

Article

Data-Driven Robust Kalman Filter-Based Fault Detection for Traction Drive Systems

Caixin Fu ¹, Changhong Jiang ^{2,*}, Zhiwei Wan ^{2,*} , Peng Cheng ³ and Shenquan Wang ²

¹ School of Mechanical and Electrical Engineering, Changchun University of Technology, Changchun 130012, China; 1202301017@stu.ccut.edu.cn

² College of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130012, China; jiangchanghong@ccut.edu.cn (C.J.); shenquanwang@ccut.edu.cn (S.W.)

³ CRRC Changchun Railway Vehicles Co., Ltd., Changchun 130011, China; chengpeng.ck@crrcgc.com

* Correspondence: 1202503002@stu.ccut.edu.cn

Abstract

This article addresses the fault detection (FD) problem for traction drive systems in the presence of unknown noise covariances. The dynamic behavior of the traction drive system, affected by actuator and sensor faults, is first formulated. Following the philosophy of the subspace identification, the system matrices are identified directly from collected process data using QR decomposition and singular value decomposition. Based on the identified model, a robust Kalman filter (KF)-based FD scheme is developed. By exploiting the iterative interaction between the estimator and measurement data within the KF framework, the noise covariance matrices are adaptively estimated, which alleviates the adverse effects caused by empirical covariance selection in conventional KF-based FD methods. Experimental results obtained from a real traction drive system verify the effectiveness and reliability of the proposed approach.

Keywords: data-driven; robust KF; subspace identification; fault detection; traction drive systems

1. Introduction

Ensuring the safe and reliable operation of industrial systems has become increasingly important with the rapid development of automation and intelligent technologies. Traction drive systems serve a key role by transforming electrical energy into mechanical motion in industrial applications. The reliability of the systems is therefore of great importance, which has motivated extensive research on fault detection (FD) techniques for traction drive systems [1–3]. Advances in modern control theory have contributed to the design and implementation of model-based FD approaches in traction drive systems [4–6].

Model-based FD methods are attractive because of their high sensitivity and capability for early detection of system anomalies [7]. Depending on the structure of the residual generator, model-based FD methods for traction drive systems are typically classified into fault detection filter-based approaches and diagnostic observer-based approaches [8]. In parallel with these developments, recent studies, such as those by Cheng et al. [9], Sun et al. [10], and Xia et al. [11], have explored data-driven and model–data fusion approaches for FD in traction drive systems. Despite the diversity of these methodologies, accurate characterization of system dynamics and fault-induced residual variations remains crucial for reliable monitoring. In this context, fault detection filters are particularly appealing, as they explicitly account for state evolution and measurement updates in the residual generation. Among them, Kalman filter (KF)-based



Academic Editor: Ignacio González-Prieto

Received: 19 March 2026

Revised: 10 April 2026

Accepted: 19 April 2026

Published: 22 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

approaches are especially attractive due to the recursive estimation structure, which is well suited for dynamic systems operating under stochastic noises.

In recent years, KF-based FD methods and their variants have emerged as the dominant approach among fault detection filter-based techniques for traction drive systems [12–14]. Cheng et al. [12] developed a sigma-mixed unscented KF-based FD algorithm for addressing performance degradation. The approach constructs a mixture distribution using sigma points in the unscented KF framework and incorporates a Lévy process with jump characteristics to model degradation dynamics. A moving average interstate standard deviation index is finally developed for FD purposes. Foo et al. [13] introduced an extended KF-based approach for sensor FD and isolation, which ensures fault-resilient control when sensor faults occur. For stator inter-turn faults, Namdar et al. [14] proposed a KF-based FD method. In this approach, the KF is employed to extract signal features, and the standard deviation of the extracted signatures is subsequently used as an indicator for FD. Miniach et al. [15] developed a current sensor fault-tolerant control scheme based on an extended KF to achieve detection and compensation of current sensor faults in induction motor drives.

It should be noted that the practical deployment of aforementioned methods still faces several challenges. First, most of these methods rely on prior knowledge of system matrices. However, due to the complex internal structure of traction drive systems, constructing an accurate mathematical model is often difficult in practice, and the required system matrices are typically unavailable. As a result, the applicability of KF-based FD methods in real monitoring scenarios is limited. Second, the process and measurement noise covariance matrices in most existing KF and its variant-based FD approaches are usually specified based on prior knowledge or engineering experience. In practical operating environments, however, the actual noise statistics are generally unknown, which may lead to a mismatch between the assumed and actual noise characteristics during the iterative estimation process, thereby degrading the FD performance. Therefore, a key challenge lies not only in implementing KF-based FD for traction drive systems, but also in ensuring its effectiveness when both the system matrices and noise statistics are not accurately known a priori.

Although traction drive systems can be modeled in a state-space model form [8], the system matrices are often unavailable in practice. To address this issue, a subspace identification approach [16,17] is employed to construct the state-space model directly from measured input–output (I/O) data. The observability matrix is estimated via QR decomposition and singular value decomposition (SVD), from which the system matrices are identified. This enables KF and variant-based FD approaches to be implemented in a fully data-driven manner while retaining its capability for dynamic modeling and recursive estimation, without relying on prior model knowledge. Furthermore, to address the case where the process and measurement noise covariance matrices are unknown in traction drive systems, the proposed approach is inspired by the principle of iterative generalized least squares estimation [18]. The noise statistics are learned through iterative interactions between the estimator and the measurements, enabling reliable estimation of the noise covariance matrices. This mitigates the adverse impact on detection performance caused by relying on the priori noise covariance matrices that deviate from the true noise characteristics. The main contributions of this work are outlined as follows.

1. A data-driven KF-based approach is developed for FD of traction drive systems, which does not require a first-principles system model.
2. An iterative scheme is proposed to estimate the covariance matrices of noises from measured and estimated data, mitigating the adverse effects caused by mismatches between a priori covariance assumptions and the true noise statistics.
3. The proposed method accounts for the dynamic characteristics of the systems while maintaining satisfactory monitoring results.

The remainder of this study is organized as follows: Section 2 presents preliminaries, including a brief introduction to traction drive systems, the KF, and the problem formulation. A detailed description of the proposed FD approach is provided in Section 3. Section 4 presents the corresponding experimental study on a traction drive system. The study concludes with a summary in Section 5.

Notations: All notations used in this paper follow standard conventions. R^k represents the k -dimensional Euclidean space. The symbol $[\cdot]^-$ denotes the pseudo-inverse of the matrix $[\cdot]$. For a vector ω , $\omega_s(k) = [\omega^T(k-s) \ \omega^T(k-s+1) \cdots \omega^T(k)]^T$, $\Omega_k = [\omega(k-s) \ \omega(k-s+1) \cdots \omega(k-s+N-1)]$, $\Omega_{k,s} = [\omega_s(k) \ \omega_s(k+1) \cdots \omega_s(k+N-1)]$, where N and s are positive integers. The symbol $\hat{\xi}$ denotes the estimate of ξ . The operator $vec([\cdot])$ denotes the vectorization of the matrix $[\cdot]$. $\otimes$ denotes the Kronecker product.

2. Preliminaries and Problem Formulation

This section first presents the traction drive system along with its general mathematical model, followed by an introduction to the KF. Based on this foundation, the FD problem addressed in this work is formulated.

2.1. Traction Drive Systems

Traction drive systems serve as fundamental devices for converting electrical energy into mechanical motion, delivering the required traction to the equipment in a controlled manner. The system includes a DC-link, a traction inverter, a traction motor, and a traction control unit, as illustrated in Figure 1.

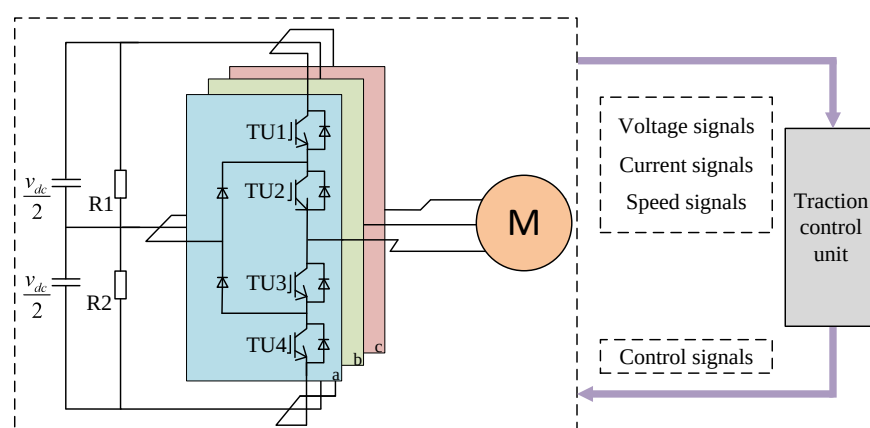


Figure 1. Schematic diagram of the traction drive system.

A variety of sensors, such as those measuring speed, current, and voltage, are incorporated to provide data for control and monitoring purposes. A total of seven sensors are used to collect operational data, as summarized in Table 1. Figure 1 further emphasizes the pivotal role of the traction control unit in maintaining safe and stable system operation.

Table 1. Sensors used in the traction drive system.

Variable	Description	Unit
I_a	A-phase current	A
I_b	B-phase current	A
V_a	A-phase voltage	V
V_b	B-phase voltage	V
V_c	C-phase voltage	V
s	Motor speed	r/min
V_{dc}	System input voltage	V

Effective FD requires first representing the traction drive system in a state-space model. Around a fixed operating point, the system can be described by a generalized linear time-invariant model that incorporates fault-related effects [1]:

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

where $f(k) \in R^{k_f}$, $u(k) \in R^{k_u}$, $x(k) \in R^{k_x}$, and $y(k) \in R^{k_y}$ represent the actuator or sensor faults, system inputs, state variables, and process outputs, respectively. Fault matrices E_f and F_f have compatible dimensions. The system matrices A , B , C , and D are structured to satisfy the dimensional requirements of the state-space representation. $v(k) \in R^{k_y}$ and $w(k) \in R^{k_x}$ represent measurement and process noises, respectively, which are assumed to be statistically independent of $u(k)$ and the initial state $x(0)$ and satisfy

$$\mathbb{E} \left(\begin{bmatrix} w(i) \\ v(i) \end{bmatrix} \begin{bmatrix} w(j) \\ v(j) \end{bmatrix}^T \right) = \begin{bmatrix} \Sigma_w \delta_{ij} & 0 \\ 0 & \Sigma_v \delta_{ij} \end{bmatrix}, \quad \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}.$$

For designing data-driven FD schemes, an I/O model reflecting the relation between system inputs and outputs is essential. By iteratively expanding (1), the associated I/O data model can be systematically represented as [19]:

$$Y_{k,s} = \Gamma_s X_k + H_{u,s} U_{k,s} + H_{f,s} F_{k,s} + H_{w,s} W_{k,s} + V_{k,s} \quad (2)$$

with $Y_{k,s} \in R^{(s+1)k_y \times N}$ and

$$\begin{aligned}\Gamma_s &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^s \end{bmatrix}, \quad H_{u,s} = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{s-1}B & \cdots & CB & D \end{bmatrix} \\ H_{w,s} &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ C & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{s-1} & \cdots & C & 0 \end{bmatrix}, \quad H_{f,s} = \begin{bmatrix} F_f & 0 & \cdots & 0 \\ CE_f & F_f & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ CA^{s-1}E_f & \cdots & CE_f & F_f \end{bmatrix}\end{aligned}$$

in which $H_{w,s} \in R^{(s+1)k_y \times (s+1)k_x}$, $H_{f,s} \in R^{(s+1)k_y \times (s+1)k_f}$, $\Gamma_s \in R^{(s+1)k_y \times k_x}$, and $H_{u,s} \in R^{(s+1)k_y \times (s+1)k_u}$.

2.2. Kalman Filter

For the fault-free case of system (1), the KF is adopted in an innovation-based predictor form. The innovation is defined as

$$r(k) = y(k) - C\hat{x}(k) - Du(k) = y(k) - \hat{y}(k), \quad (3)$$

and its covariance matrix is denoted by

$$\Sigma_{r,k} = \mathbb{E}[r(k)r^T(k)] = \Sigma_v + CP_k C^T, \quad (4)$$

where $P_k = \mathbb{E}[(x(k) - \hat{x}(k))(x(k) - \hat{x}(k))^T]$ denotes the state estimation error covariance matrix, and Σ_v is the measurement noise covariance matrix. Then, KF gain is calculated by

$$K_k = AP_k C^T \Sigma_{r,k}^{-1}, \quad (5)$$

and the state estimation error covariance and state estimate are recursively updated as

$$\begin{aligned} P_{k+1} &= AP_k A^T + \Sigma_w - K_k \Sigma_{r,k} K_k^T, \\ \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + K_k r(k), \end{aligned} \quad (6)$$

where Σ_w denotes the process noise covariance matrix. The innovation sequence $r(k)$ represents the deviation between the actual measurement and the prediction, which can be employed as a residual signal for FD.

2.3. Problem Formulation

In practical operation of traction drive systems, the exact covariance matrices of the noises are often unknown, and the system parameters in the state-space representation (1) are difficult to accurately determine [8]. Consequently, conventional KF-based FD methods become less applicable. This motivates the development of a robust data-driven scheme that mitigates the impact of mismatches between assumed and true noise covariance matrices during iterative estimation, while avoiding complex system modeling. To achieve FD, the residual vector can be constructed based on (3) as

$$r(k) = y(k) - \hat{y}(k). \quad (7)$$

According to (3)–(6), the residual generation process requires not only the process I/O data but also the knowledge of system matrices A , B , C , and D , as well as the noise covariance matrices Σ_w and Σ_v . However, due to the complexity of system modeling and the lack of accurate noise statistical information in practice, these quantities are often unavailable. Consequently, residual generation relying on parameters specified from prior knowledge may lead to unsatisfactory detection performance.

To achieve reliable FD, the proposed framework aims to identify system matrices and estimate noise covariance matrices directly from process data. Therefore, the objective of this work is to design a data-driven method capable of learning system parameters and noise statistics from I/O measurements.

3. Proposed FD Method

Since the state vector $x(k)$ is typically unavailable in practice, the I/O data model (2) cannot be identified and employed directly for generating residual signals. Consequently, the reliance on the state variable $x(k)$ needs to be eliminated. Considering the innovation sequence $e(k) = y(k) - \hat{y}(k)$ and the gain matrix K , the I/O relationship of the process can alternatively be described by [20]

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + Ke(k) = A_K \hat{x}(k) + B_K u(k) + K y(k) \\ \hat{y} &= C\hat{x}(k) + Du(k), \quad B_K = B - KD, \quad A_K = A - KC. \end{aligned} \quad (8)$$

According to (8), the following can be obtained

$$\hat{x}(k) = A_K^{s_p} \hat{x}(k - s_p) + \sum_{i=1}^{s_p} A_K^{i-1} [B_K \ K] \begin{bmatrix} u(k-i) \\ y(k-i) \end{bmatrix}. \quad (9)$$

Given that all eigenvalues of A_K lie strictly within the unit circle, it follows that $A_K^{s_p} \rightarrow 0$ for a sufficiently large positive integer s_p . This leads to

$$\hat{x}(k) \approx \sum_{i=1}^{s_p} A_K^{i-1} [B_K \ K] \begin{bmatrix} u(k-i) \\ y(k-i) \end{bmatrix}. \quad (10)$$

This implies that the state vector $x(k)$ can be inferred from historical I/O data. Therefore, in the absence of faults, the model (2) can be equivalently reformulated as

$$Y_{k,s} \approx \Gamma_s L_p Z_p + H_{u,s} U_{k,s} + H_{w,s} W_{k,s} + V_{k,s} \quad (11)$$

where

$$L_p = [A_K^{s_p-1} B_K \cdots B_K A_K^{s_p-1} K \cdots K],$$

$$Z_p = \begin{bmatrix} U_{k-s_p, s_p-1} \\ Y_{k-s_p, s_p-1} \end{bmatrix}.$$

Applying a QR decomposition of the form

$$\begin{bmatrix} Z_p \\ U_{k,s} \\ Y_{k,s} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix}. \quad (12)$$

Then, one can obtain

$$\begin{bmatrix} Z_p \\ U_{k,s} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, \quad (13)$$

$$Y_{k,s} = [R_{31} \ R_{32}] \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} + R_{33} Q_3. \quad (14)$$

Specially, the I/O data model (11) can be further expressed as

$$Y_{k,s} = [\Gamma_s L_p \ H_{u,s}] \begin{bmatrix} Z_p \\ U_{k,s} \end{bmatrix} + H_{w,s} W_{k,s} + V_{k,s}. \quad (15)$$

From (13)–(15), the following equivalent form can be obtained

$$\begin{aligned} [\Gamma_s L_p \ H_{u,s}] \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} + H_{w,s} W_{k,s} + V_{k,s} &= [R_{31} \ R_{32}] \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} + R_{33} Q_3 \\ \Rightarrow [\Gamma_s L_p \ H_{u,s}] \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} &\triangleq [R_{31} \ R_{32}] \\ H_{w,s} W_{k,s} + V_{k,s} &\triangleq R_{33} Q_3 \end{aligned} \quad (16)$$

In line with subspace identification methods [21], and without loss of generality, the following result can be obtained

$$\text{rank} \left(\begin{bmatrix} Z_p \\ U_{k,s} \end{bmatrix} \right) = \text{number of the row} \Rightarrow \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \text{ is of full row rank.}$$

Therefore, one can obtain

$$[\Gamma_s L_p \ H_{u,s}] = [R_{31} \ R_{32}] \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}^{-} \quad (17)$$

where

$$\begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}^{-} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}^T \left(\begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}^T \right)^{-1}.$$

With the identified matrix $[\Gamma_s L_p \ H_{u,s}]$, the matrices A , B , C , and D can be reliably estimated. To extract A and C from the observability matrix, an estimation $\hat{\Gamma}_s$ is obtained by applying SVD to $\Gamma_s L_p Z_p$

$$\Gamma_s L_p Z_p = [U_1 \ U_2] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \approx 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}. \quad (18)$$

Based on the decomposition, the estimation $\hat{\Gamma}_s$ is given by $\hat{\Gamma}_s = U_1 S_1^{1/2}$. The matrix C is estimated by

$$\hat{C} = \hat{\Gamma}_s(1 : k_y, :). \quad (19)$$

Concurrently, the matrix A can be estimated through the following calculation

$$\begin{aligned} \hat{\Gamma}_s(k_y + 1 : (s+1)k_y, :) &= \begin{bmatrix} CA \\ \vdots \\ CA^s \end{bmatrix} = \begin{bmatrix} C \\ \vdots \\ CA^{s-1} \end{bmatrix} A \triangleq \hat{\Gamma}_{s-1} A \\ \Rightarrow \hat{A} &= (\hat{\Gamma}_{s-1}^T \hat{\Gamma}_{s-1})^{-1} \hat{\Gamma}_{s-1}^T \hat{\Gamma}_s(k_y + 1 : (s+1)k_y, :). \end{aligned} \quad (20)$$

Furthermore, B and D can be estimated based on the Toeplitz matrix $H_{u,s}$. Specifically, the estimations of B and D are obtained as follows:

$$\hat{D} = H_{u,s}(sk_y + 1 : (s+1)k_y, sk_u + 1 : (s+1)k_u), \quad (21)$$

$$H_{u,s}(k_y + 1 : (s+1)k_y, 1 : k_u) = \begin{bmatrix} CB \\ \vdots \\ CA^{s-1}B \end{bmatrix} \triangleq \hat{\Gamma}_{s-1} B \quad (22)$$

$$\Rightarrow \hat{B} = (\hat{\Gamma}_{s-1}^T \hat{\Gamma}_{s-1})^{-1} \hat{\Gamma}_{s-1}^T H_{u,s}(k_y + 1 : (s+1)k_y, 1 : k_u). \quad (23)$$

Although the measurement noise $v(k)$ and process noise $w(k)$ are commonly modeled as zero-mean Gaussian vectors, their covariance matrices Σ_v and Σ_w are generally unavailable in practice. This uncertainty may limit the monitoring performance when the KF-based residual generation is applied. To mitigate this difficulty, the covariance matrices Σ_w and Σ_v are estimated within the KF framework using the identified matrices. The estimation is achieved through iterative interactions between the estimator and the measurement data.

Remark 1. For notational simplicity, the matrices A , B , C , and D used in the subsequent derivations denote their estimated counterparts $\hat{A}$, $\hat{B}$, $\hat{C}$, and $\hat{D}$, respectively.

Based on the KF formulation in (3)–(6), the following expression can be derived:

$$\Sigma_v = \Sigma_{r,k} - CP_k C^T \Rightarrow \begin{cases} \Sigma_v + C\Sigma_w C^T = \Sigma_{r,k+1} - \Phi_{k,0}, \\ \Phi_{k,0} = C(AP_k A^T - K_k \Sigma_{r,k} K_k^T)C^T. \end{cases} \quad (24)$$

Within the KF framework, the following formulation is obtained based on the idea of iterative interaction:

$$\begin{aligned}\Sigma_v + \sum_{i=0}^j CA^i \Sigma_w (CA^i)^T &= \Sigma_{r,k+j+1} - \Phi_{k,j}, \quad j = 0, 1, \dots \\ \Phi_{k,j} &= C \left(A^{j+1} P_k (A^T)^{j+1} - \sum_{i=0}^j A^{j-i} K_{k+i} \Sigma_{r,k+i} K_{k+i}^T (A^{j-i})^T \right) C^T.\end{aligned}\quad (25)$$

Applying the vectorization operation to (24)–(25) yields

$$\begin{aligned}\text{vec}(\Sigma_v) &= \text{vec}(\Sigma_{r,k} - CP_k C^T) \\ \text{vec}(\Sigma_v + \sum_{i=0}^j CA^i \Sigma_w (CA^i)^T) &= \text{vec}(\Sigma_v) + \sum_{i=0}^j [(CA^i) \otimes (CA^i)] \text{vec}(\Sigma_w) \\ &= \text{vec}(\Sigma_{r,k+j+1}) - \text{vec}(\Phi_{k,j}), \quad j = 0, 1, \dots, q.\end{aligned}\quad (26)$$

From (26), the following equivalent form can be derived

$$\begin{bmatrix} I & 0 \\ I & C \otimes C \\ \vdots & \vdots \\ I & \sum_{i=0}^q (CA^i) \otimes (CA^i) \end{bmatrix} \begin{bmatrix} \text{vec}(\Sigma_v) \\ \text{vec}(\Sigma_w) \end{bmatrix} = \begin{bmatrix} \text{vec}(\Sigma_{r,k}) - \text{vec}(CP_k C^T) \\ \text{vec}(\Sigma_{r,k+1}) - \text{vec}(\Phi_{k,0}) \\ \vdots \\ \text{vec}(\Sigma_{r,k+q+1}) - \text{vec}(\Phi_{k,q}) \end{bmatrix}.\quad (27)$$

It is worth noting that Σ_v and Σ_w are diagonal matrices. Therefore, they contain k_y and k_x independent elements, respectively. These independent elements are denoted by ϑ_v and ϑ_w . Accordingly, there exists a matrix $\Psi \in \mathbb{R}^{(k_y^2 + k_x^2) \times (k_y + k_x)}$ such that

$$\begin{bmatrix} \text{vec}(\Sigma_v) \\ \text{vec}(\Sigma_w) \end{bmatrix} = \Psi \begin{bmatrix} \vartheta_v \\ \vartheta_w \end{bmatrix}\quad (28)$$

where

$$\begin{aligned}\vartheta_v &= [\sigma_{v,1}, \sigma_{v,2}, \dots, \sigma_{v,k_y}]^T, \\ \vartheta_w &= [\sigma_{w,1}, \sigma_{w,2}, \dots, \sigma_{w,k_x}]^T,\end{aligned}$$

and $\sigma_{v,i}$ represents the i -th diagonal element of Σ_v , $i = 1, 2, \dots, k_y$, and $\sigma_{w,j}$ denotes the j -th diagonal element of Σ_w , $j = 1, 2, \dots, k_x$.

Substituting (28) into (27) yields

$$\underbrace{\begin{bmatrix} I & 0 \\ I & C \otimes C \\ \vdots & \vdots \\ I & \sum_{i=0}^q (CA^i) \otimes (CA^i) \end{bmatrix}}_{A_{v,w}} \Psi \begin{bmatrix} \vartheta_v \\ \vartheta_w \end{bmatrix} = \begin{bmatrix} \text{vec}(\Sigma_{r,k}) - \text{vec}(CP_k C^T) \\ \text{vec}(\Sigma_{r,k+1}) - \text{vec}(\Phi_{k,0}) \\ \vdots \\ \text{vec}(\Sigma_{r,k+q+1}) - \text{vec}(\Phi_{k,q}) \end{bmatrix}.\quad (29)$$

It should be noted that (29) provides a general formulation for estimating the structured parameters of Σ_v and Σ_w . Based on (29), the estimations of Σ_v and Σ_w can be iteratively updated by solving the following expression:

$$A_{v,w} \begin{bmatrix} \vartheta_{v,i} \\ \vartheta_{w,i} \end{bmatrix} = \begin{bmatrix} \varphi(k) - \text{vec}(CP_k C^T) \\ \varphi(k+1) - \text{vec}(\Phi_{k,0}) \\ \vdots \\ \varphi(k+q+1) - \text{vec}(\Phi_{k,q}) \end{bmatrix} \quad (30)$$

where i represents the iteration index in the covariance update procedure and

$$\varphi(k+j) = \text{vec}([y(k+j) - \hat{y}(k+j)][y(k+j) - \hat{y}(k+j)]^T), \quad j = 0, 1, \dots, q.$$

Accordingly, the estimations of Σ_v and Σ_w are given by

$$\begin{aligned} \hat{\Sigma}_v &= \text{diag}(\hat{\vartheta}_v), \\ \hat{\Sigma}_w &= \text{diag}(\hat{\vartheta}_w). \end{aligned} \quad (31)$$

To enable FD, a residual vector is required for the construction of the test statistic. Using the identified matrices A , B , C , and D , together with the estimated covariance matrices Σ_v and Σ_w , the residual vector $r(k)$ is derived from the KF algorithm (3)–(6). Based on this residual signal, the test statistic for FD is formulated as

$$J(r) = r^T(k) \Sigma_{r,0} r(k) \sim \chi^2(k_y) \quad (32)$$

where $\Sigma_{r,0}$ denotes the covariance matrix of the residual vector under fault-free conditions. According to the statistic in (32), the threshold J_{th} is given by

$$J_{th} = \chi^2_{\alpha}(k_y) \quad (33)$$

where α is a user-selected significance level that sets the upper limit of the acceptable false alarm rate (FAR).

Based on (32) and (33), the following detection logic is proposed for FD; that is,

$$\begin{cases} J(r) - J_{th} > 0, \implies \text{fault alarm} \\ J(r) - J_{th} \leq 0, \implies \text{no alarm.} \end{cases} \quad (34)$$

To construct a reliable FD scheme for traction drive systems, the matrices A , B , C , and D are identified, and covariance matrices Σ_v and Σ_w are estimated in an offline procedure. During online operation, process measurements are utilized to compute the residual signals, which serve as the basis for FD. The overall design of the developed FD system is summarized in Algorithms 1 and 2.

Algorithm 1 Design of the proposed FD scheme: Off-line Learning

Input: Fault-free process I/O data.

Output: $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, $\hat{\Sigma}_v$, and $\hat{\Sigma}_w$.

1: **begin**

2: Load process data and construct matrices $U_{k,s}$, $Y_{k,s}$, and Z_p ;

3: Execute (12)–(17) to obtain $\Gamma_s L_p$ and $H_{u,s}$;

4: Identify $\hat{A}$, $\hat{B}$, $\hat{C}$, and $\hat{D}$ using (19)–(23);

5: Obtain $\hat{\Sigma}_v$ and $\hat{\Sigma}_w$ using (30) and (31).

6: **end**

Algorithm 2 Design of the proposed FD scheme: Online FD**Input:** Online I/O data and offline-estimated parameters.**Output:** Online FD results.

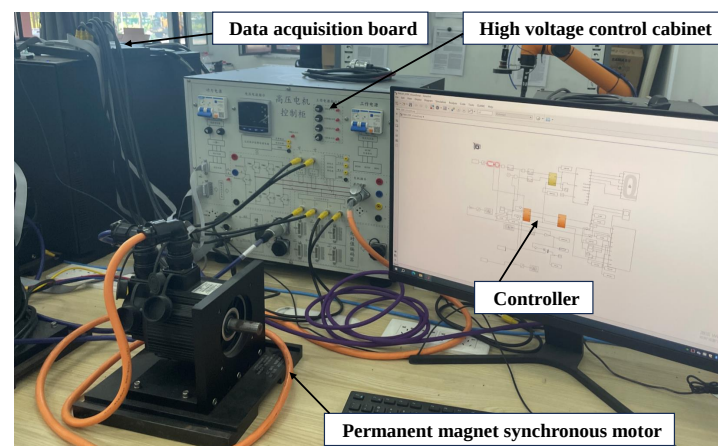
```

1: begin
2:   Load process I/O data;
3:   Execute (3)–(6) to obtain the residual signal;
4:   Design the residual evaluation unit using (32) and (33);
5:   Execute FD based on the detection logic (34).
6: end

```

4. Experiment and Discussion

Experimental validation is conducted on a practical traction drive system to evaluate the proposed approach. The test bench consists of a high-voltage control cabinet, a data acquisition board, a computer for the program implementation, and a permanent magnet synchronous motor, as depicted in Figure 2. The main parameters of the traction motor are listed in Table 2. It should be noted that this work focuses on algorithm validation rather than sensor metrology; therefore, no separate study on sensor uncertainty was conducted. Prior to the experiments, routine checks confirmed that the sensor outputs were stable and within the expected operating ranges under normal conditions. Consequently, remaining sensor inaccuracies are treated as part of the measurement noise in collected data.

**Figure 2.** Experimental setup of the traction drive system.**Table 2.** Key parameters of the traction motor.

Symbol	Parameter	Value (Unit)
T_e	rated torque	2.43 (N·m)
J_m	moment of inertia	0.425×10^{-3} (kg·m ²)
L_a	inductance of motor coil	2.96×10^{-3} (H)
T_L	load torque	3.645 (N·m)
O_a	magnet flux	9.92×10^{-2} (Wb)
p	pole pairs	4
R_a	resistance of motor coil	0.985 (Ω)
D_m	viscosity friction coefficient	1×10^{-4}

In the experimental setup, the voltage measurements are used as input variables, while the speed and current measurements are considered as output variables; that is,

$$\begin{aligned}
 u &= [V_{dc} \ V_a \ V_b]^T, \\
 y &= [I_a \ I_b \ s]^T.
 \end{aligned} \tag{35}$$

I/O data for the parameter estimation are collected from monitoring nodes of the traction drive system, with representative samples shown in Figure 3. To assess the performance of the proposed approach, two distinct fault conditions are introduced in the traction drive system for analysis. These scenarios capture different operational challenges that may arise in practical applications.

1. An offset fault with a magnitude of 0.1 A is injected into I_a at $k = 1001$ st.
2. A drift fault, defined as $f_2 = 0.15 (k - 1001)$ A, is applied to I_b starting from $k = 1001$ st.

To verify the effectiveness of the developed method, a conventional KF-based FD approach [7], a unscented KF-based FD approach [12], and an extended KF-based FD approach [15] are selected as benchmark methods for comparison, with the system matrices considered to be known.

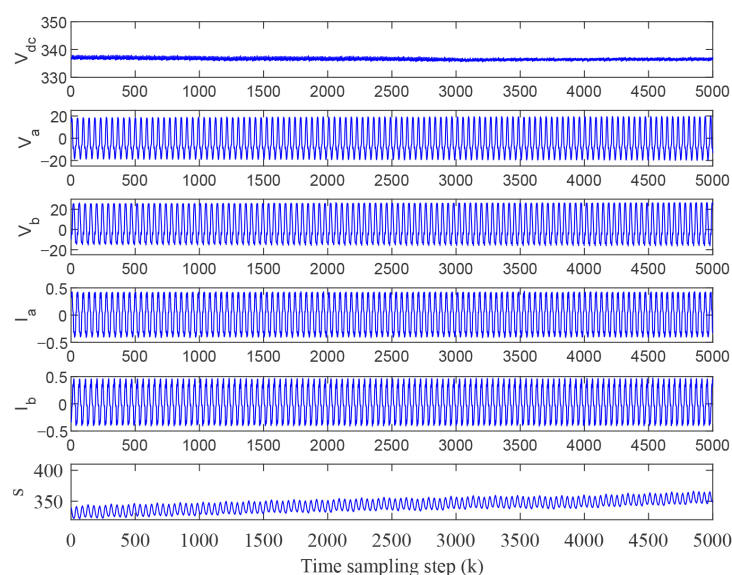


Figure 3. I/O signals of the traction drive system.

Figures 4–7 present the FD results for the offset fault scenario obtained using the three benchmark methods and the proposed data-driven approach, respectively. Figures 4–6 indicate that deviations between the prior-assumed noise statistics and the true system noise statistics can impair the FD performance. Although the conventional KF-based, unscented KF-based, and extended KF-based FD methods, as well as the proposed method, exhibit comparable FAR, the three benchmark methods suffer from limited fault detection rates (FDR), resulting in unbalanced detection performance. It is worth noting that the abrupt increase in the test statistic at the beginning mainly results from the initialization transient of the KF rather than the fault itself. The discrepancy between the initial state estimate and the true system state produces large innovations during the initial sampling instants. Because the residual sequence has not yet converged to its steady-state distribution, the test statistic is temporarily amplified, which leads to the observed initial spike. For the proposed method, the FAR remains within the acceptable limit, while the desired detection performance is effectively maintained.

Figures 8–11 show the detection results for the drift fault scenario using the three benchmark methods and the proposed approach, respectively. A similar initial spike of the test statistic can also be observed. This phenomenon mainly results from the transient behavior of the iterative algorithm during its initialization stage. Since the parameter estimates have not yet converged in the early iterations, the corresponding test statistic may be temporarily amplified. As observed from the figures, the proposed method exhibits a relatively low FDR for the drift fault scenario. This is mainly attributed to the gradual nature

of the drift fault, where the short-term variation in system outputs may still remain within the normal fluctuation range, making early identification difficult. Nevertheless, compared with the conventional KF-based, unscented KF-based, and extended KF-based FD methods, the proposed data-driven method achieves more reliable FD, leading to improved detection performance.

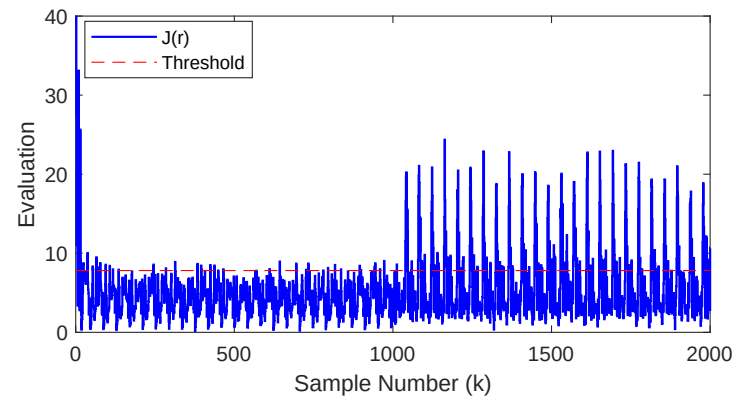


Figure 4. Offset fault detection results using the conventional KF-based method.

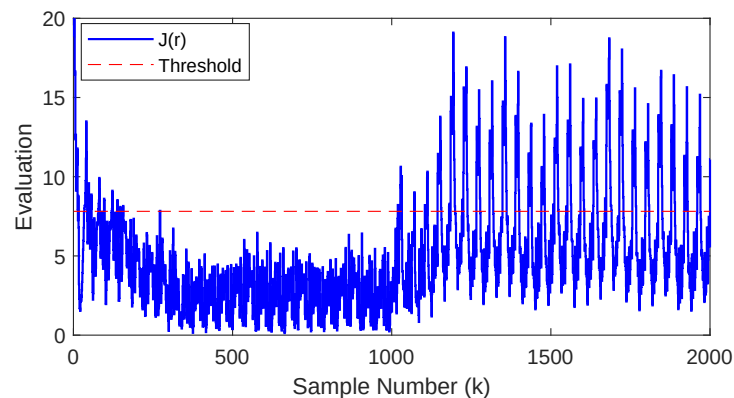


Figure 5. Offset fault detection results using the unscented KF-based method.

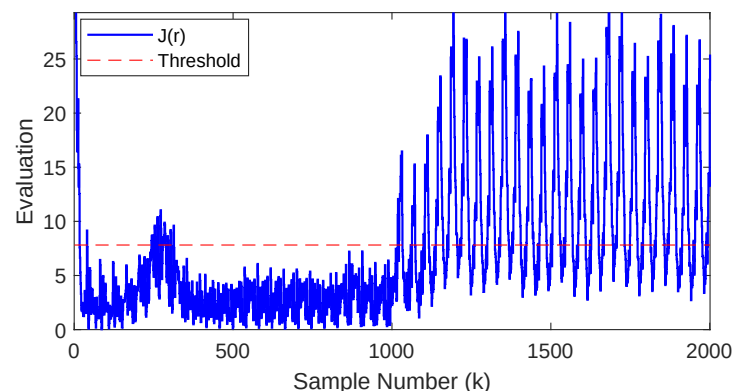


Figure 6. Offset fault detection results using the extended KF-based method.

To further demonstrate the effectiveness of the proposed method, Table 3 provides a quantitative comparison of the performance achieved by the benchmark FD methods and the proposed method. Although using the same identified system matrices as the proposed method, the conventional KF-based, unscented KF-based, and extended KF-based FD methods neglect the mismatch between assumed and actual noise statistics, resulting in degraded monitoring performance. In contrast, the proposed method effectively balances the FAR and FDR. Moreover, it does not rely on precise system modeling, thereby avoid-

ing complex model construction while still achieving the desired detection performance. The results in Figures 4–11 and Table 3 demonstrate that, despite using the same identified system matrices, the conventional KF-based, unscented KF-based, and extended KF-based FD methods neglect the mismatch between assumed and actual noise statistics, resulting in degraded detection performance. In contrast, the proposed approach alleviates this limitation through iterative interaction and learning, effectively suppressing the adverse effects caused by noise statistics mismatch.

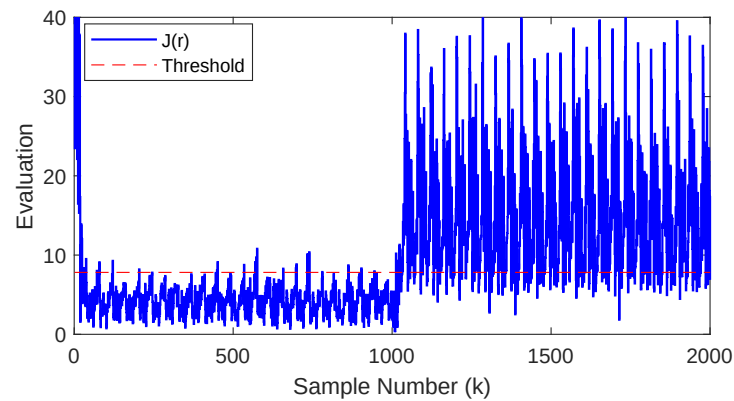


Figure 7. Offset fault detection results using the proposed method.

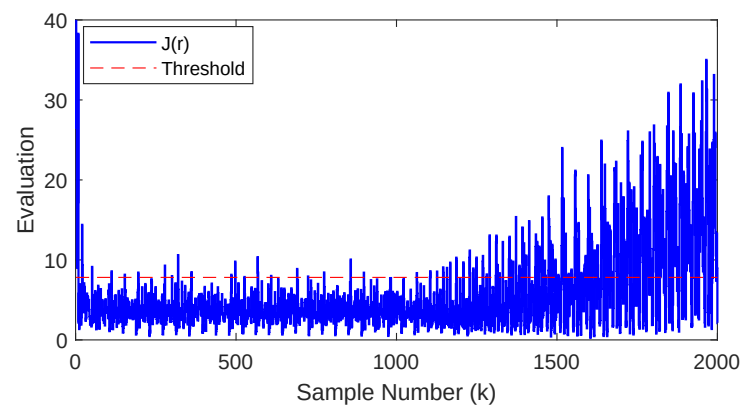


Figure 8. Drift fault detection results using the conventional KF-based method.

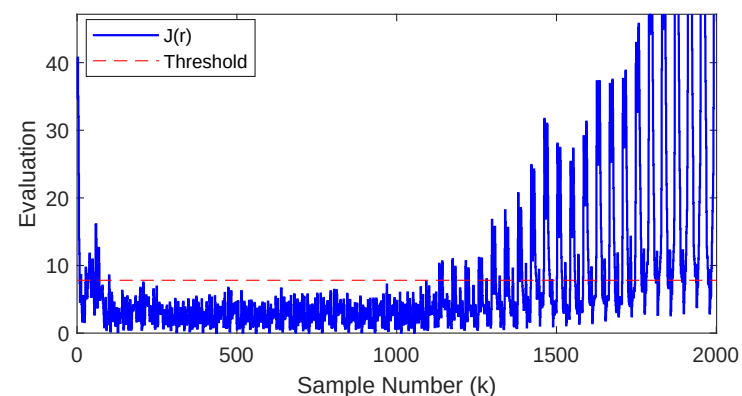


Figure 9. Drift fault detection results using the unscented KF-based method.

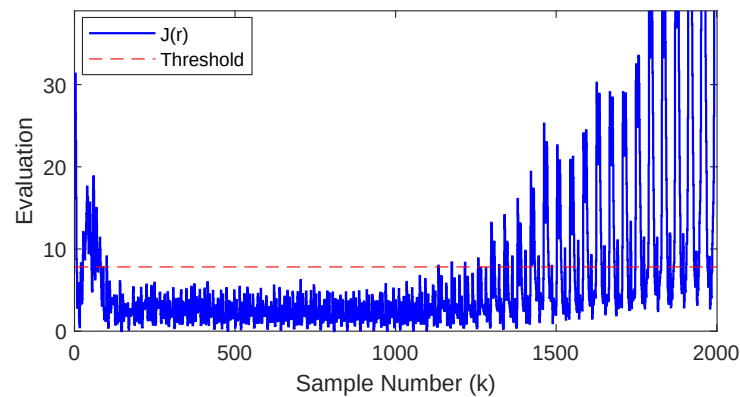


Figure 10. Drift fault detection results using the extended KF-based method.

Although the proposed method achieves satisfactory FD performance, several limitations still remain. First, the current framework is developed based on a linear state-space model, which is derived by linearizing the traction drive system around a stable operating point. Its applicability may be limited under strongly nonlinear conditions when the operating state of the traction drive system varies significantly. On the other hand, the proposed method focuses on FD and does not explicitly address fault-tolerant control. In practical traction drive systems, reliable monitoring should be further integrated with control reconfiguration to ensure safe operation under complex conditions.

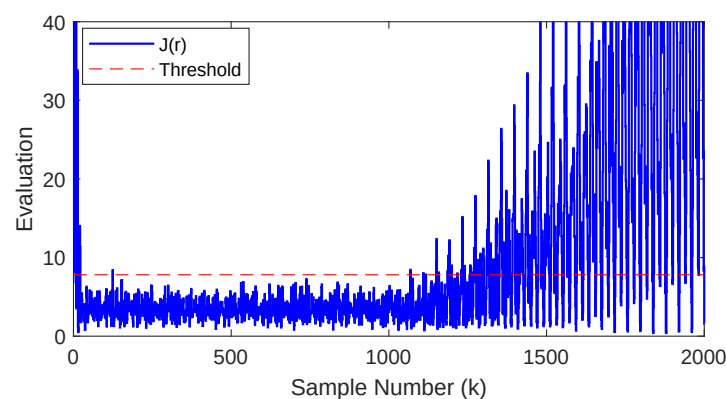


Figure 11. Drift fault detection results using the proposed data-driven method.

Table 3. Comparison of FD performance for different methods.

Methods	f_1 : Offset Fault		f_2 : Drift Fault	
	FAR	FDR	FAR	FDR
The method in [7]	0.0510	0.2430	0.0290	0.3340
The method in [12]	0.0450	0.2840	0.0310	0.4240
The method in [15]	0.0460	0.5920	0.0560	0.3360
The developed method	0.0460	0.8040	0.0150	0.5340

5. Conclusions

In this work, a data-driven FD scheme has been developed for traction drive systems in the presence of unknown noise covariances. To avoid reliance on accurate system modeling, the system matrices were directly identified from collected I/O data based on the principle of subspace identification. To address the challenge posed by unknown noise statistics, an adaptive covariance estimation strategy was incorporated into the KF framework. By exploiting the iterative interaction between the estimator and measurement data, the noise covariance matrices can be progressively updated, thereby mitigating the per-

formance impairment caused by empirically specified covariance matrices. The proposed approach was experimentally validated on a real traction drive system. It achieves an FDR of 0.8040 for offset faults under comparable FAR and an FAR of 0.0150 for drift faults under acceptable FDR, both outperforming the benchmark methods. These results demonstrate that the proposed method delivers reliable and satisfactory FD performance. It should be noted that the developed method is established for linear dynamic systems. Future research will aim to extend the developed framework to nonlinear systems and enhance fault-tolerant control capability under more complex operating conditions.

Author Contributions: Conceptualization, C.F. and C.J.; methodology, C.F. and C.J.; software, C.F., S.W. and Z.W.; validation, C.F., S.W. and Z.W.; formal analysis, C.F.; investigation, C.F.; resources, C.J.; data curation, C.F. and S.W.; writing—original draft preparation, C.F., S.W. and Z.W.; writing—review and editing, P.C. and C.F.; visualization, C.F. and P.C.; supervision, C.J.; project administration, C.J.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the National Natural Science Foundation of China under Grant 62273058.

Data Availability Statement: Requests for access to the data supporting this study should be directed to the authors.

Acknowledgments: The authors would like to thank the editors and reviewers for their constructive comments, which has helped improve the quality of this work.

Conflicts of Interest: Author Peng Cheng was employed by the company CRRC Changchun Railway Vehicles Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FD	Fault detection
KF	Kalman filter
I/O	Input–output
SVD	Singular value decomposition
FAR	False alarm rate
FDR	Fault detection rate

References

1. Chen, H.; Jiang, B.; Ding, S.X.; Huang, B. Data-Driven Fault Diagnosis for Traction Systems in High-Speed Trains: A Survey, Challenges, and Perspectives. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 1700–1716. [\[CrossRef\]](#)
2. Xu, J.; Zhong, M.; Li, L.; Wu, Y.; Song, B. A Fuzzy H_i/H_∞ Optimization Approach to Fault Detection of High-Speed Train Traction Motor Systems. *IEEE Trans. Ind. Inform.* **2025**, *21*, 3655–3665. [\[CrossRef\]](#)
3. Zhong, M.; Zhang, K.; Zhang, L.; Zhong, N.; Guo, H. An innovation re-organised parity space approach to fault detection for rectifier in high-speed train electrical traction systems. *Int. J. Syst. Sci.* **2026**. [\[CrossRef\]](#)
4. Mao, Z.; Tao, G.; Jiang, B.; Yan, X.G. Adaptive Compensation of Traction System Actuator Failures for High-Speed Trains. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 2950–2963. [\[CrossRef\]](#)
5. Zhang, K.; Jiang, B.; Yan, X.G.; Shen, J. Interval Sliding Mode Observer Based Incipient Sensor Fault Detection with Application to a Traction Device in China Railway High-Speed. *IEEE Trans. Veh. Technol.* **2019**, *68*, 2585–2597. [\[CrossRef\]](#)
6. Kolpakhchyan, P.G.; Pakhomin, S.A.; Kochin, A.E.; Evstaf'ev, A.M.; Andreev, V. Traction Induction Motor State Observer Based on an Luenberger Filter. In *Proceedings of the International Conference on Intelligent Information Technologies for Industry, St. Petersburg, Russia, 25–30 September 2023*; Springer: Cham, Switzerland, 2023; pp. 260–270.
7. Ding, S.X. *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*; Springer: Berlin/Heidelberg, Germany, 2008.
8. Chen, H.; Jiang, B. A Review of Fault Detection and Diagnosis for the Traction System in High-Speed Trains. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 450–465. [\[CrossRef\]](#)

9. Cheng, C.; Wan, Z.; Wang, W.; Sun, W.; Fu, C.; Chen, H. A Robust Data-Driven Sensor Fault Detection Method for Traction Drive Systems. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3562009. [[CrossRef](#)]
10. Sun, X.; Song, C.; Zhang, Y.; Sha, X.; Diao, N. An Open-Circuit Fault Diagnosis Algorithm Based on Signal Normalization Preprocessing for Motor Drive Inverter. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3513712. [[CrossRef](#)]
11. Xia, L.; Liang, Y.; Zheng, P.; Huang, X. Residual-Hypergraph Convolution Network: A Model-Based and Data-Driven Integrated Approach for Fault Diagnosis in Complex Equipment. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3501811. [[CrossRef](#)]
12. Cheng, C.; Wang, W.; Meng, X.; Shao, H.; Chen, H. Sigma-mixed unscented Kalman filter-based fault detection for traction systems in high-speed trains. *Chin. J. Electron.* **2023**, *32*, 982–991. [[CrossRef](#)]
13. Foo, G.H.B.; Zhang, X.; Vilathgamuwa, D.M. A Sensor Fault Detection and Isolation Method in Interior Permanent-Magnet Synchronous Motor Drives Based on an Extended Kalman Filter. *IEEE Trans. Ind. Electron.* **2013**, *60*, 3485–3495. [[CrossRef](#)]
14. Namdar, A.; Samet, H.; Allahbakhshi, M.; Tajdinian, M.; Ghanbari, T. A robust stator inter-turn fault detection in induction motor utilizing Kalman filter-based algorithm. *Measurement* **2022**, *187*, 110181. [[CrossRef](#)]
15. Miniach, M.; Orłowska-Kowalska, T. Innovative extended Kalman filter in a coherent system for detection and compensation of various current sensor faults in the induction motor drive. *IEEE Trans. Ind. Electron.* **2025**, *72*, 7772–7784. [[CrossRef](#)]
16. Huang, B.; Kadali, R. *Dynamic Modeling, Predictive Control and Performance Monitoring: A Data-Driven Subspace Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008; Volume 374.
17. Favoreel, W.; De Moor, B.; Van Overschee, P. Subspace state space system identification for industrial processes. *J. Process Control* **2000**, *10*, 149–155. [[CrossRef](#)]
18. Goldstein, H. Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika* **1986**, *73*, 43–56. [[CrossRef](#)]
19. Ding, S.X.; Yang, Y.; Zhang, Y.; Li, L. Data-driven realizations of kernel and image representations and their application to fault detection and control system design. *Automatica* **2014**, *50*, 2615–2623. [[CrossRef](#)]
20. Ding, S.X. *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*; Springer: London, UK, 2014.
21. Qin, S.J. An overview of subspace identification. *Comput. Chem. Eng.* **2006**, *30*, 1502–1513. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.