Physics-Informed Feature Generation for Fault Classification in Industrial Drive Systems

Rohan Arun Nawale[1], Martin Grotjahn[2], Timo von Marcard[2], Charlotte Tkany[2] and Moritz Fehsenfeld[3]

***Abstract*—Condition monitoring of industrial drive systems is challenging because mechanical faults are often difficult to identify directly from motor-side measurements, while purely data-driven approaches require large labeled datasets and show limited generalization and poor interpretability. To address these limitations, this paper proposes a physics-informed feature generation (PiFG) workflow for supervised fault classification. In contrast to approaches that embed physics into the learning objective, the proposed method incorporates physical knowledge at the feature-generation level using an Extended Kalman Filter (EKF). A simplified dynamic model is used by an EKF to estimate physically meaningful parameters directly from standard motor signals. Experimental validation on an industrial belt-drive testbench demonstrates improved feature separability leading to substantial improvements over a purely data-driven baseline. Consequently, increasing fault classification accuracy under limited training data from 62.25% to 100% while eliminating the severe fault class confusion and improving cross-condition generalization accuracy from 41.38% to 75.77% across varying operating speeds.**

***Index Terms*—Condition monitoring, fault classification, physics-informed machine learning, Extended Kalman Filter, physical parameter estimation, feature extraction, elastic-drive systems.**

I. Introduction

Belt-driven and elastic drive systems are core components of modern industrial automation due to their mechanical simplicity, flexibility, and cost-effectiveness. During operation, however, these systems are subject to gradual mechanical degradation caused by wear, changing load conditions and environmental influences. Such degradation typically manifests as changes in physical system properties, for example increased friction, belt deterioration, or load imbalance, which may progressively reduce efficiency and reliability and lead to costly maintenance interventions. Reliable condition monitoring is therefore required to detect these changes at an early stage. In practice, however, the available measurements in industrial drive systems, such as motor torque, position, or velocity, are strongly influenced by operating trajectories and control inputs. The underlying mechanical degradation is often difficult to identify directly from raw measurement signals, making robust fault classification particularly challenging under varying operating regimes and limited labeled fault data.

To detect such incipient physical faults early, existing condition monitoring approaches can generally be classified into three categories [1], [2]:

1) Data-driven (Black-box) models;
2) Physics-based and/or observer-based models;
3) Physics-informed machine learning (PIML) models.

Data-driven condition monitoring models can achieve high classification accuracy when large labeled datasets are available [3], [4]. In industrial settings, however, faults are rare events, and fault data is often scarce, which limits robust training and transferability [5]. Since models are trained directly on raw measurements, they may overfit to operating conditions, sensor placement, or noise characteristics rather than the underlying physical degradation mechanisms, and degrade when applied to new machines and unseen operating regimes [6], [7]. The limited interpretability of black-box models can be problematic in safety-critical applications [3].

Physics- and observer-based approaches provide physically interpretable diagnostic indicators by explicitly modeling the system dynamics [8]. Observers such as Kalman Filter and sliding mode observers are widely used for state and parameter estimation and have been applied in industrial condition monitoring [10]–[12]. Similarly, robust control frameworks based on Kalman filter-enhanced 6-DOF quadrotor control have demonstrated effective state estimation and precise position control under noise, disturbances, and modeling uncertainties [9]. Such approaches typically rely on accurate dynamic models and require extensive controller and estimator tuning for specific operating regimes, limiting their scalability [15], [16]. In practice, nonlinearities, parameter uncertainty, degradation, and unmodeled dynamics make accurate modeling difficult [13].

Physics-informed machine learning (PIML) aims to combine physical knowledge with data-driven learning to improve robustness and reduce data requirements [17], [18]. Integrated workflows based on vibration feature extraction, decision trees, and extreme learning machines remain largely dependent on signal-domain statistical descriptors, lack physically interpretable parameter representations linked to the underlying system dynamics, and are primarily limited to degradation-state monitoring rather than discrimination of multiple physical fault types [19]. Physics-informed neural networks (PINNs) enforce physical constraints during training but can require careful loss weighting and extensive tuning [20], [21]. Neural network-based observers have also been explored for

*This work was supported by European Union (EU) at its European Regional Development Fund (ERDF) and the Ministry of Science and Culture of Lower Saxony (MWK) via research program "Innovation through universities and research institutions" (87013588, PiFD-CoM).

[1]Rohan Arun Nawale is with Institute of Engineering Design, Mechatronics and Electromobility, University of Applied Sciences, 30173 Hannover, Germany rohan-arun.nawale@hs-hannover.de

[2]Martin Grotjahn, Timo von Marcard and Charlotte Tkany are with Institute of Engineering Design, Mechatronics and Electromobility, University of Applied Sciences, 30173 Hannover, Germany {martin.grotjahn, timo.marcard, charlotte.tkany}@hs-hannover.de

[3]Moritz Fehsenfeld is with Lenze SE, 31855 Aerzen, Germany moritz.fehsenfeld@lenze.com

nonlinear friction estimation [22], yet generalization under limited physics knowledge remains challenging [23]. Related hybrid approaches based on observer residuals and machine learning can improve diagnostic performance, but often depend on accurate observer design and multiple sensor signals, reducing robustness under model mismatch [24].

The preceding discussion highlights that integrating physical knowledge into machine learning (ML) can reduce dependency on large labeled datasets while improving physical interpretability. Furthermore, lightweight frameworks that integrate simplified physical models directly into the feature-generation process of ML pipelines remain insufficiently explored.

To address this gap, this paper proposes a physics-informed feature generation strategy for fault classification of elastic drive systems. An Extended Kalman Filter (EKF) utilizing the simplified physical system model is used to estimate physically meaningful parameters such as equivalent inertia and friction directly from motor-side measurements. These estimated parameters characterize healthy and faulty operating conditions and are used to augment conventional signal-based features for machine learning classification.

The main contributions of this paper are:

- A physics-informed feature generation (PiFG) strategy that utilizes an Extended Kalman Filter (EKF) with a simplified physical system model to estimate fault-sensitive parameters, such as equivalent inertia and friction, directly from standard motor-side measurements.
- A hybrid fault-classification pipeline that combines features extracted from raw measurement signals with features derived from EKF-estimated physical parameters for supervised machine learning-based fault classification.
- Experimental validation on an industrial belt-drive test-bench under varying operating conditions and limited training-data scenarios to evaluate fault classification capability and cross-condition generalization.

II. Industrial drive systems

The system considered in this work represents a class of elastic drive systems commonly used in industrial automation, where the load is connected elastically (e.g. belt, gearbox) to the motor. In such systems, mechanical degradation typically manifests as changes in physical parameters such as friction or inertia. To describe the dynamic behavior of these systems and to enable parameter estimation, elastic drives are commonly modeled as two-mass rotational systems, as illustrated in Fig. 1. Equations (1)–(4) describe the corresponding system dynamics:

$$J_{\mathrm{M}}\ddot{\varphi}_{\mathrm{M}} = M_{\mathrm{M}} - M_{\mathrm{M,fric}} - M_{\mathrm{coupling}}, \tag{1}$$

$$J_{\mathrm{L}}\ddot{\varphi}_{\mathrm{L}} = M_{\mathrm{coupling}} - M_{\mathrm{L,fric}}, \tag{2}$$

$$M_{\mathrm{coupling}} = k(\varphi_{\mathrm{M}} - \varphi_{\mathrm{L}}) + d(\dot{\varphi}_{\mathrm{M}} - \dot{\varphi}_{\mathrm{L}}), \tag{3}$$

$$M_{\mathrm{i,fric}} = b_{\mathrm{i}}\dot{\varphi}_{\mathrm{i}} + M_{\mathrm{i,c}}\ \mathrm{sgn}(\dot{\varphi}_{\mathrm{i}}), \quad \mathrm{i} \in \{\mathrm{M}, \mathrm{L}\}. \tag{4}$$

The two-mass system consists of the motor inertia J_{M} and load inertia J_{L}, connected by an elastic transmission element modeled as a spring-damper pair. The angular positions of the motor and load are denoted by φ_{M} and φ_{L}, respectively, while M_{M} represents the motor torque. The torque transmitted through the elastic coupling is denoted by M_{coupling} and is described by (3). Friction effects on both the motor and load sides are modeled using a Coulomb-viscous friction model consisting of viscous damping b and Coulomb friction torque M_{c}, as defined in (4).

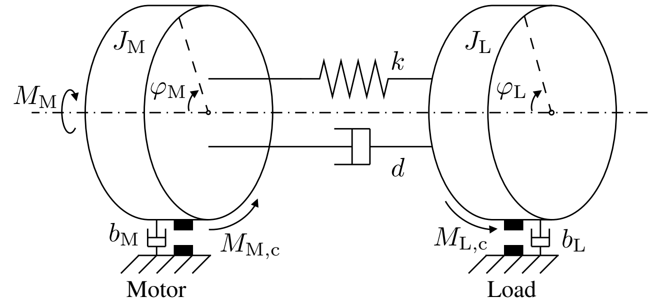

Fig. 1. Two-mass rotational drive model used to describe the motor-load dynamics of elastic transmission systems.

Although the two-mass formulation captures the torsional elasticity of the drive system, in typical industrial motion profiles, the transmission stiffness k is sufficiently high such that the torsional deflection remains small and only the slow rigid-body mode is observable in motor-side measurements [25]. Consequently, the relative angular deflection satisfies $\varphi_{\mathrm{M}} - \varphi_{\mathrm{L}} \approx 0$, allowing the system to be approximated by a stiff one-mass model. The objective is not high-fidelity physical simulation, but estimation of fault-sensitive physical parameters. Therefore, a simplified rigid-body representation is employed to estimate equivalent physical parameters directly from motor-side measurements.

This simplification reasonably neglects internal elasticity but significantly reduces the model complexity. In the resulting second-order model, the system dynamics are described using an equivalent inertia J_{total} and combined friction parameters b_{total} and $M_{\mathrm{total,c}}$. This simplified dynamic model forms the basis for the recursive state and parameter estimation used in the proposed physics-informed feature generation approach.

III. Physics-informed machine learning approach

Motor-side measurements are processed along two parallel feature generation paths. In the first path, statistical features are extracted directly from the measurement signals, forming the data-driven feature vector $\mathbf{z}_{\mathrm{meas}}$. In the second path, an Extended Kalman Filter based on a simplified physical model is used to infer fault-sensitive physical parameters. Statistical features extracted from these estimated parameters form the physics-informed feature vector $\mathbf{z}_{\mathrm{p}}$. Finally, the two feature sets are concatenated to obtain the combined feature vector $\mathbf{z}_{\mathrm{hybrid}} = [\mathbf{z}_{\mathrm{meas}}, \mathbf{z}_{\mathrm{p}}]$, which is used within a supervised learning pipeline. The resulting hybrid processing pipeline is illustrated in Fig. 2.

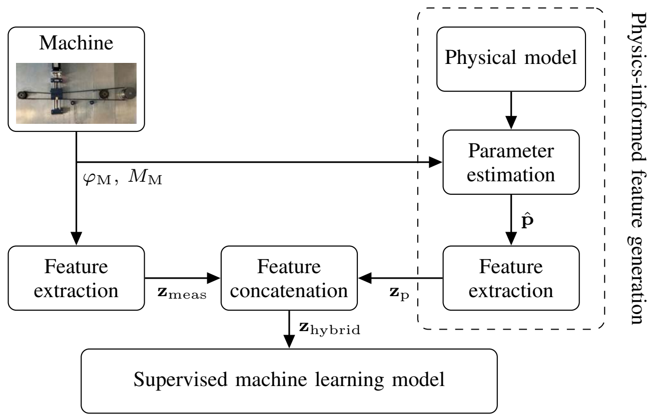

Fig. 2. Schematic overview of the proposed approach with parallel data-driven and physics-informed feature generation (PiFG) paths fused for supervised fault classification.

A. EKF-based State and Parameter Estimation

To enable physics-informed feature generation, an Extended Kalman Filter (EKF) is formulated based on a simplified rigid-body equation of motion as:

$$J_\mathrm{total}\ddot{\varphi}_\mathrm{M} = M_\mathrm{M} - b_\mathrm{total}\dot{\varphi}_\mathrm{M} - M_\mathrm{total,c}\,\mathrm{sgn}(\dot{\varphi}_\mathrm{M}) + M_\mathrm{d}, \quad (5)$$

where M_d is an augmented unmodeled disturbance term. The estimator processes the available motor-side measurements and recursively estimates system states and parameters. The EKF operates on an augmented state vector defined as:

$$\hat{\mathbf{x}} = \begin{bmatrix}\varphi_\mathrm{M} & \dot{\varphi}_\mathrm{M} & \hat{J}_\mathrm{total} & \hat{b}_\mathrm{total} & \hat{M}_\mathrm{total,c} & \hat{M}_\mathrm{d}\end{bmatrix}^\top, \quad (6)$$

where φ_M and $\dot{\varphi}_\mathrm{M}$ denote the motor position and velocity states, while $\hat{J}_\mathrm{total}$, $\hat{b}_\mathrm{total}$, and $\hat{M}_\mathrm{total,c}$ represent the estimated physical parameters. Finally, $\hat{M}_\mathrm{d}$ is included as a lumped disturbance torque. These parameters correspond to physical parameters and quantities that are directly affected by common mechanical faults such as increased friction or load imbalance, and therefore provide fault-sensitive descriptors of the system condition.

The nonlinear state-space model is formulated in discrete time as:

$$\hat{\mathbf{x}}_\mathrm{k+1} = f(\hat{\mathbf{x}}_\mathrm{k}, \mathbf{u}_\mathrm{k}) + \mathbf{w}_\mathrm{k}, \quad (7)$$

$$\mathbf{y}_\mathrm{k} = h(\hat{\mathbf{x}}_\mathrm{k}) + \mathbf{v}_\mathrm{k}, \quad (8)$$

where $f(\cdot)$ and $h(\cdot)$ denote the nonlinear process and measurement models, respectively. $\mathbf{u}_\mathrm{k}$ denotes the system input (motor torque M_M), while $\mathbf{y}_\mathrm{k}$ represents the available motor-side measurements (e.g., angular position φ_M). $\mathbf{w}_\mathrm{k}$ and $\mathbf{v}_\mathrm{k}$ represent process and measurement noise. These terms are assumed to be zero-mean Gaussian with covariance matrices Q and R, respectively, i.e., $\mathbf{w}_\mathrm{k} \sim \mathcal{N}(0, Q)$ and $\mathbf{v}_\mathrm{k} \sim \mathcal{N}(0, R)$.

State and parameter estimation are performed using the standard EKF prediction and measurement update steps; the detailed formulation can be found in [26]. The Coulomb friction term involves the non-differentiable function $\mathrm{sgn}(\cdot)$, a smooth approximation $\tanh(\dot{\varphi}_\mathrm{M}/\epsilon)$ with small $\epsilon \ll 1$ is employed to ensure differentiability during linearization.

The process noise covariance matrix Q models uncertainty in the system dynamics and slowly varying physical quantities. The measurement noise covariance matrix R represents the uncertainty of the motor-side measurements and is chosen according to sensor characteristics. Both Q and R are calibrated empirically to ensure stable convergence and physically consistent state and parameter estimates under nominal operating conditions.

B. Feature Extraction

The fault classification task is formulated based on statistical features extracted from time-series signals. Continuous time-series measurements are segmented into fixed-length, non-overlapping windows, where the window length N is treated as a hyperparameter.

The motor-side measurement vector used for feature extraction is defined as:

$$\mathbf{x}_\mathrm{meas} = \begin{bmatrix}\varphi_\mathrm{M} & \dot{\varphi}_\mathrm{M} & \ddot{\varphi}_\mathrm{M} & M_\mathrm{M}\end{bmatrix}^\top \in \mathbb{R}^{n_\mathrm{z}}, \quad (9)$$

where φ_M and $\dot{\varphi}_\mathrm{M}$ denote the motor angular position and angular velocity, respectively. The acceleration $\ddot{\varphi}_\mathrm{M}$ is obtained by discrete differentiation of the position signal followed by first-order low-pass filtering with a time constant $T_\mathrm{f} = 0.6\,\mathrm{ms}$. The corresponding motor torque is denoted by M_M.

In addition, the EKF provides estimates of the physical system parameters, forming the parameter vector described as:

$$\hat{\mathbf{p}} = \begin{bmatrix}\hat{J}_\mathrm{total} & \hat{b}_\mathrm{total} & \hat{M}_\mathrm{total,c} & \hat{M}_\mathrm{d}\end{bmatrix}^\top \in \mathbb{R}^{n_\mathrm{p}}. \quad (10)$$

Thus, a total of $n_\mathrm{z} = 4$ measurement signals and $n_\mathrm{p} = 4$ estimated parameter signals are available. For each signal, $n_\mathrm{f} = 13$ statistical descriptors are extracted from the corresponding time window.

The extracted features include both time-domain and frequency-domain descriptors. In the time domain, standard statistical descriptors are computed to characterize the signal distribution and amplitude variations . These include measures of central tendency (mean, median), dispersion (standard deviation, interquartile range), amplitude bounds (minimum and maximum), and signal energy (RMS), as well as higher-order statistics such as skewness and kurtosis to capture distribution asymmetry and impulsive behavior.

In the frequency domain, features are extracted from the zero-mean windowed signal using its spectral representation obtained via the discrete Fourier transform. The spectral centroid characterizes the dominant frequency content, while the spectral entropy, computed from the power spectral density estimated via Welch's method, quantifies the concentration or dispersion of spectral energy [27].

Let $\mathcal{F}(\cdot)$ denote the feature extraction operator. The data-driven feature vector is defined as:

$$\mathbf{z}_\mathrm{meas} = \mathcal{F}(\mathbf{x}_\mathrm{meas}) \in \mathbb{R}^{n_\mathrm{z} n_\mathrm{f}}, \quad (11)$$

while the physics-informed feature vector extracted from the EKF-estimated parameter vector is described as:

$$\mathbf{z}_{\mathrm{p}} = \mathcal{F}(\hat{\mathbf{p}}) \in \mathbb{R}^{n_{\mathrm{p}} n_{\mathrm{f}}}. \tag{12}$$

The final feature vector used for classification is obtained by concatenating both representations:

$$\mathbf{z}_{\mathrm{hybrid}} = \begin{bmatrix} \mathbf{z}_{\mathrm{meas}} & \mathbf{z}_{\mathrm{p}} \end{bmatrix} \in \mathbb{R}^{(n_{\mathrm{z}}+n_{\mathrm{p}})n_{\mathrm{f}}}. \tag{13}$$

For the present study, this results in a feature vector of dimension $\mathbf{z}_{\mathrm{hybrid}} \in \mathbb{R}^{104}$. Finally, to ensure uniform scaling across all inputs for the downstream classifier, all features are standardized using Z-score normalization.

C. Machine Learning-Based Fault Classification

The fault classification problem is formulated as a supervised learning task mapping the combined feature vector $\mathbf{z}_{\mathrm{hybrid}}$ to discrete class labels $c \in \mathcal{C}$ from the finite class set $\mathcal{C}$. A Random Forest (RF) classifier is employed to evaluate the combined feature vector $\mathbf{z}_{\mathrm{hybrid}}$. The classifier consists of an ensemble of N_{tree} decision trees $\{g_i\}_{i=1}^{N_{\mathrm{tree}}}$, where the final prediction $\hat{c}$ is obtained by majority voting:

$$\hat{c} = \mathrm{mode}\{g_1(\mathbf{z}_{\mathrm{hybrid}}), \ldots, g_{N_{\mathrm{tree}}}(\mathbf{z}_{\mathrm{hybrid}})\}. \tag{14}$$

Hyperparameters, including the number of trees and tree depth, are optimized via grid search combined with cross-validation exclusively on the training data. The trained model is subsequently evaluated on unseen test data to assess classification performance.

IV. Experimental Setup

To evaluate the proposed physics-informed condition monitoring approach, experiments were conducted on a toothed belt-drive testbench shown in Fig. 3. The system is driven by an industrial synchronous servo motor equipped with an integrated position encoder. The recorded signals used are described in equation (9).

Four operating conditions were experimentally induced to represent fault-free and faulty system states:

- **Healthy:** fault-free operation;
- **Friction (M):** increased motor-side friction;
- **Friction (L):** increased load-side friction;
- **Imbalance:** eccentric load.

The induced faults in the training and testing experiments have identical physical magnitudes (e.g., added imbalance mass or increased friction). The system was excited using trapezoidal velocity point-to-point trajectories with direction reversals, representative of standard industrial operation. The resulting motor velocity $\dot{\varphi}_{\mathrm{M}}$ and motor torque M_{M} for a representative $27\,\mathrm{s}$ segment are shown in Fig. 4. The trajectory comprises motion segments at angular velocity magnitudes of $\pm 34.91, 104.72, 174.53$, and $244.35\,\mathrm{rad/s}$.

The training and testing data originate from independent measurements. Each recording consists of 600,000 samples acquired at a sampling rate of $1\,\mathrm{kHz}$. For feature extraction, the recording is partitioned into 24 non-overlapping windows of 25,000 samples each. This window length ensures that each window contains multiple motion segments of the excitation trajectory while maintaining sufficient temporal resolution for fault detection. The windowing is performed within each run and therefore does not affect the run-level separation between training and testing data.

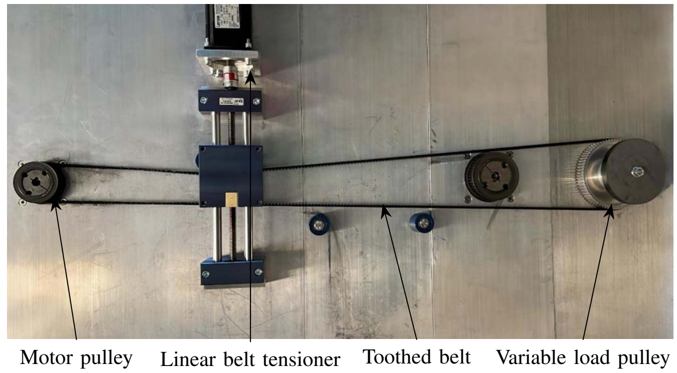

Fig. 3. Toothed belt-drive testbench.

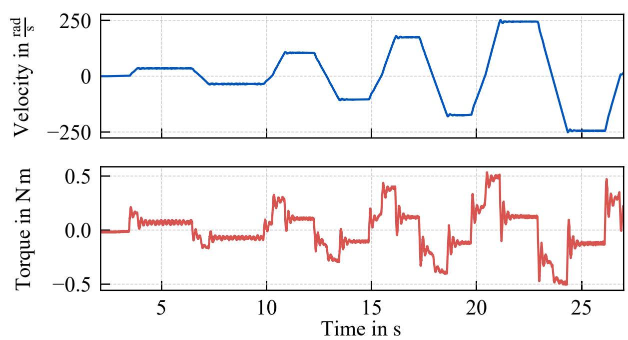

Fig. 4. Measured excitation trajectory of the belt-drive system.

V. Results

To evaluate the impact of physics-informed feature generation, two classification models are considered:

- **Data-driven baseline:** a classifier operating only on measurement-based features $\mathbf{z}_{\mathrm{meas}}$ defined in (11).
- **Proposed hybrid approach:** a classifier operating on the augmented feature vector $\mathbf{z}_{\mathrm{hybrid}}$ defined in (13).

The experimental evaluation focuses on classification performance, data efficiency under limited training data, and cross-condition generalization.

1) Classification Performance: The resulting classification performance is summarized in the confusion matrices shown in Fig. 6. The matrices represent the average prediction results over 20 independent model realizations, obtained by training the classifier without fixing the random seed.

The data-driven model exhibits significant class confusion, particularly for the load-side friction fault, where misclassification occurs across multiple conditions. In contrast, the

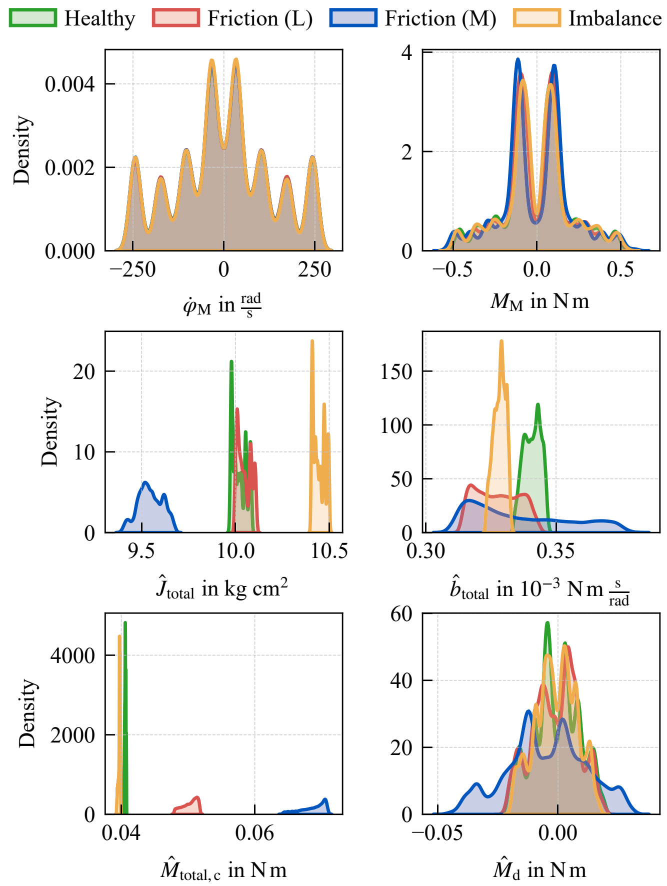

Fig. 5. Probability density distributions of raw sensor measurements ($\dot{\varphi}_{\mathrm{M}}$ and M_{M}) and EKF-estimated physical parameters across different fault conditions (color-coded) over a 600 s recording. The estimated parameters exhibit improved class separation compared to the measurement signals.

proposed approach achieves perfect separation across all fault classes.

This difference can be explained by the distribution of the extracted features. As shown in Fig. 5, raw motor-side measurements such as the angular velocity $\dot{\varphi}_{\mathrm{M}}$ and motor torque M_{M} exhibit substantial overlap across different fault conditions. In contrast, the EKF-estimated parameters form well-separated clusters, enabling the classifier to establish clear decision boundaries.

To quantitatively evaluate feature separability, the Fisher Discriminant Ratio (FDR) [28] was computed for all extracted features. The FDR measures the ratio between the distance of the class means and the corresponding within-class variance, where higher values indicate better class separability. The proposed physics-informed feature set achieved a significantly higher average FDR of 244.18 compared to 17.79 for the motor-side measurement-based features.

2) Data Efficiency: To evaluate this aspect, the available training windows are progressively reduced while keeping the test dataset unchanged.

For each training data fraction, windows are randomly sampled from every healthy and fault class to maintain balanced

Fig. 6. Confusion matrices for the data-driven and proposed models averaged over 20 realizations.

training sets. Since both the window selection and the classifier initialization are stochastic, the experiment is repeated over 60 independent realizations to capture the resulting variability in performance. The resulting data-efficiency behavior is shown in Fig. 7.

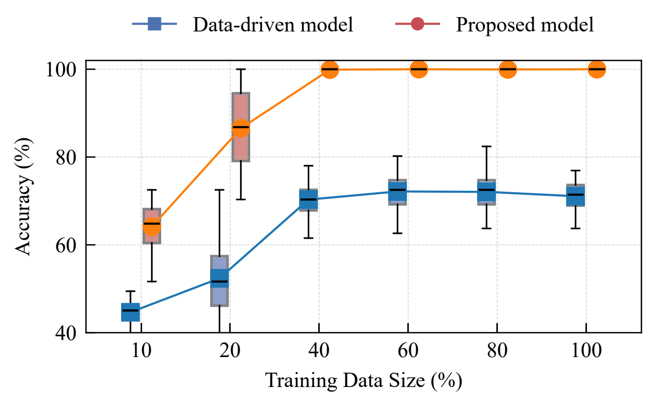

Fig. 7. Classification accuracy as a function of available training data. Box plots represent 60 independent realizations.

Even in the most extreme low-data (20%) scenario, the proposed model significantly outperforms the data-driven model even when the latter is trained on the full dataset. This behavior indicates that physics-informed features substantially reduce the statistical complexity of the classification task. The data-driven model shows a strong degradation in low-data regimes, dropping below 50% accuracy when only 20% of the training data is available. In contrast, the proposed approach maintains stable performance with minimal variance down to 40% training data. Since the classifier operates on low-dimensional physically meaningful parameter descriptors rather than highly nonlinear raw signal mappings, substantially fewer training samples are required.

3) Cross-Condition Generalization: The models are trained at a single angular velocity magnitude and evaluated across all velocity levels. For each experiment, training is performed using data from a single velocity magnitude, while testing is conducted across all four operating speeds. Table I summarizes the resulting classification performance.

TABLE I
CROSS-CONDITION GENERALIZATION PERFORMANCE

Train Speed (rad/s)	Test Speeds	Data-driven (%)	Proposed (%)
± 34.91	All	31.52	74.21
± 104.72	All	33.84	69.43
± 174.53	All	55.25	84.16
± 244.35	All	44.89	75.27
Overall		**41.38**	**75.77**

The data-driven model exhibits strong sensitivity to operating velocity, as motor-side measurement features scale directly with excitation amplitude. In contrast, the proposed approach substantially improved generalization performance across different operating speeds. This behavior can be explained by the nature of the extracted features. Physical parameters such as inertia and friction are largely independent of the operating velocity, whereas signal-based features scale directly with the excitation amplitude.

VI. CONCLUSIONS

This paper presents a physics-informed machine learning workflow for condition monitoring of industrial drive systems with elastic transmissions. By utilizing an EKF within the PiFG workflow, the approach extracts physically meaningful parameters only from motor-side measurements. These parameters are directly related to the underlying mechanical degradation, enabling the approach to overcome the major limitations of purely data-driven methods.

Furthermore, the proposed physics-informed features yield a substantially higher average FDR, ensuring distinct class separability. Consequently, a marked improvement in fault classification performance is achieved compared to a purely data-driven baseline. In addition, the proposed approach improves robustness under limited training data and increases cross-condition generalization accuracy from 41.38% to 75.77%. Since the estimated physical parameters are largely independent of the excitation velocity, the trained classifier can be transferred more reliably across different operating regimes.

Overall, the results confirm the high potential of feature-level integration of physical knowledge in machine learning workflows for robust, transferable, and data-efficient industrial condition monitoring.

REFERENCES

[1] R. R. Kumar, M. Andriollo, G. Cirrincione, M. Cirrincione, and A. Tortella, “A comprehensive review of conventional and intelligence-based approaches for the fault diagnosis and condition monitoring of induction motors,” *Energies*, vol. 15, no. 23, p. 8938, Nov. 2022.

[2] D. An, J. H. Choi, and N. H. Kim, “Options for prognostics methods: A review of data-driven and physics-based prognostics,” *AIAA Journal*, vol. 51, no. 12, pp. 1940–1950, 2013.

[3] Y. Lei, B. Yang, X. Shen, and F. Jia, “Machinery health prognostics: A systematic review,” *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, May 2018.

[4] W. Zhang, G. Peng, C. Sun, and Y. Chen, “A deep learning-based approach for bearing fault diagnosis with good anti-noise and domain adaptation ability,” *IEEE Trans. Ind. Electron.*, vol. 64, no. 7, pp. 5474–5485, Jul. 2017.

[5] D. Latil, R. Houe Ngouna, K. Medjaher, and S. Lhuisset, “Robust data-driven fault diagnostics for rotating machinery operating under varying working conditions,” in *Proc. 9th Int. Conf. Control, Decision and Information Technologies (CoDIT)*, Rome, Italy, Jul. 2023, pp. 1912–1917.

[6] R. Yan, R. X. Gao, and X. Chen, “Wavelets for fault diagnosis of rotary machines: A review,” *Mech. Syst. Signal Process.*, vol. 42, no. 1–2, pp. 1–18, Jan. 2014.

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

[8] J. Chen and R. J. Patton, *Robust model-based fault diagnosis for dynamic systems*. Boston, MA, USA: Springer, 1999.

[9] H. Dalimarta, et al., *Design and Implementation of a Robust 6-DOF Quadrotor Controller Based on Kalman Filter for Position Control*, Journal of Dynamic Systems, Measurement, and Control, 2024.

[10] D. Simon, *Optimal state estimation: Kalman, H-infinity, and nonlinear approaches*. Hoboken, NJ, USA: Wiley, 2006.

[11] M. S. Grewal and A. P. Andrews, *Kalman filtering: Theory and practice using MATLAB*, 4th ed. Hoboken, NJ, USA: Wiley, 2015.

[12] R. Isermann, *Fault-diagnosis systems: An introduction from fault detection to fault tolerance*. Berlin, Germany: Springer, 2006.

[13] L. Ljung, *System identification: Theory for the user*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[14] S. X. Ding, ”Model-based fault diagnosis techniques: design schemes, algorithms, and applications,” Springer Science and Business Media, 2008.

[15] Z. Gao, C. Cecati, and S. X. Ding, ”Survey on fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches,” IEEE Trans. Ind. Electron., vol. 62, no. 6, pp. 3757–3767, Jun. 2015.

[16] S. Zhang et al., ”A review on fault diagnosis of rotating machinery,” IEEE Access, vol. 7, pp. 16537–16557, 2019.

[17] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nature Rev. Phys.*, vol. 3, pp. 422–440, 2021.

[18] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems,” *J. Comput. Phys.*, vol. 378, pp. 686–707, Feb. 2019.

[19] M. Euldji, et al., *Optimizing condition monitoring of ball bearings: An integrated approach using decision tree and extreme learning machine for effective decision-making*, Open Physics, vol. 21, no. 1, Jun. 2023. doi: 10.1515/phys-2022-0239.

[20] Antonaggio, F. and et al., “Understanding Physics-Informed Neural Networks: Techniques, Applications, Trends, and Challenges,” *Machine Learning and Knowledge Extraction*, vol. 6, pp. 1780–1810, 2024.

[21] Zhou, Y. and et al., “Fundamental flaws of physics-informed neural networks and explainability methods in engineering systems,” *Reliability Engineering & System Safety*, vol. 254, pp. 110609, 2025.

[22] A. H. Jafari, R. Dhaouadi, and A. Jhemi, “Nonlinear friction estimation in elastic drive systems using a dynamic neural network-based observer,” *J. Adv. Comput. Intell. Intell. Informat.*, vol. 17, no. 4, pp. 637–648, 2013.

[23] M. A. Chao, C. Kulkarni, K. Goebel, and O. Fink, ”Fusing physics-based and deep learning models for prognostics,” Reliability Engineering and System Safety, vol. 217, p. 107961, 2022.

[24] S. Ullah, M. F. Siddique, and J.-M. Kim, “Multi-sensor observer-based residual learning with Auto-Permutation Feature Importance for fault diagnosis of multistage centrifugal pumps under variable pressures,” *Sci. Rep.*, vol. 15, no. 45735, 2025. [Online].

[25] R. O. B. Zarouala, C. Vivas, J. A. Acosta, and L. El Bakkali, “On singular perturbations of flexible and variable-speed wind turbines,” in *Proc. 11th Int. Conf. Informatics Control Autom. Robot. (ICINCO)*, vol. 2, Vienna, Austria, 2014, pp. 568–575.

[26] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: John Wiley & Sons, 2006.

[27] Y. Lei, *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Machines*. Berlin, Germany: Springer, 2016.

[28] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.