. Inf. Autom.*, Jun. 2009, pp. 928–932.

[93] X. Xie, B. Jiang, J. Liu, and H. Chen, "Fault detection and isolation based on MM- ICA with application to high-speed railway," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2018, pp. 111–115.

[94] S. Kharbech, I. Dayoub, M. Zwingelstein-Colin, and E. P. Simon, "Blind digital modulation identification for MIMO systems in railway environments with high-speed channels and impulsive noise," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7370–7379, Aug. 2018.

[95] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, nos. 3–4, pp. 321–377, Dec. 1936.

[96] W. E. Larimore, "Canonical variate analysis in identification, filtering, and adaptive control," in *Proc. 29th IEEE Conf. Decis. Control*, Dec. 1990, pp. 596–604.

[97] H. Chen, B. Jiang, T. Zhang, and N. Lu, "Data-driven and deep learning-based detection and diagnosis of incipient faults with application to electrical traction systems," *Neurocomputing*, vol. 396, pp. 429–437, Jul. 2020.

[98] H. Chen, B. Jiang, H. Yi, and N. Lu, "Data-driven fault diagnosis for dynamic traction systems in high-speed trains," *SCIENTIA SINICA Informationis*, vol. 50, no. 4, pp. 496–510, Apr. 2020.

[99] Y. Jiang and S. Yin, "Recent advances in Key-Performance-Indicator oriented prognosis and diagnosis with a MATLAB toolbox: DB-KIT," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2849–2858, May 2019.

[100] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor, "Canonical correlation analysis: An overview with application to learning methods," *Neural Comput.*, vol. 16, no. 12, pp. 2639–2664, Dec. 2004.

[101] M. Borga, *Learning Multidimensional Signal Processing*. London, U.K.: Linköping University Electronic Press, 1998.

[102] K. Zhang, K. Peng, S. X. Ding, Z. Chen, and X. Yang, "A correlation-based distributed fault detection method and its application to a hot tandem rolling mill process," *IEEE Trans. Ind. Electron.*, vol. 67, no. 3, pp. 2380–2390, Mar. 2020.

[103] H. Chen, B. Jiang, and S. X. Ding, "A broad learning aided data-driven framework of fast fault diagnosis for high-speed trains," *IEEE Intell. Transp. Syst. Mag.*, early access, Jan. 7, 2020, doi: 10.1109/MITS.2019.2907629.

[104] Z. Chen *et al.*, "A data-driven ground fault detection and isolation method for main circuit in railway electrical traction system," *ISA Trans.*, vol. 87, pp. 264–271, Apr. 2019.

[105] L. Gasparetto, S. Alfi, and S. Bruni, "Data-driven condition-based monitoring of high-speed railway bogies," *Int. J. Rail Transp.*, vol. 1, nos. 1–2, pp. 42–56, Feb. 2013.

[106] N. Qin, Y. Sun, P. Gu, and L. Ma, "Bogie fault identification based on EEMD information entropy and manifold learning," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 315–318, Jul. 2017.

[107] G. Li, W. Jin, P. Jiang, X. Fu, D. Xiong, and P. Gu, "Research on fault diagnosis technology of high-speed rail for large scale monitoring data," *J. Syst. Simul.*, vol. 26, pp. 2458–2464, Oct. 2014.

[108] H. Chen, J. Wu, B. Jiang, and W. Chen, "A modified neighborhood preserving embedding-based incipient fault detection with applications to small-scale cyber–physical systems," *ISA Trans.*, vol. 104, pp. 175–183, Sep. 2020.

[109] Y. Ma, Z. Wang, H. Yang, and L. Yang, "Artificial intelligence applications in the development of autonomous vehicles: A survey," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 2, pp. 315–329, Mar. 2020.

[110] E. Principi, D. Rossetti, S. Squartini, and F. Piazza, "Unsupervised electric motor fault detection by using deep autoencoders," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 2, pp. 441–451, Mar. 2019.

[111] K. Zhang, S. X. Ding, Y. A. W. Shardt, Z. Chen, and K. Peng, "Assessment of T^2- and Q-statistics for detecting additive and multiplicative faults in multivariate statistical process monitoring," *J. Franklin Inst.*, vol. 354, no. 2, pp. 668–688, Jan. 2017.

[112] D. Tang, W. Jin, N. Qin, and H. Li, "Feature selection and analysis of single lateral damper fault based on SVM-RFE with correlation bias reduction," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 3840–3845.

[113] S. Vallachira, M. Orkisz, M. Norrlof, and S. Butail, "Data-driven gearbox failure detection in industrial robots," *IEEE Trans. Ind. Informat.*, vol. 16, no. 1, pp. 193–201, Jan. 2020.

[114] G. Ding, L. Wang, P. Shen, and P. Yang, "Sensor fault diagnosis based on ensemble empirical mode decomposition and optimized least squares support vector machine," *J. Comput.*, vol. 8, no. 11, pp. 2916–2924, Nov. 2013.

[115] M. Fei, L. Ning, M. Huiyu, P. Yi, S. Haoyuan, and Z. Jianyong, "Online fault diagnosis model for locomotive traction inverter based on wavelet transform and support vector machine," *Microelectron. Rel.*, vols. 88–90, pp. 1274–1280, Sep. 2018.

[116] J. Liu, Y.-F. Li, and E. Zio, "A SVM framework for fault detection of the braking system in a high speed train," *Mech. Syst. Signal Process.*, vol. 87, pp. 401–409, Mar. 2017.

[117] G. E. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. Hoboken, NJ, USA: Wiley, 2011.

[118] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: The role of machine learning," *IEEE Access*, vol. 5, pp. 20590–20616, 2017.

[119] H. Chen, B. Jiang, and N. Lu, "A newly robust fault detection and diagnosis method for high-speed trains," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2198–2208, Jun. 2019.

[120] 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. 5590–5600, Jul. 2017.

[121] H. Wang, A. Nunez, Z. Liu, D. Zhang, and R. Dollevoet, "A Bayesian network approach for condition monitoring of high-speed railway catenaries," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4037–4051, Oct. 2020.

[122] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[123] C. L. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.

[124] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.

[125] Y. Wu and W. Jin, "CNN-based fault diagnosis of high-speed train with imbalance data: A comparison study," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 5053–5058.

[126] J. Yin and W. Zhao, "Fault diagnosis network design for vehicle onboard equipments of high-speed railway: A deep learning approach," *Eng. Appl. Artif. Intell.*, vol. 56, pp. 250–259, Nov. 2016.

[127] W. Zhang, X. Li, and Q. Ding, "Deep residual learning-based fault diagnosis method for rotating machinery," *ISA Trans.*, vol. 95, pp. 295–305, Dec. 2019.

[128] L. Yang, K. Li, D. Zhao, S. Gu, and D. Yan, "A network method for identifying the root cause of high-speed rail faults based on text data," *Energies*, vol. 12, no. 10, p. 1908, May 2019.

[129] J. Zhang, J. Zhao, D. Zhou, and C. Huang, "High-performance fault diagnosis in PWM voltage-source inverters for vector-controlled induction motor drives," *IEEE Trans. Power Electron.*, vol. 29, no. 11, pp. 6087–6099, Nov. 2014.

[130] Z. Chen *et al.*, "A Just-in-time-learning aided canonical correlation analysis method for multimode process monitoring and fault detection," *IEEE Trans. Ind. Electron.*, early access, Apr. 28, 2020, doi: 10.1109/TIE.2020.2989708.

[131] A. Astolfi and L. Menini, "Input/output decoupling problems for high speed trains," in *Proc. Amer. Control Conf.*, 2002, pp. 549–554.

[132] J. W. Finch and D. Giaouris, "Controlled AC electrical drives," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 481–491, Feb. 2008.

[133] H. Dong, B. Ning, B. Cai, and Z. Hou, "Automatic train control system development and simulation for high-speed railways," *IEEE Circuits Syst. Mag.*, vol. 10, no. 2, pp. 6–18, May 2010.

[134] V. V. Krylov, *Noise Vibratation from High-Speed Trains*, London, U.K.: Thomas Telford, 2001.

[135] D. Zhou, Y. Zhao, Z. Wang, X. He, and M. Gao, "Review on diagnosis techniques for intermittent faults in dynamic systems," *IEEE Trans. Ind. Electron.*, vol. 67, no. 3, pp. 2337–2347, Mar. 2020.

[136] S. J. Qin, "Process data analytics in the era of big data," *AIChE J.*, vol. 60, no. 9, pp. 3092–3100, Sep. 2014.

[137] Q. Xu *et al.*, "A platform for fault diagnosis of high-speed train based on big data," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 309–314, 2018.

[138] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 383–398, Jan. 2019.

[139] S. Haykin, *Neural Network Learning Machine*, London, U.K.: Pearson, 2010.

[140] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, "Intelligent fault diagnosis of the high-speed train with big data based on deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2106–2116, Aug. 2017.

[141] Q. Liu, D. Kong, S. J. Qin, and Q. Xu, "Map-reduce decentralized PCA for big data monitoring and diagnosis of faults in high-speed train bearings," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 144–149, 2018.

[142] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1624–1639, Dec. 2011.

[143] R. Boutaba *et al.*, "A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities," *J. Internet Services Appl.*, vol. 9, no. 1, pp. 1–99, Jun. 2018.

[144] S. Jialin Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[145] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, p. 9, May 2016.

[146] K. Zhou and J. C. Doyle, *Essentials Robust Control*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[147] Y. Zhang, J. An, and C. Ma, "Fault detection of non-Gaussian processes based on model migration," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 5, pp. 1517–1526, Sep. 2013.

[148] Z. Liu *et al.*, "Industrial AI enabled prognostics for high-speed railway systems," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2018, pp. 1–8.

[149] Z. Yang, V. Cheung, C. Gao, and Q. Zhang, "Train intelligent detection system based on convolutional neural network," in *Adv. Hum. Factors Simulat.*, Jun. 2019, pp. 157–165.

[150] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[151] C. Cheng, J. Wang, Z. Zhou, W. Teng, Z. Sun, and B. Zhang, "A BRB-based effective fault diagnosis model for high-speed trains running gear systems," *IEEE Trans. Intell. Transp. Syst.*, early access, Jul. 22, 2020, doi: 10.1109/TITS.2020.3008266.

[152] S. J. Qin, "An overview of subspace identification," *Comput. Chem. Eng.*, vol. 30, nos. 10–12, pp. 1502–1513, Sep. 2006.

[153] S. X. Ding, "Data-driven design of monitoring and diagnosis systems for dynamic processes: A review of subspace technique based schemes and some recent results," *J. Process Control*, vol. 24, no. 2, pp. 431–449, 2014.

[154] W. Du, Y. Zhang, and W. Zhou, "Modified non-Gaussian multivariate statistical process monitoring based on the Gaussian distribution transformation," *J. Process Control*, vol. 85, pp. 1–14, Jan. 2020.

[155] Q. Song and Y.-D. Song, "Data-based fault-tolerant control of high-speed trains with traction/braking notch nonlinearities and actuator failures," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2250–2261, Dec. 2011.

[156] H. Luo, X. Yang, M. Krueger, S. X. Ding, and K. Peng, “A plug-and-play monitoring and control architecture for disturbance compensation in rolling mills,” *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 200–210, Feb. 2018.

[157] Y. Chen, Y. Lv, and F.-Y. Wang, “Traffic flow imputation using parallel data and generative adversarial networks,” *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1624–1630, Apr. 2020.

[158] Y. Cao, P. Li, and Y. Zhang, “Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing,” *Future Gener. Comput. Syst.*, vol. 88, pp. 279–283, Nov. 2018.

[159] V. J. Hodge, S. O’Keefe, M. Weeks, and A. Moulds, “Wireless sensor networks for condition monitoring in the railway industry: A survey,” *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1088–1106, Jun. 2015.

[160] C. Cheng *et al.*, “A semi-quantitative information based fault diagnosis method for the running gears system of high-speed trains,” *IEEE Access*, vol. 7, pp. 38168–38178, 2019.

[161] Y. Lv, Y. Chen, L. Li, and F.-Y. Wang, “Generative adversarial networks for parallel transportation systems,” *IEEE Intell. Transp. Syst. Mag.*, vol. 10, no. 3, pp. 4–10, Jun. 2018.

[162] C. Cheng *et al.*, “Health status prediction based on belief rule base for high-speed train running gear system,” *IEEE Access*, vol. 7, pp. 4145–4159, 2019.

[163] O. Chapelle, B. Schölkopf, and A. Zien, *Semi-Supervised Learning*. Cambridge, MA, USA: MIT Press, 2006.

[164] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[165] K. Hu *et al.*, “Cost-effective prognostics of IGBT bond wires with consideration of temperature swing,” *IEEE Trans. Power Electron.*, vol. 35, no. 7, pp. 6773–6784, Jul. 2020.

[166] D. Feng, S. Lin, Q. Yang, X. Lin, Z. He, and W. Li, “Reliability evaluation for traction power supply system of high-speed railway considering relay protection,” *IEEE Trans. Transport. Electrific.*, vol. 5, no. 1, pp. 285–298, Mar. 2019.

[167] Q. Wang, S. Bu, and Z. He, “Achieving predictive and proactive maintenancefor high-speed railway power equipment with LSTM-RNN,” *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6509–6517, Oct. 2020.

[168] M. Karakose, O. Yaman, M. Baygin, K. Murat, and E. Akin, “A new computer vision based method for rail track detection and fault diagnosis in railways,” *Int. J. Mech. Eng. Robot. Res.*, vol. 6, no. 1, pp. 22–27, Jan. 2017.

[169] Z. Liu, K. Liu, J. Zhong, Z. Han, and W. Zhang, “A high-precision positioning approach for catenary support components with multi-scale difference,” *IEEE Trans. Instrum. Meas.*, vol. 69, no. 3, pp. 700–711, Mar. 2020.

[170] X. Yang, C. Yang, T. Peng, Z. Chen, B. Liu, and W. Gui, “Hardware-in-the-Loop fault injection for traction control system,” *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 6, no. 2, pp. 696–706, Jun. 2018.

Hongtian Chen (Member, IEEE) received the B.S. and M.S. degrees from the School of Electrical and Automation Engineering, Nanjing Normal University, China, in 2012 and 2015, respectively, and the Ph.D. degree from the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, China, in 2019.

He had been a Visiting Scholar at the Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Germany, in 2018. He is currently a Post-Doctoral Fellow with the Department of Chemical and Materials Engineering, University of Alberta, Canada. His research interests include process monitoring and fault diagnosis, data mining and analytics, machine learning, and quantum computation; and their applications in high-speed trains, new energy systems, and industrial processes. He was a recipient of the Grand Prize of Innovation Award of Ministry of Industry and Information Technology of the People’s Republic of China in 2019, and the Excellent Ph.D. Thesis Award of Jiangsu Province in 2020.

Bin Jiang (Fellow, IEEE) received the Ph.D. degree in automatic control from Northeastern University, Shenyang, China, in 1995.

He had been a Post-Doctoral Fellow, a Research Fellow, an Invited Professor, and a Visiting Professor in Singapore, France, USA, and Canada, respectively. He is currently a Chair Professor of Cheung Kong Scholar Program with the Ministry of Education and the Vice President of the Nanjing University of Aeronautics and Astronautics, Nanjing, China. He has authored eight books and over 200 referred international journal articles and conference papers. His current research interests include intelligent fault diagnosis and fault-tolerant control and their applications to helicopters, satellites, and high-speed trains.

Dr. Jiang is a fellow of the Chinese Association of Automation (CAA). He was a recipient of the Second Class Prize of National Natural Science Award of China in 2018. He is the Chair of the Control Systems Chapter in IEEE Nanjing Section, and a member of the IFAC Technical Committee on Fault Detection, Supervision, and Safety of Technical Processes. He currently serves as an Associate Editor or an Editorial Board Member for a number of journals, such as the IEEE TRANSACTIONS ON CYBERNETICS, *Journal of the Franklin Institute*, and *Neurocomputing*.

Steven X. Ding received the Ph.D. degree in electrical engineering from the Gerhard-Mercator University of Duisburg, Duisburg, Germany, in 1992.

From 1992 to 1994, he was an Research and Development Engineer with Rheinmetall GmbH. From 1995 to 2001, he was a Professor of Control Engineering with the University of Applied Science Lausitz, Senftenberg, Germany, where he served as the Vice-President from 1998 to 2000. Since 2001, he has been a Professor of Control Engineering and the Head of the Institute for Automatic Control and Complex Systems (AKS), University of Duisburg-Essen, Duisburg. His research interests are model-based and data-driven fault diagnosis, fault-tolerant systems, and their applications in the industry with a focus on automotive systems and chemical processes.

Biao Huang (Fellow, IEEE) received the B.Sc. and M.Sc. degrees in automatic control from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1983 and 1986, respectively, and the Ph.D. degree in process control from the University of Alberta, Edmonton, AB, Canada, in 1997.

In 1997, he joined the University of Alberta as an Assistant Professor with the Department of Chemical and Materials Engineering. He is currently a Professor with the Natural Sciences and Engineering Research Council Industrial Research Chair in Control of Oil Sands Processes and was also the Alberta Innovates Technology Futures Industry Chair in Process Control. He has applied his expertise extensively in industrial practice, particularly in the oil sands industry. His research interests include process control, system identification, control performance assessment, Bayesian methods, and state estimation.

Dr. Huang is a fellow of the Canadian Academy of Engineering and the Chemical Institute of Canada. He was a recipient of Germany’s Alexander von Humboldt Research Fellowship, the Canadian Chemical Engineer Society’s Syncrude Canada Innovation and D. G. Fisher Awards, APEGA Summit Research Excellence Award, University of Alberta McCalla and Killam Professorship Awards, Petro-Canada Young Innovator Award, and a Best Paper Award from the Journal of Process Control.