Measurement 270 (2026) 120862

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Cross-correlation heat-maps synthesis with Walsh–Fourier transformation in articulated robotic systems diagnostics☆

Dong Wu [a], Yazheng Wang [b], Peng Chen [b,*], Chunhua Guo [c], Renpu Li [d], Ge Xin [e], Changbo He [f]

[a] Department of Information Engineering, Changzhou Vocational Institute of Industry Technology, Changzhou, 213164, Jiangsu, PR China
[b] College of Engineering, Shantou University, Shantou, 515063, Guangdong, PR China
[c] Basic teaching department, Changzhou Vocational Institute of Industry Technology, Changzhou, 213164, Jiangsu, PR China
[d] School of Economics and Management, Changzhou Vocational Institute of Engineering, Changzhou, 213164, Jiangsu, PR China
[e] School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, PR China
[f] College of Electrical Engineering and Automation, Anhui University, Hefei, 230601, PR China

ARTICLE INFO

Keywords:
Articulated robotic manipulator
Fault diagnosis
Walsh transform
Signal processing
Condition monitoring

ABSTRACT

Harmonic drives, critical actuators in articulated robotic manipulators, are strain-wave gear transmissions comprising a wave generator, a thin-walled flexspline, and a rigid circular spline, and they operate under nonstationary dynamics and variable speed/load conditions that confound vibration-based diagnostics. Existing methods often become unstable across speed variations; moreover, practical deployments lack a priori characteristic fault frequencies (wave-generator, flexspline, tooth-mesh) and complete operating context, which undermines robustness and reliability. We introduce Cross-correlation Heatmap Synthesis via Walsh–Fourier Transform (CHS-WFT), which models cross-signal structure and adaptively segments time series to accommodate unknown regimes while suppressing nonstationary interference. The approach hinges on speed-aware partitioning with frame-level decomposition, together with a joint Walsh–Fourier cross-correlation representation that fuses Fourier's sensitivity to harmonic content with Walsh's responsiveness to abrupt, piecewise transients, thereby revealing inter-signal dependencies obscured in Fourier-only analyses. On this representation, multi-scale energy entropy quantifies irregularity and complexity across scales, yielding compact, discriminative features that separate healthy and faulty states despite large speed fluctuations. Consequently, across dynamic operating regimes, CHS-WFT increases discriminability, stabilizes detection with reduced dependence on labeled data and detailed operating context, and enables early-warning fault detection to support condition-based maintenance in articulated robotic systems.

1. Introduction

Harmonic drives are strain-wave gear transmissions that serve as key actuators in articulated robotic manipulators and comprise a wave generator, a thin-walled flexspline, and a rigid circular spline [1,2]. Owing to their large reduction ratio, low mass, and high transmission accuracy, their condition directly affects robot precision and process performance. Torque is transmitted as the wave generator, via an elliptical flexible bearing, elastically deforms the flexspline so its teeth engage the circular spline. However, this mechanism concentrates contact stress and imposes repeated strain, which leads to abrasion and pitting of the bearing and to wear, microcracking, or fracture of the flexspline teeth, thereby increasing backlash, reducing efficiency, and shortening service life. Once failure occurs, prompt fault diagnosis and condition monitoring are required to prevent secondary damage, minimize unplanned downtime, and restore safe operation.

☆ This work was supported of the National Natural Science Foundation of China (Grant 52105111), the Guangdong Basic and Applied Basic Research Foundation (Grant 2025A1515012256), the Scientific Research Foundation of Changzhou Vocational Institute of Engineering (No. 11130900120001), the Scientific Research Foundation of Changzhou Vocational Institute of Industry Technology (No. BS202213101002), the Professional Virtual Teaching and Research Room Construction Practices and Exploration Studies (No. XHYBLX2023004), the Changzhou Intelligent Connected Vehicle Driverless Driving and Network Security Technology Key Laboratory (No. CM2024007), and the Research on the Construction of Artificial Intelligence Technology Application Specialties and the Training Model for Industrial Talents in Changzhou Science and Education Town (No. CDGZ202526).

* Corresponding author.
E-mail addresses: pengchen@alu.uestc.edu.cn, dr.pengchen@foxmail.com (P. Chen).

https://doi.org/10.1016/j.measurement.2026.120862
Received 26 September 2025; Received in revised form 10 February 2026; Accepted 13 February 2026
Available online 14 February 2026

In robotic systems, machine learning and deep learning technologies have revolutionized fault diagnosis approaches for harmonic drives by enabling sophisticated data-driven detection across multiple sensing modalities and diverse operating conditions. The field has witnessed significant advancements in representation learning, graph-based modeling, generative methods, and automated architecture optimization [1,3–5]. For instance, Yang et al. [6] introduced an innovative approach that transforms multiple vibration signals into Symmetrized Dot Pattern images and utilizes a ConvNeXt model with Transformer training procedures. This groundbreaking integration of SDP with ConvNeXt for harmonic drive diagnosis demonstrates superior performance, interpretability, and practicality compared to traditional ConvNets and Transformers [7,8] under both stable and variable conditions. Chen et al. [9] further advanced the field by creating a self-supervised framework that incorporates physics-informed prior knowledge across time, frequency, and time–frequency domains into a multiscale convolutional autoencoder. This approach significantly reduces the reliance on labeled data while enhancing generalization capabilities in few-shot industrial applications. Addressing networked applications, Bilal et al. [10] developed an advanced Internet of Robotic Things architecture that combines transfer learning with a hybrid 1D-MCNN-RNN model. This comprehensive approach achieves remarkable 99.03% accuracy in cross-domain diagnosis of manipulator joints by incorporating matrix kernels, recurrent modeling of failure patterns, and innovative timestamp mapping for sensor alignment. However, these advanced methods face several critical challenges in practical applications. They typically require extensive high-quality labeled datasets that are expensive and challenging to obtain. The scarcity of well-curated datasets and potential issues with measurement accuracy present significant obstacles. Additional challenges include limited model transparency, interpretability, and complications arising from domain shifts and sensor variations. Furthermore, factors such as class imbalance and high computational demands hinder the successful deployment and certification of these systems in real-world robotic environments.

In modern robotics, signal analysis techniques for condition monitoring and fault diagnosis of harmonic drives have evolved significantly, incorporating advanced signal processing methodologies and multi-modal data fusion approaches to extract meaningful diagnostic information from complex mechanical systems [11–16]. Against this technological backdrop, Zhang et al. [4] proposed a robust defect diagnosis framework based on motor current analysis, utilizing advanced signal processing paradigms to successfully decouple fault signatures from environmental interference. The integration of sophisticated vibration analysis, acoustic emission monitoring, and dynamic modeling has enabled more precise and reliable fault detection mechanisms across diverse operational conditions. To address the non-stationary nature of robotic operations, Wang et al. [17] investigated the dynamic modeling of flexible thin-walled bearings, shifting the diagnostic paradigm from traditional stationary spectral analysis to advanced time–frequency decomposition strategies. For instance, He et al. [18] developed a comprehensive yet lightweight monitoring framework that innovatively combines a multichannel vibration imaging technique (Wavetrizorn) with a lossy reconstruction algorithm (WPCAE) to enhance fault diagnosis capabilities under challenging noise conditions. Their framework introduces a novel canonical correlation feature to establish precise correlations between fault patterns and grinding quality through vibration-to-force signal regression, thereby enabling accurate quality assessment using solely vibration measurements. Furthermore, Wang et al. [19] constructed a trackable multi-domain collaborative generative adversarial network, demonstrating that feature extraction spanning time, frequency, and time–frequency domains is indispensable for characterizing the transient evolution of mechanical faults. Contributing to the advancement of multi-sensor fusion techniques, Zhang et al. [20] proposed an innovative adaptive information fusion methodology based on the Newton–Raphson Optimizer that constructs a comprehensive Health Indicator for harmonic reducers. Their approach uniquely combines acoustic emission and micro-vibration signals while implementing a Dynamic Monotonicity Strength Index and Composite Evaluation Index to optimize feature selection and fusion processes, resulting in more robust degradation assessment capabilities. Complementing these extraction methods, effective de-noising strategies are essential to mitigate the masking effects of heavy background noise prior to feature learning. Li et al. [21] introduced a multiscale symbolic fuzzy entropy algorithm that functions as an effective de-noising filter, stabilizing feature distributions against heavy background noise under varying operational loads. However, these signal analysis methodologies face substantial challenges particularly in non-stationary operating environments where frequent speed variations and operational transitions occur. The time-varying dynamics inherent in such conditions often mask important non-stationary fault signatures while simultaneously compromising the stability and reliability of learned signal representations, making accurate fault diagnosis significantly more challenging. Additionally, these methods must contend with issues such as sensor noise, environmental interference, signal coupling effects, and the complex interactions between different mechanical components that can obscure fault-related features in the measured signals.

Motivated by the scarcity of high-quality labeled data for machine learning and deep learning based diagnostic methods, and by the persistent challenges of vibration signal analysis under non-stationary dynamics and variable operating conditions, current approaches remain vulnerable to poor generalization and unstable performance. Moreover, practical deployments of harmonic drives in articulated robotic systems are often hindered by the lack of a priori knowledge regarding characteristic fault frequencies, such as those from the wave generator, flexspline, and tooth mesh. This issue is compounded by incomplete contextual information about operating conditions, which collectively degrades the robustness of fault detection and diagnosis. To overcome these constraints, we propose Cross-correlation Heatmap Synthesis via Walsh–Fourier Transform (CHS-WFT), a framework specifically engineered for articulated robotic systems under dynamic speed variations that leverages cross-correlation structure in the Walsh–Fourier domain to enhance discriminability under non-stationary regimes. This approach integrates adaptive signal partitioning with multiscale feature extraction to improve robustness and sensitivity. It estimates instantaneous speed to segment nonstationary motion into quasi-steady regimes, then applies frame-level decomposition within each segment. Cross-correlation heatmaps are synthesized in a joint Walsh–Fourier domain: Fourier captures harmonic structure while Walsh encodes abrupt, piecewise transients, revealing inter-signal relationships missed by Fourier alone. Multi-scale energy entropy then quantifies irregularity and complexity across scales, yielding compact, discriminative features that separate healthy and faulty states. Together, these steps enable CHS-WFT to sustain diagnostic accuracy under large speed variations, strengthen condition monitoring, and support early-warning fault detection in articulated robotic systems with harmonic drives.

The principal contributions of this research are threefold:

1. Adaptive speed-aware segmentation: Estimates instantaneous rotational speed and partitions non-stationary motion into quasi-steady regimes, followed by frame-level decomposition to enable locally stationary analysis and consistent comparison across variable operating conditions.
2. Joint Walsh–Fourier cross-correlation: Synthesizes cross-correlation heatmaps in the combined Walsh–Fourier domain, where Fourier captures harmonic structure (tooth-mesh, flexspline, wave-generator) and Walsh encodes abrupt and piecewise transients, revealing inter-signal dependencies obscured in Fourier-only analyses.
3. Multiscale entropy-based diagnosis: Computes multi-scale energy entropy over the heatmaps to quantify deviations from nominal behavior, yielding compact, discriminative features robust to speed fluctuations for reliable fault detection and early warning in harmonic drives.

To provide a coherent and logically progressive narrative, this study is organized as follows. Section 2 presents the theoretical preliminaries, formally defining the mathematical properties of the Walsh–Fourier transform and justifying the signal representation strategies utilized in the subsequent diagnostic pipeline. Building on this foundation, Section 3 introduces the CHS-WFT methodology. It first formalizes the adaptive rotational-speed partitioning and the cross-correlation heatmap synthesis pipeline. Subsequently, the section details a complementary feature-based diagnosis module that utilizes multi-scale energy entropy to enhance pattern separability and interpretability. Subsequently, Section 4 presents a rigorously designed experimental evaluation that enables controlled and comparative analysis under variable operating conditions, thereby assessing robustness, sensitivity, and practical relevance. Finally, Section 5 concludes by distilling the main findings, reflecting on methodological limitations, and outlining directions for future research and potential applications.

2. Theoretical framework for hybrid orthogonal analysis

The diagnosis of articulated robotic manipulators requires robust feature extraction capable of handling non-stationary dynamics. This section provides the theoretical justification for the proposed framework, strictly distinguishing the morphological advantages of the Walsh transform [22] from conventional time–frequency methods and mathematically deriving the mechanism by which the cross-correlation heatmap reveals the inter-signal dependencies between impulsive and stationary harmonics.

2.1. Analysis of basis mismatch and spectral convergence

The fundamental limitation of conventional Time–Frequency Analysis (TFA) methods, such as the Short-Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT), lies in the "basis mismatch" between their analyzing functions and the fault signatures of harmonic drives. Faults in these mechanisms (e.g., flexspline tooth breakage or bearing pitting) manifest as sharp, rectangular-like impulsive shock sequences. STFT utilizes sinusoidal basis functions $e^{j\omega t}$, which are globally smooth. When approximating a sharp discontinuity (a step function or impulse), the Fourier series converges slowly ($O(k^{-1})$), resulting in the Gibbs phenomenon where energy leaks into high-frequency sidebands, smearing the fault feature in the time–frequency plane [23]. Similarly, while CWT improves resolution using variable-width windows, standard mother wavelets (e.g., Morlet) are typically decaying oscillatory functions. This mismatch forces the transform to use varying scales to approximate a single rectangular edge, leading to "ringing" artifacts and dispersed energy representation.

In contrast, the Walsh Transform employs an orthogonal set of rectangular basis functions (Walsh functions, $\mathrm{wal}(k,t)$) taking values in $\{+1,-1\}$. This creates a morphological match with the piecewise-constant nature of mechanical impact shocks. Mathematically, the Approximation Error ϵ_M for a fault transient $x_{imp}(t)$ using the first M basis coefficients is defined as:

$$\epsilon_M(\mathcal{B}) = \left\| x_{imp}(t) - \sum_{k=1}^{M} c_k^{\mathcal{B}} \phi_k^{\mathcal{B}}(t) \right\|^2 \tag{1}$$

For smooth bases ($\mathcal{B}_{STFT}, \mathcal{B}_{CWT}$), ϵ_M decays asymptotically. However, for the Walsh basis ($\mathcal{B}_{Walsh}$), since a rectangular pulse can be exactly represented by a finite linear combination of Walsh functions, the error ϵ_M drops to zero for a finite M_0. This property, known as sequency sparsity, proves that the Walsh transform concentrates fault energy into fewer coefficients than competing TFA methods, thereby providing a more discriminative and noise-robust feature set for impulsive fault detection.

2.2. Modulation decoupling mechanism via cross-correlation synthesis

The discriminative power of the proposed Cross-correlation Heatmap Synthesis (CHS) stems from its ability to reveal the non-linear inter-signal dependencies that single-domain spectra obscure. A harmonic drive fault is fundamentally a modulation phenomenon where the stationary rotational dynamics interact with transient structural impacts. We model the fault signal $x_{fault}(t)$ as a multiplicative coupling between a stationary carrier (rotational harmonic) and a transient modulator (impulsive train):

$$x_{fault}(t) = \underbrace{\cos(2\pi f_{rot} t)}_{\text{Stationary Harmonic}} \cdot \underbrace{\left[1 + \sum_i \alpha_i \mathrm{wal}(k_i, t)\right]}_{\text{Impulsive Modulator}} \tag{2}$$

where f_{rot} is the rotation frequency and $\mathrm{wal}(k_i,t)$ represents the ith Walsh mode of the impact train.

If analyzed by Fourier Transform alone, this multiplication becomes a convolution $X(f) = \mathcal{F}\{Carrier\} * \mathcal{F}\{Modulator\}$, which disperses the modulation energy α_i into complex sidebands $f_{rot} \pm f_i$, making it difficult to distinguish from noise. The proposed CHS framework explicitly recovers this dependency by computing the cross-correlation between the Fourier-reconstructed carrier and the Walsh-projected modulator. The heatmap intensity $S_{CHS}(\omega,\kappa)$ is derived as:

$$S_{CHS}(\omega,\kappa) = \left| \int \left(x_{fault}(t) \cdot e^{-j\omega t} \right) \cdot \mathrm{wal}(\kappa, t)\, dt \right| \tag{3}$$

Substituting the fault model into this equation, the integral serves as a matched filter for the modulation interaction. When the scanning coordinates align with the physical fault parameters ($\omega = 2\pi f_{rot}$ and $\kappa = k_i$), the orthogonal properties of the bases filter out uncorrelated noise, and the integral maximizes to proportional to the coupling strength α_i. Unlike standard spectra that display energy in isolation, this heatmap mathematically isolates the specific "locking" between the periodic rotation and the impulsive impacts. Consequently, the resulting high-intensity clusters in the heatmap are direct representations of the inter-signal dependencies, making the feature highly discriminative compared to smeared single-domain representations.

3. Cross-correlation heatmap synthesis via Walsh–Fourier transform

This section elucidates the methodology of Cross-correlation Heatmap Synthesis via Walsh–Fourier Transform (CHS–WFT) as applied to the diagnostics of articulated robotic systems under conditions of dynamic speed variations. As illustrated in Fig. 1, the proposed framework unfolds through four distinct yet interconnected stages. Initially, as detailed in Section 3.1, the process involves adaptive rotational speed partitioning, which segments motion signals into stable, steady stages for further analysis. Subsequently, Section 3.2 addresses the synthesis of cross-correlation heatmaps by leveraging the combined Fourier and Walsh domains to capture intricate signal relationships. This is followed by the procedures outlined in Section 3.3, which focus on time-series frame decomposition and heatmap construction, effectively dividing steady signals into discrete frames to facilitate localized analysis. Finally, Section 3.4 introduces a feature-based fault diagnosis mechanism that employs multi-scale energy entropy to quantitatively distinguish between normal and aberrant states. This approach leverages the multi-scale decomposition of signal energy distribution, where the entropy metric captures the irregularity and complexity of energy variations across different scales—effectively encoding the subtle differences between healthy and faulty states in a compact, discriminative feature. Through this systematic and integrated approach, the methodology ensures a robust differentiation between healthy and faulty conditions, thereby enhancing the efficacy of fault diagnosis in complex robotic systems.

3.1. Adaptive rotational speed partitioning

This section elaborates on the analytical process following wavelet denoising of the original signal and the subsequent segmentation into static and motion segments using short-time energy analysis.

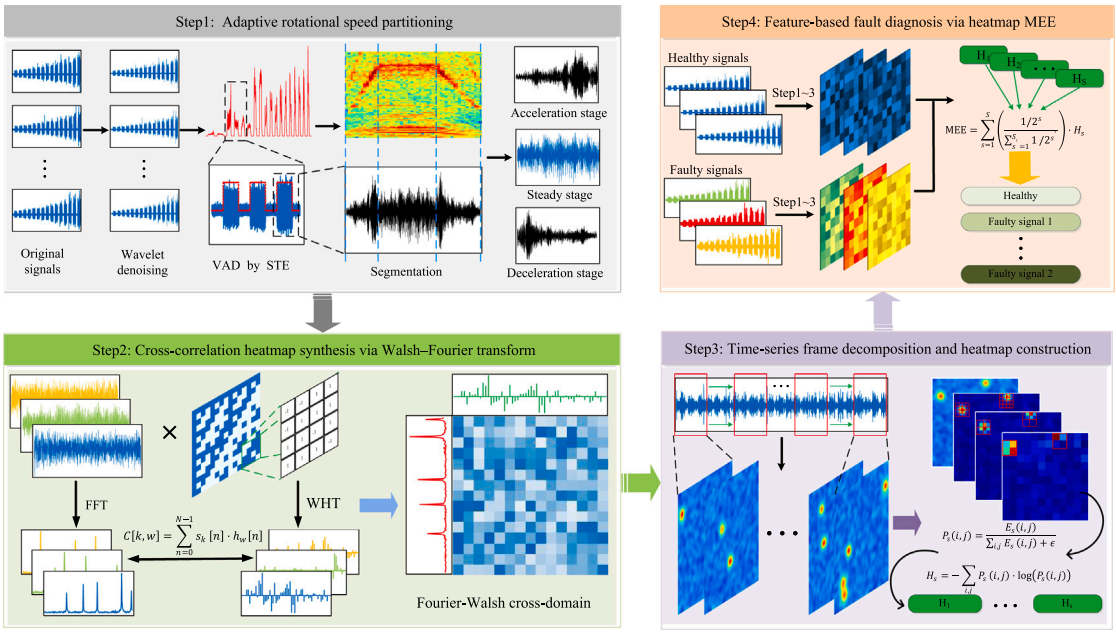

Fig. 1. Architectural overview of CHS–WFT: (1) adaptive rotational speed partitioning; (2) cross-correlation heatmap synthesis via Walsh–Fourier transform; (3) time-series frame decomposition and heatmap construction; and (4) feature-based diagnosis using multi-scale energy entropy.

Prior to segmentation, the raw vibration signal $x_{\text{raw}}[n]$ is subjected to Discrete Wavelet Transform (DWT) [24,25] for noise suppression. The signal is decomposed into approximation coefficients $c_{j,k}$ and detail coefficients $d_{j,k}$ at level J. A soft-thresholding operator is applied to the detail coefficients to remove Gaussian noise components:

$$\hat{d}_{j,k} = \begin{cases} \operatorname{sgn}(d_{j,k})(|d_{j,k}| - \lambda) & |d_{j,k}| \geq \lambda \\ 0 & |d_{j,k}| < \lambda \end{cases} \tag{4}$$

where λ is the universal threshold derived from the noise variance estimation. The denoised signal $x_{\text{den}}[n]$ is reconstructed using the inverse DWT.

Subsequently, to distinguish the active motion phase from static idle periods, the Short-Time Energy (STE) [26] of the denoised signal is computed as:

$$E_{\text{STE}}(m) = \sum_{i=0}^{W_E-1} \left| x_{\text{den}}[mR_E + i] \right|^2 \tag{5}$$

where W_E and R_E denote the window length and hop size for the energy calculation, respectively. Frames where $E_{\text{STE}}(m)$ exceeds an adaptive threshold T_E are concatenated to form the effective motion segment signal $x_o[n]$.

It is important to note that the signal within this motion segment often exhibits non-stationary behavior. This is primarily caused by the robotic arm's motor undergoing an acceleration phase, during which it must overcome significant resistance. Such dynamics result in considerable fluctuations in the acquired vibration signals, which, in turn, pose challenges for the accurate extraction of fault features. Consequently, to address this variability and enhance diagnostic precision, the motion segment is further categorized into three distinct stages – namely acceleration, steady, and deceleration – with the analytical emphasis predominantly placed on the steady stage, where signal stability is more conducive to feature extraction.

To facilitate this detailed segmentation, a relatively small temporal window is initially utilized to apply the Short-Time Fourier Transform (STFT) to the signal within the motion segment. For the input motion segment signal denoted as $x_o[n]$, the STFT is computed as follows:

$$\mathrm{S}_t(m,u) = \sum_{n=0}^{N_w-1} x_o[n + mR_s]\, w[n]\, e^{-j2\pi un/N_{\text{fft}}}, \quad u = 0, \ldots, N_{\text{fft}} - 1 \tag{6}$$

where m is the frame index, u is the frequency bin index, R_s is the frame shift, N_{fft} is the DFT length, and $w[n]$ is the window function of length N_w.

The power spectral density (PSD) is

$$P(m,u) = \left| \mathrm{S}_t(m,u) \right|^2, \tag{7}$$

and the corresponding frequency axis is $f_u = u\,\frac{f_s}{N_{\text{fft}}}$, where f_s is the sampling rate. The dominant frequency per frame is

$$f_{\text{dom}}(m) = \arg\max_u \; P(m,u) \cdot \frac{f_s}{N_{\text{fft}}}. \tag{8}$$

The rate of frequency change is

$$\Delta f(m) = \left| f_{\text{dom}}(m) - f_{\text{dom}}(m-1) \right|. \tag{9}$$

To delineate stages, thresholds $T_1 < T_2$ are set, and each frame is classified as

$$\text{Stage}(m) = \begin{cases} +1, & \Delta f(m) > T_2 \;\wedge\; f_{\text{dom}}(m) > f_{\text{dom}}(m-1), \\ 0, & \Delta f(m) < T_1 \;\wedge\; \left| f_{\text{dom}}(m) - \bar{f}(m) \right| < \delta_f, \\ -1, & \Delta f(m) > T_2 \;\wedge\; f_{\text{dom}}(m) < f_{\text{dom}}(m-1), \end{cases} \tag{10}$$

where $\bar{f}(m)$ is the local mean of f_{dom} over a sliding window, and δ_f is a stability tolerance. Frequency stability is quantified via the coefficient of variation

$$\text{CV}(m) = \frac{\sigma\big(f_{\text{dom}}(m-r:m+r)\big)}{\mu\big(f_{\text{dom}}(m-r:m+r)\big)}, \tag{11}$$

where r is the half-width (in frames) of the sliding window, and $\sigma(\cdot)$, $\mu(\cdot)$ denote standard deviation and mean, respectively.

The steady stage is precisely identified by its start and end frame indices (m_s, m_e) using a stability criterion:

$$\text{Find the smallest } m_s \text{ such that } \text{CV}(m') < \text{CV}_{\text{th}} \quad \forall m' \in [m_s,\, m_s + N_c - 1],$$

$$\text{and the largest } m_e \text{ such that } \text{CV}(m') < \text{CV}_{\text{th}} \quad \forall m' \in [m_e - N_c + 1,\, m_e],$$

where CV_{th} is a threshold and N_c is the minimum run-length of stable frames. This yields a time–frequency map (see "segmentation" in Fig. 1) where the dominant frequency increases (acceleration), remains flat (steady), and decreases (deceleration), mirroring the motion profile.

Mapping frames to samples, the steady segment is

$$n_s = m_s R_s, \quad n_e = m_e R_s + N_w - 1, \quad x_s[n] = x_o[n], \quad n = n_s, \dots, n_e, \tag{12}$$

with steady length $N_s = n_e - n_s + 1$. The steady signal $x_s[n]$ is used for subsequent feature extraction and diagnosis.

3.2. Cross-correlation heatmap synthesis via Walsh–Fourier transform

This section details the construction of a spectral correlation heatmap that elucidates the interaction between sinusoidal (Fourier) and square-wave (Walsh) components for diagnostic analysis. The algorithm computes zero-lag cross-correlations between reconstructed Fourier sinusoids and coefficient-weighted Walsh functions, thus retaining the original amplitude information through coefficient weighting and directly quantifying their interaction.

For a steady signal segment $x_s[n]$ of length N, the FFT is

$$F[k] = \sum_{n=0}^{N-1} x_s[n]\, e^{-j2\pi kn/N}, \quad k = 0, \dots, N-1, \tag{13}$$

with amplitude spectrum for positive frequencies

$$A[k] = \left| \frac{F[k]}{N} \right|, \quad k = 1, \dots, \frac{N}{2}, \tag{14}$$

and frequency axis

$$f[k] = k \frac{f_s}{N}. \tag{15}$$

In parallel, the Walsh–Hadamard transform (WHT) uses the $N \times N$ Hadamard matrix $\mathbf{H} \in \{\pm 1\}^{N \times N}$ with orthogonality. The normalized WHT is

$$W[w] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} H[w, n]\, x_s[n], \quad w = 0, \dots, N-1, \tag{16}$$

where $H[w, n]$ represents the element of the Hadamard matrix corresponding to the wth Walsh function at time index n.

The magnitude is given by

$$B[w] = |W[w]|, \quad w = 1, \dots, N-1, \tag{17}$$

excluding the DC term when focusing on transients.

To isolate the periodic oscillations at bin k, reconstruct the sinusoid

$$s_k[n] = \mathrm{Re}\left(\frac{F[k]}{N}\, e^{j2\pi f[k]\, n/f_s} \right), \quad n = 0, \dots, N-1, \tag{18}$$

where $\mathrm{Re}(\cdot)$ denotes the real part operator, ensuring the reconstruction of the specific time-domain waveform corresponding to frequency $f[k]$.

For each Walsh index w, define a coefficient-weighted Walsh function

$$h_w[n] = W[w]\, H[w, n], \quad n = 0, \dots, N-1. \tag{19}$$

where $h_w[n]$ is the weighted rectangular basis function scaled by its spectral contribution $W[w]$.

The zero-lag cross-correlation (inner product) between $s_k[n]$ and $h_w[n]$ is

$$C[k, w] = \sum_{n=0}^{N-1} s_k[n]\, h_w[n]. \tag{20}$$

The matrix $C[k, w]$ constitutes the heatmap, inherently weighted by the Fourier and Walsh coefficients. It focuses on interactions between physically significant steady-state harmonics and transient Walsh components, forming a robust basis for subsequent fault feature extraction.

3.3. Time-series frame decomposition and heatmap construction

This section details time-series decomposition and local heatmap construction for the steady signal $x_s[n]$ from Section 3.1. To preserve subtle time variations, the signal is segmented into shorter frames that are processed independently, aligning with the multi-domain fusion in Section 3.2.

Let $x_s[n]$ have length N_s. Choose a frame length $L = 2^q$ (for FFT/WHT efficiency) and a frame shift S_f (typically $S_f = L/2$ for 50% overlap). The number of frames is

$$P = \left\lfloor \frac{N_s - L}{S_f} \right\rfloor + 1, \tag{21}$$

and the pth frame is

$$x_p[n] = x_s[(p-1)S_f + n], \quad n = 0, \dots, L-1, \quad p = 1, \dots, P. \tag{22}$$

Before multi-domain fusion, energy normalization removes amplitude-only differences while preserving morphology. The frame energy and normalized frame are

$$E_p = \sum_{n=0}^{L-1} x_p[n]^2, \tag{23}$$

$$\hat{x}_p[n] = \frac{x_p[n]}{\sqrt{E_p + \zeta}}, \quad n = 0, \dots, L-1, \tag{24}$$

where $\zeta > 0$ prevents division by zero and ensures $\sum_n \hat{x}_p[n]^2 \approx 1$.

For each frame $\hat{x}_p[n]$, compute FFT and WHT as in Section 3.2 with $N = L$: obtain $F_p[k]$, $A_p[k]$, $W_p[w]$, $B_p[w]$; reconstruct sinusoids $s_{p,k}[n]$ from $F_p[k]/L$; form weighted Walsh functions $h_{p,w}[n] = W_p[w]\, H[w, n]$; and compute the frame-specific cross-correlation heatmap

$$C_p[k, w] = \sum_{n=0}^{L-1} s_{p,k}[n]\, h_{p,w}[n]. \tag{25}$$

The set $\{C_1, \dots, C_P\}$ preserves local variations, avoids feature smearing, and provides a spatiotemporal view suitable for downstream diagnostic feature extraction.

3.4. Feature-based fault diagnosis via heatmap multi-scale energy entropy

This section delineates the application of Multi-scale Energy Entropy (MEE) as a statistical feature to facilitate fault diagnosis using frame-specific heatmaps $C_p[k, w]$ generated per Section 3.3. In particular, MEE quantifies the regularity of energy distribution across multiple scales in dynamic patterns [27,28], and, by design, it captures structural uniformity that systematically differs between healthy and faulty operational states. Consequently, this measure serves as an objective descriptor of spatial energy organization in the heatmap domain. This capability enables robust discrimination without relying on ad hoc thresholds or hand-crafted features.

For the pth heatmap $C_p \in \mathbb{R}^{K \times W}$ with $K = \frac{L}{2}$ and $W = L$, analyze scales $s = 1, \dots, S$. To this end, at each scale, partition C_p into non-overlapping $2^s \times 2^s$ blocks so as to aggregate localized energy in a multi-resolution manner that is consistent with the spatial support of heatmap structures. Define the block energy as

$$E_{p,s}(a, b) = \sum_{(k', w') \in \mathcal{B}(a,b,s)} C_p[k', w']^2, \tag{26}$$

where $\mathcal{B}(a, b, s)$ is the set of indices belonging to the (a, b)-th block at scale s, and a, b are the block coordinates along the frequency and sequency axes, respectively.

Normalize to probabilities

$$P_{p,s}(a, b) = \frac{E_{p,s}(a, b)}{\sum_{a,b} E_{p,s}(a, b) + \epsilon_m}, \tag{27}$$

with a small $\epsilon_m = 10^{-8}$ for numerical stability. The single-scale entropy is defined as:

$$H_{p,s} = -\sum_{a,b} P_{p,s}(a, b)\, \log\big(P_{p,s}(a, b)\big). \tag{28}$$

Aggregate across scales with weights $w_s = \frac{2^{-s}}{\sum_{s'=1}^{S} 2^{-s'}}$ to obtain

$$\mathrm{MEE}_p = \sum_{s=1}^{S} w_s \, H_{p,s}. \tag{29}$$

where w_s ensures that finer scales contribute more significantly to the final entropy score.

MEE reflects energy concentration: higher values indicate dispersed energy (flat heatmaps), while lower values indicate concentrated spikes. Accordingly, healthy systems typically yield structured yet broadly distributed correlations and thus higher MEE_p; by contrast, faulty systems exhibit transient, spike-dominated patterns and thus lower MEE_p. To form a sample-level metric,

$$\overline{\mathrm{MEE}} = \frac{1}{P} \sum_{p=1}^{P} \mathrm{MEE}_p. \tag{30}$$

In practice, healthy samples exhibit large $\overline{\mathrm{MEE}}$, whereas faulty samples yield small $\overline{\mathrm{MEE}}$, thereby providing a robust and interpretable discriminator for fault diagnosis in articulated robotic manipulators.

4. Experimental validation and comparative analysis

4.1. Experimental apparatus and data acquisition protocol for the articulated robotic manipulator

In this real-world experimental investigation, the configuration of the articulated robotic manipulator and the corresponding data acquisition protocols were deliberately engineered to maximize precision, repeatability, and reliability. In particular, the synchronous belt tension within the test rig was rigorously controlled and maintained between 70 to 80 N, since this parameter is critical for ensuring both drivetrain stability and positional accuracy during operation. Moreover, the sampling posture adopted by the manipulator was determined on the basis of accumulated shop-floor expertise and extensive practical experience, thereby aligning the configuration with representative industrial practice and operational constraints; this standardized posture is illustrated in Fig. 2 for reference. To capture high-fidelity vibration data, a Quatronix ULT2101-50 piezoelectric triaxial accelerometer was employed. Key performance specifications include a sensitivity of 100 mV/g, a measurement range of ±50 g, and a frequency response band of 0.5–5000 Hz. Additionally, the sensor features a resonant frequency of 50 kHz, a maximum shock resistance of 2000 g, and an operating temperature range of −40 to +120°C, ensuring robustness in harsh industrial environments. The optimal mounting position was determined through a preliminary RMS-based energy analysis across seven candidate process holes (labeled ①–⑦ in Fig. 2). A comparative analysis across the full operational speed range revealed that the total vibration energy (RMS_{total}) peaked at 2500 rpm for all candidate points; consequently, this speed was established as the reference condition for sensitivity evaluation. Results at this reference speed indicated that the topmost position (Point ④) exhibited the highest total vibration energy, thereby identifying it as the most sensitive location for characterizing the manipulator's dynamic response. With respect to sampling parameters, the data acquisition frequency was set to 7680 Hz to capture high-fidelity signals, and the sampling duration was chosen to match one complete operational cycle of the robotic arm – namely 80 seconds – so as to encompass both transient and steady-state behaviors.

During the in-situ measurement campaign, the operational dynamics of the articulated manipulator were systematically varied to span the full range of working rotational speeds in a controlled and repeatable manner. Starting from a designated initial position, the robotic arm paused for 1 s, proceeded to the endpoint at an initial speed of 1000 r/min, paused again for 1 s at the endpoint, and then returned to the starting position at the same speed, followed by an additional 1-second pause. This round-trip cycle was then repeated with the rotational speed increased by 500 r/min in each subsequent cycle – beginning at 1500 r/min – until the manipulator completed the protocol at its maximum designated speed and returned to the starting point. Such a progressive speed escalation not only ensures comprehensive coverage of the operational envelope but also facilitates a thorough assessment of speed-dependent phenomena and performance under varying conditions. It is important to note that the harmonic reducers utilized in this study were harvested from industrial manipulators undergoing natural degradation cycles, ensuring the fault signatures represent realistic failure modes rather than idealized artificial defects. The classification of severity levels is determined based on the cumulative operating hours of the components and maintenance records. Physically, the rationale for the observed impulsive vibrations stems from the specific failure mechanism of the harmonic drive: as the flexspline undergoes fatigue wear and tooth profile deviation, the meshing interface develops geometric irregularities. When the elliptical wave generator rotates, these irregularities cause momentary stiffness variations and repetitive mechanical impacts (shocks) during the tooth engagement process. These tribological interactions manifest as high-frequency impulsive vibrations that modulate the carrier signal, the intensity of which correlates directly with the degradation severity (Minor, Moderate, and Severe).

4.2. Multi-condition vibration acquisition and preliminary observations

In this case study, vibration signals from four distinct sets of articulated robotic manipulators were meticulously acquired using high-quality sensors to enable comparative analysis across health states. The sets comprise one baseline (healthy) sample and three faulty samples reflecting graded severities: minor, severe, and moderate. The selection of these four specific categories—Healthy, Minor, Moderate, and Severe—is grounded in industrial Condition-Based Maintenance (CBM) protocols. Notably, data representing a fifth state of "Catastrophic Failure"—characterized by intense physical jitter and audible abnormal noise, with vibration amplitudes orders of magnitude higher—was excluded from this study. Since such terminal states are easily detectable via simple thresholding or visual inspection, this research prioritizes the non-trivial diagnostic range where advanced signal processing is requisite. These four levels correspond to the critical decision-making thresholds in practical engineering: baseline operation (Healthy), initial degradation requiring monitoring (Minor), degradation requiring scheduled maintenance (Moderate), and critical degradation requiring urgent maintenance planning (Severe). Consequently, this classification task rigorously evaluates the model's sensitivity to subtle signal energy changes during the degradation process, serving as a validation for potential prognostic health management applications. Specifically, the minor-fault condition manifests as occasional abnormal noise during operation, although the auditory cues are subtle and thus easily overlooked without close monitoring. By contrast, the severe-fault condition is characterized by pronounced abnormal vibrations—especially at higher rotational speeds—often accompanied by noticeable mechanical instability, whereas the moderate-fault condition exhibits a combination of faint abnormal noise and minor vibrations that lies between the minor and severe cases.

To ensure comprehensive data capture, a triaxial sensor was employed to synchronously record vibration acceleration along the X, Y, and Z axes. For the present analysis, emphasis is placed on the X-axis acceleration, as it provides diagnostic information that is most directly aligned with the observed mechanical behaviors.

The time-domain representations of the acquired acceleration signals are illustrated in Fig. 3. Subpanels (a_1)~(d_1) depict the raw signals for the healthy, minor, severe, and moderate fault conditions, respectively. As observed, the raw waveforms exhibit significant noise interference, particularly during the static intervals, rendering it difficult to discern distinct fault characteristics solely through time-domain inspection. To mitigate measurement noise and enhance signal interpretability, we applied wavelet denoising to generate the refined

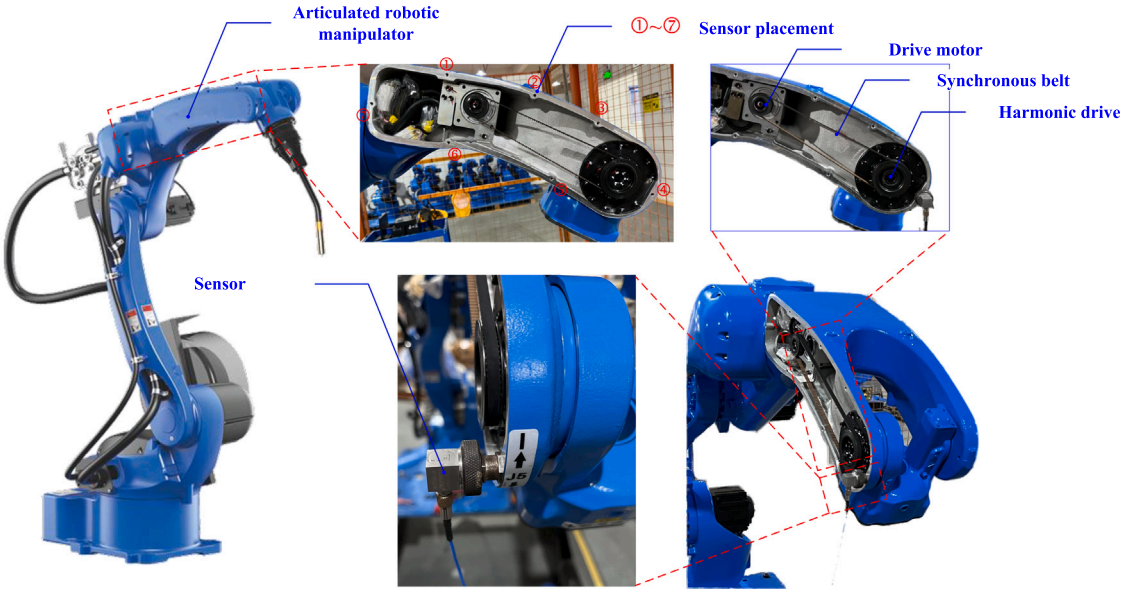

Fig. 2. Test rig and standardized sampling posture for the articulated robotic manipulator.

signals shown in subpanels (a_2)~(d_2). The denoising process utilized a decomposition level of 5, the 'db8' mother wavelet, and soft thresholding with a value of 0.005. The 'db8' (Daubechies 8) wavelet was selected for its inherent asymmetry and regularity, which provide a closer morphological match to the transient mechanical impulses of harmonic drives compared to symmetric or lower-order wavelets. A decomposition level of 5 was determined empirically to effectively separate the high-frequency sensor noise from the structural vibration bands of the manipulator. Furthermore, soft thresholding was employed to suppress noise coefficients while maintaining continuity in the reconstructed signal, thereby avoiding the pseudo-Gibbs artifacts often introduced by hard thresholding. Although the global waveform morphology remains largely preserved, a closer inspection of the static intervals reveals a reduction in signal 'fuzziness' in (a_2) relative to (a_1). This reduction in ambient noise variance, though visually marginal over the full 60-second cycle, represents a critical enhancement in signal fidelity required for accurate fault feature isolation.

Following denoising, Short-Time Energy (STE) analysis was employed to partition the signals into motion and static phases in a manner that is both adaptive and sensitive to speed variations. Because the motor speed increases progressively during acquisition, the segmentation threshold was correspondingly adjusted to track the evolving signal amplitude, thereby optimizing discrimination performance. The STE profiles for the four signal sets are presented in panels (a_1)~(d_1) of Fig. 4, and the corresponding segmentation outcomes are shown in panels (a_2)~(d_2). This approach yields a precise delineation of active motion periods from static intervals, providing a reliable foundation for more granular, stage-specific analysis.

To quantitatively validate the precision of the STE segmentation, the extracted motion intervals were compared against the ground-truth rotational speed data recorded by the robot's motor encoders. Table 1 summarizes the performance metrics across the four health states. In this evaluation, the missing detection rate refers to the proportion of missed active frames caused by a delayed detection start or early termination, while the false detection rate represents the proportion of falsely detected frames resulting from an early start or delayed termination. The overall accuracy is derived from these two metrics. As evidenced by the average accuracy of 92.58%, the adaptive thresholding method reliably isolates active operation periods without requiring external tachometer synchronization.

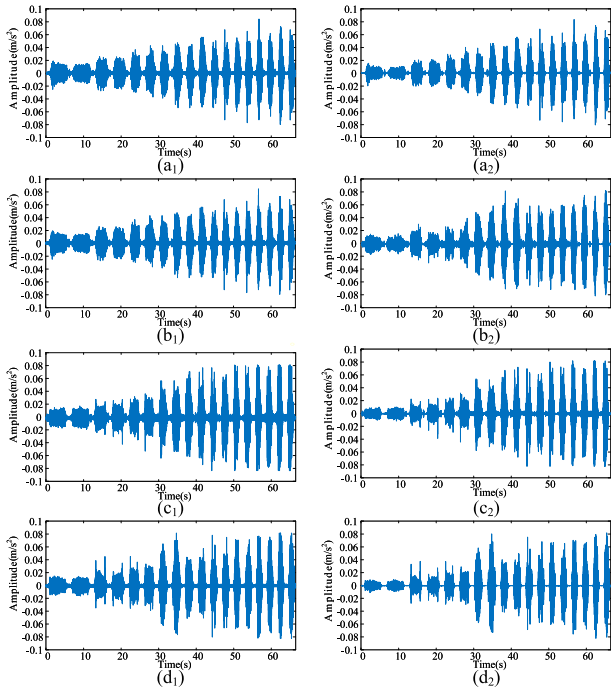

Fig. 3. Comparison of original (a_1)~(d_1) and wavelet-denoised (a_2)~(d_2) vibration signals across four health states: (a) Healthy, (b) Minor fault, (c) Severe fault, and (d) Moderate fault.

Table 1
Quantitative metrics of STE-based motion segmentation.

Fault Level	Missing Rate (%)	False Detection Rate (%)	Accuracy (%)
Healthy	6.57	0.31	93.12
Minor fault	8.12	0.15	91.73
Severe fault	7.69	0.29	92.02
Moderate fault	6.38	0.17	93.45
Average	**7.19**	**0.23**	**92.58**

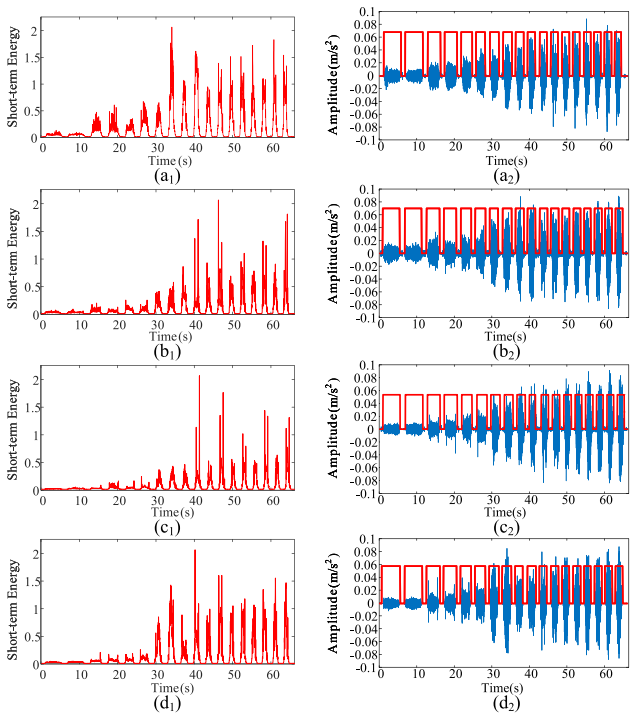

Fig. 4. STE profiles (a_1)~(d_1) and adaptive segmentation outcomes (a_2)~(d_2) across four health states: (a) Healthy, (b) Minor fault, (c) Severe fault, and (d) Moderate fault.

4.3. STFT-based stage delineation and steady-state extraction

Subsequent to segmentation, a specific motion segment was selected for in-depth analysis. Given that faults in articulated manipulators tend to manifest more prominently at elevated rotational speeds, the 9th motion segment – corresponding to 3000 r/min – was chosen for examination. Within a single motion segment, the signal exhibits non-stationarity, and three stages acceleration, steady-state, and deceleration—are clearly identifiable, as shown in panels (a_1)~(d_1) of Fig. 5. To characterize these temporal variations, the Short-Time Fourier Transform (STFT) was computed using Eq. (6) with a window length of 256 samples, and the resulting spectrograms for the four samples are presented in panels (a_2)~(d_2) of Fig. 5. The stage-specific frequency boundaries derived from these spectrograms were subsequently used to isolate the steady-state interval for further analysis; employing Eqs. (7) through (12), the signals in Fig. 5 were segmented accordingly, with the results summarized in Fig. 6, where panels (a_2)~(d_2) denote the extracted steady-state phases.

To quantitatively evaluate the quality of the phase segmentation, particularly the isolation of the steady-state regime from the transient phases, we employed the Coefficient of Variation (CV) of the dominant frequency as a stability metric. As transition boundaries in electromechanical systems are inherently continuous, the CV provides a robust measure of stationarity. Our analysis reveals that the extracted steady-state phases exhibit high frequency stability with an average CV of 0.012. In contrast, the acceleration and deceleration phases demonstrate significantly higher variability, with average CV values of 0.165 and 0.158, respectively. This order-of-magnitude disparity quantitatively confirms that the proposed STFT-based thresholding successfully isolates the stationary segments required for reliable Walsh–Fourier analysis.

4.4. Comparative analysis and method validation

In this section, we conduct a systematic validation of the proposed CHS-WFT framework. To rigorously establish the necessity of this novel approach, we first perform a comparative analysis against conventional diagnostic techniques, including Short-Time Fourier Transform (STFT), Continuous Wavelet Transform (CWT), and single-domain spectral analysis (FFT and Walsh Transform). By identifying the specific limitations of these established methods in characterizing impulsive fault features under non-stationary conditions, we provide the empirical justification for the cross-correlation heatmap strategy implemented in the proposed framework.

4.4.1. Diagnostic limitations of conventional TFA methods

Before evaluating the proposed framework, it is essential to establish why conventional Time–Frequency Analysis (TFA) methods are insufficient for this specific diagnostic task. Fig. 5 provides a comprehensive feature comparison across the Time (Column 1), STFT (Column 2), and CWT (Column 3) domains for all health conditions. While the gross energy intensity generally increases with fault severity (transitioning from Row a to Row d), standard TFA methods fail to resolve the discrete impulsive signatures associated with harmonic drive degradation.

As seen in the STFT spectrograms (Figs. 5(a_2)~(d_2)), the analysis suffers from significant spectral smearing; due to the fixed window length constraint, the short-duration impulsive energy is dispersed across the frequency axis, appearing as broad, indistinct vertical bands rather than sharp, discriminative features. Consequently, distinct fault-related sidebands are effectively masked by the dominant structural resonance band (1200–1500 Hz), which persists across all health states. To overcome these resolution limits, the Continuous Wavelet Transform (CWT) results were introduced (Figs. 5(a_3)~(d_3)), yet they reveal critical sensitivity issues. First, all health states exhibit a dominant continuous energy band around 1200–1500 Hz, likely attributable to fundamental meshing resonance, which persists regardless of fault status and masks underlying fault features. Second, while energy "hot spots" emerge in the 1500–2500 Hz range for the Minor and Moderate faults, these features manifest as faint, blurry striations rather than distinct impacts. Notably, the high-frequency energy distribution of the Moderate fault (Fig. 5(d_3)) remains statistically similar to the Healthy state (Fig. 5(a_3)), posing a significant risk of misclassification. Finally, in the Severe state (Fig. 5(c_3)), the fault signatures are heavily smeared into continuous "energy blobs". This phenomenon confirms the theoretical basis mismatch: the oscillating nature of standard wavelet bases (e.g., Morlet or Bump) cannot morphologically match the rectangular-like impact shocks of the harmonic drive, causing "ringing" artifacts that degrade temporal resolution.

These comparative results confirm that neither sinusoidal nor oscillatory bases can effectively decouple the modulation behavior of impulsive faults from the stationary background noise. The inability of CWT to distinctively separate Moderate faults from the Healthy baseline, combined with the smearing observed in Severe states, necessitates the adoption of the proposed Walsh–Fourier strategy to achieve a sparse and discriminative representation.

4.4.2. Performance of the proposed CHS-WFT framework

As noted earlier, the preliminary stage segmentation analysis focused on the 9-th motion segment. For a comprehensive demonstration of subsequent steps, we extend the analysis to include the 9-th (3000 r/min), 11-th (3500 r/min), and 13-th (4000 r/min) motion segments, thereby covering a representative range of rotational speeds and ensuring that the methodology is assessed under distinct operating conditions.

For the steady-state signals of these segments, we first normalize the energy of each frame signal using Eqs. (23) and (24) so as to enforce amplitude invariance and, in turn, ensure that the final results are not confounded by global scaling differences in signal amplitude across operating regimes. Fast Fourier Transform (FFT) was then computed using Eqs. (13) and (15) with a frame length $N = 256$, and the results are presented in Fig. 7. The figure consolidates three rotational

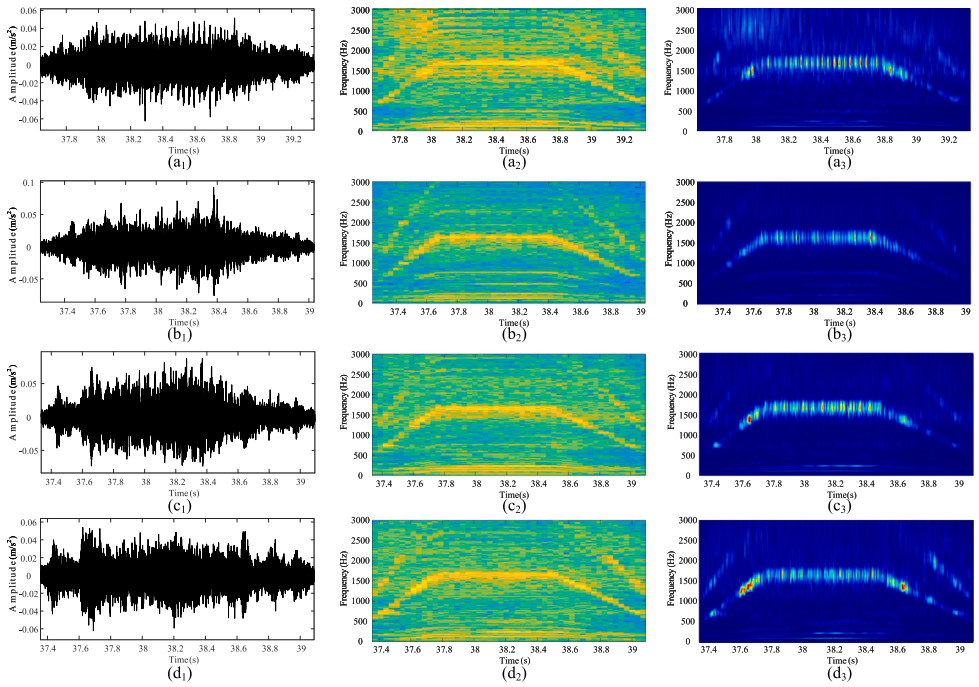

Fig. 5. Comparison of signal features in the time domain (a_1)~(d_1), STFT spectrograms (a_2)~(d_2), and CWT scalograms (a_3)~(d_3) across four health states: (a) Healthy, (b) Minor fault, (c) Severe fault, and (d) Moderate fault.

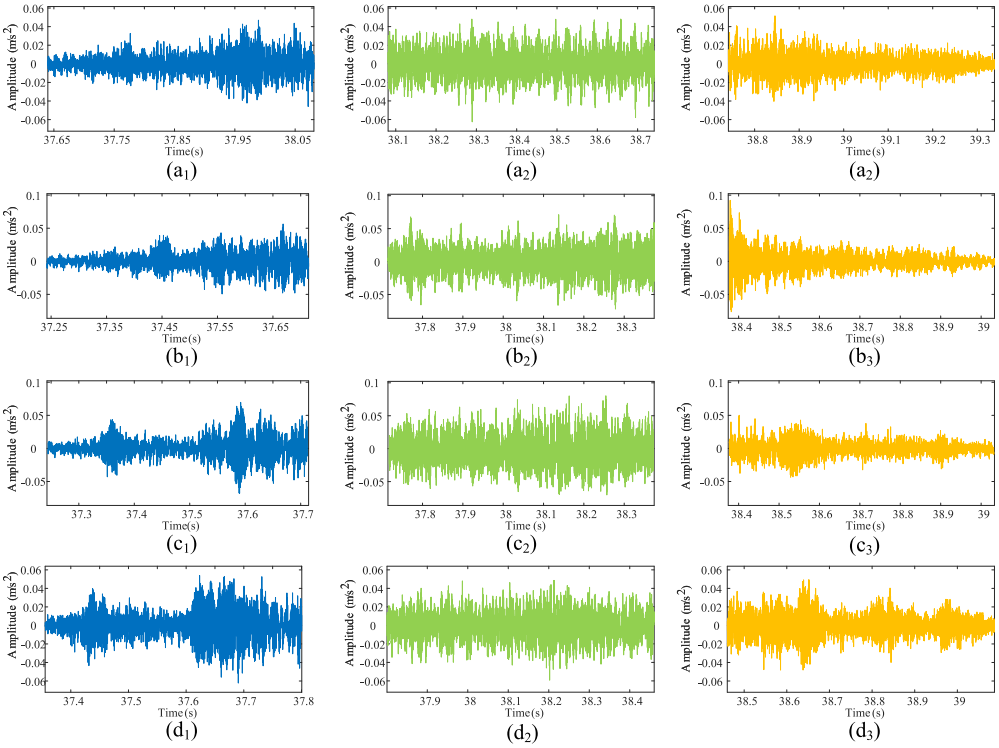

Fig. 6. Extracted motion phases including acceleration (a_1)~(d_1), steady-state (a_2)~(d_2), and deceleration (a_3)~(d_3) across four health states: (a) Healthy, (b) Minor fault, (c) Severe fault, and (d) Moderate fault.

speeds (3000 r/min, 3500 r/min, 4000 r/min) into subpanel groups (a1)~(d1), (a2)~(d2), and (a3)~(d3), each group corresponding to a particular speed; within each group, panels (a)~(d) denote healthy, minor fault, severe fault, and moderate fault samples, respectively. In parallel, the Walsh–Hadamard Transform (WHT) was applied using Eqs. (16) and (17), with outcomes shown in Fig. 8 under the same

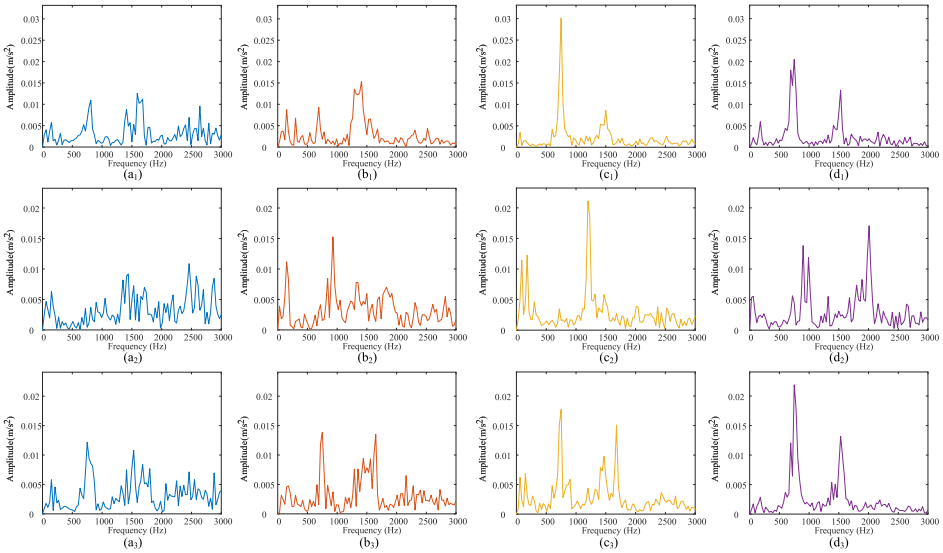

Fig. 7. Frame-wise FFT magnitude spectra at rotational speeds of 3000 r/min (a_1)~(d_1), 3500 r/min (a_2)~(d_2), and 4000 r/min (a_3)~(d_3) across four health states: (a) Healthy, (b) Minor fault, (c) Severe fault, and (d) Moderate fault.

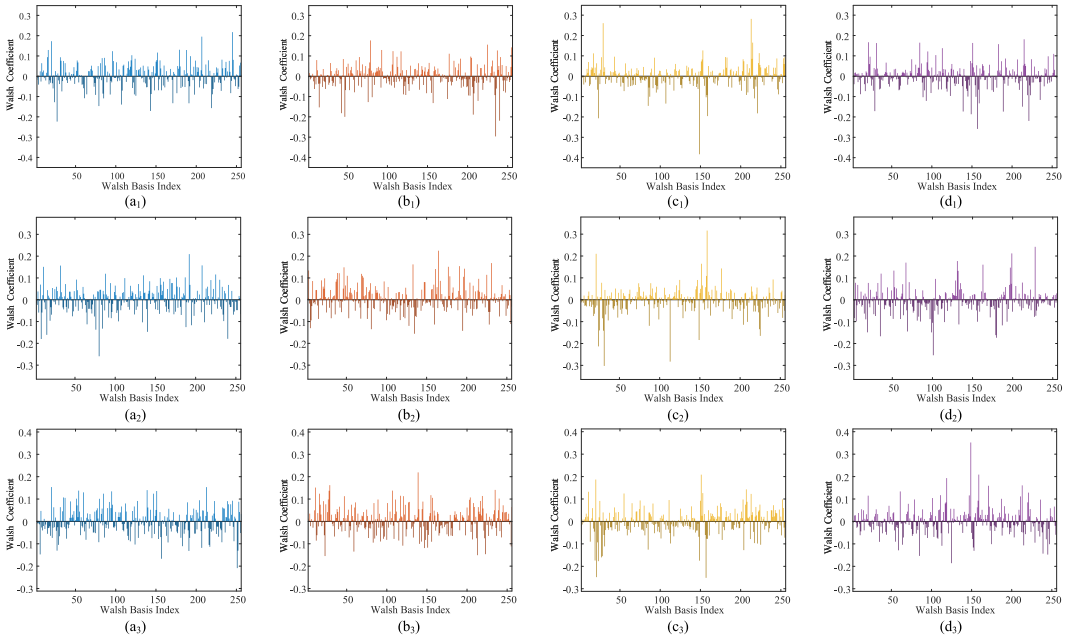

Fig. 8. Frame-wise WHT coefficient distributions at rotational speeds of 3000 r/min (a_1)~(d_1), 3500 r/min (a_2)~(d_2), and 4000 r/min (a_3)~(d_3) across four health states: (a) Healthy, (b) Minor fault, (c) Severe fault, and (d) Moderate fault.

layout—namely, subpanel groups (a1)~(d1) to (a3)~(d3) for the three speeds, and (a)~(d) for the four sample types in consistent order—thus enabling a like-for-like comparison between the Fourier and Walsh domains. Examination of both transforms reveals that across the three rotational speeds, there are no distinct differences between healthy and faulty samples in either the Fourier frequency domain or the Walsh domain. Taken together, these results do not yield sufficiently discriminative features to support a reliable, quantitative distinction among operational health states.

To further analyze the interaction between periodic and transient features, we processed the steady-state signals of the three selected motion segments (9-th at 3000 r/min, 11-th at 3500 r/min, and 13-th at 4000 r/min). For each segment, we first reconstructed sinusoidal signals for each frequency component using Eq. (18), thereby preserving the periodicity associated with specific spectral components. We then quantified the cross-correlation between these reconstructed sinusoids and weighted Walsh functions using Eqs. (19) and (20), generating heatmaps that visualize frequency–transient interaction patterns as correlation-intensity fields.

A detailed examination of the heatmaps such as those shown in Figs. 9 ~ 11 reveals striking discriminatory features that were absent in the FFT and WHT results, underscoring their value as a core output of our methodology in capturing subtle fault-related patterns. These heatmaps are derived from the cross-correlation between reconstructed periodic sinusoids using Eq. (18) and weighted Walsh functions using Eqs. (19) ~ (20). They visualize the strength of interaction between frequency-domain periodic components and time-domain transient features, with color intensity directly reflecting correlation magnitude.

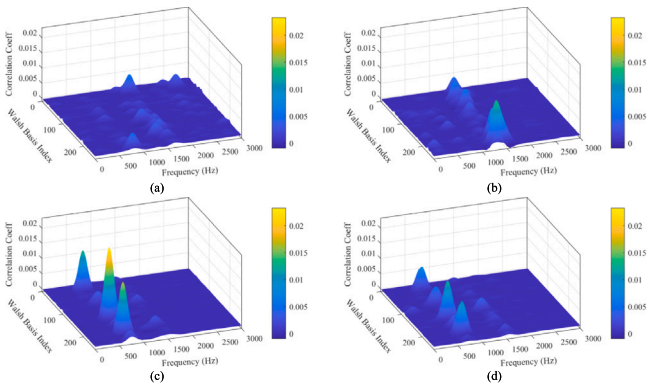

Fig. 9. Cross-correlation interaction heatmaps at 3000 r/min, computed between reconstructed periodic sinusoids (Eq. (18)) and weighted Walsh functions (Eqs. (19)~(20)). Panels: (a) Healthy; (b) Minor fault; (c) Severe fault; (d) Moderate fault.

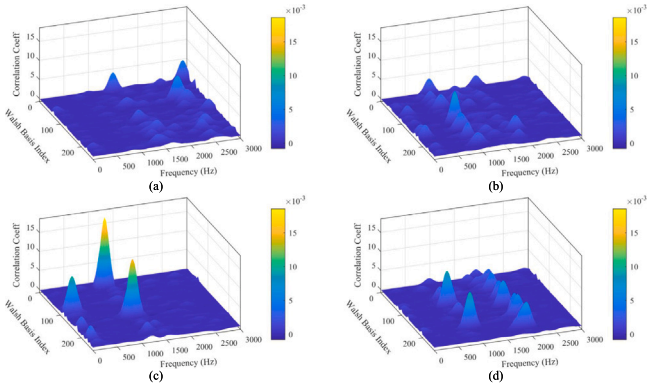

Fig. 10. Cross-correlation interaction heatmaps at 3500 r/min using the same construction as Fig. 9. Panels: (a) Healthy; (b) Minor fault; (c) Severe fault; (d) Moderate fault.

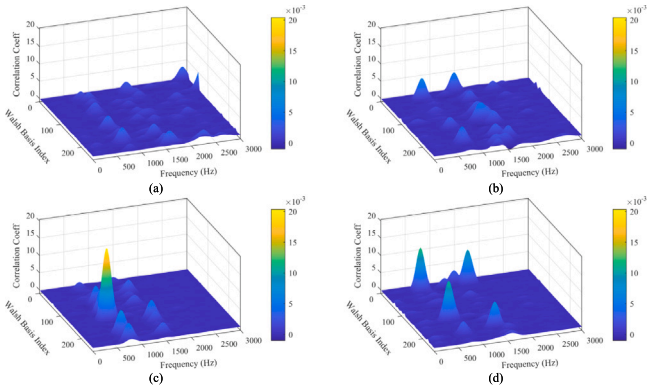

Fig. 11. Cross-correlation interaction heatmaps at 4000 r/min using the same construction as Fig. 9. Panels: (a) Healthy; (b) Minor fault; (c) Severe fault; (d) Moderate fault.

For healthy samples across all three rotational speeds such as those shown in Figs. 9(a), 10(a), 11(a), the heatmaps exhibit a remarkably flat and homogeneous distribution. There are no distinct peaks or localized high-intensity regions, indicating that periodic frequency components and transient features interact weakly and uniformly. This pattern aligns with the inherent stability of a healthy system: vibration energy is evenly distributed across multiple frequency–transient pairs, with no dominant coupling between specific periodic oscillations and impulsive behaviors. In contrast, faulty samples display distinct deviations from this baseline. For minor fault samples such as those shown in Figs. 9(b), 10(b), 11(b), the heatmaps are no longer as flat as those of healthy samples. Instead, scattered small peaks begin to appear—these peaks have relatively low energy intensity, reflecting weak but localized correlations between certain periodic components (likely early fault-induced oscillations) and transient features (possibly incipient mechanical irregularities). Though subtle, these low-intensity peaks already break the uniform energy distribution of the healthy state, marking the initial departure from normal operation. Moderate fault samples such as those shown in Figs. 9(d), 10(d), 11(d) exhibit a more pronounced pattern: the number of high-intensity peaks increases, and their spatial distribution expands slightly, while individual peak intensity rises noticeably compared to minor faults. This suggests that as fault severity progresses, the range of periodic components affected by the fault broadens, and their coupling with transient features strengthens—likely due to exacerbated mechanical wear leading to more frequent and intense impulsive interactions with periodic motion.

The most dramatic features appear in severe fault samples (Figs. 9(c), 10(c), 11(c)), where the heatmaps are dominated by a small number of extremely intense, tightly localized peaks. These peaks often span narrow ranges of both frequency and transient indices, indicating highly concentrated energy in specific periodic–transient interactions. This concentration arises from severe mechanical degradation – such as localized damage or misalignment – creating dominant fault frequencies that strongly couple with intense, repetitive transient vibrations. Collectively, these observations confirm that the cross-correlation heatmaps effectively capture the progression of fault severity by quantifying the evolving coupling between periodic and transient features—something FFT and WHT alone cannot achieve. This makes them a critical tool for distinguishing not only between healthy and faulty states but also for grading fault severity across operational speeds.

To quantitatively characterize the heatmap-derived features and enable fault severity assessment, we computed the Multi-Scale Energy Entropy (MEE) using a scale parameter $S = 3$ for each frame's heatmap across the four sample sets—healthy, minor fault, severe fault and moderate fault—and three rotational speeds (3000 r/min, 3500 r/min, 4000 r/min). The results are presented as scatter plots in Fig. 12, where subpanels (a), (b) and (c) correspond to the three rotational speeds respectively, and each data point represents the MEE of a single frame's heatmap. As shown in the figure, a consistent trend emerges across all speeds: healthy samples exhibit notably higher MEE values for most frames, which stems from the uniform energy distribution in their heatmaps. Faulty samples, by contrast, show a clear negative correlation between fault severity and MEE—severe faults have the lowest overall MEE, a result of the highly concentrated peaks in their heatmaps; moderate faults fall in the middle; and minor faults have MEE values closer to healthy samples but still distinguishable. This confirms that combining heatmap generation with MEE calculation not only reliably distinguishes healthy from faulty states but also enables objective grading of fault severity, with robust performance across different rotational speeds.

To validate the observed visual separability quantitatively, we performed independent samples t-tests between the healthy baseline and fault features across 3000–4000 r/min. As detailed in Table 2, all fault categories yielded P-values significantly below the threshold ($\alpha = 0.05$), with severe faults reaching $P < 10^{-4}$. These results statistically confirm that the extracted features are distinct from the healthy state, effectively overcoming the variance induced by speed fluctuations.

4.5. Comparative assessment of diagnostic performance across rotational speeds

To rigorously substantiate the efficacy of the proposed approach, we conducted a comparative analysis against three established condition-monitoring indices—Fuzzy Entropy [21], Envelope Kurtosis [29], and

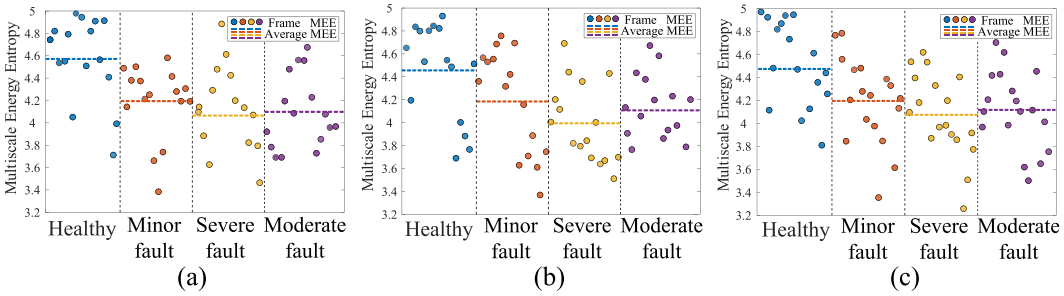

Fig. 12. Frame-wise Multi-Scale Energy Entropy (MEE, S=3) derived from interaction heatmaps across three rotational speeds and four health states. Subpanels: (a) 3000 r/min; (b) 3500 r/min; (c) 4000 r/min.

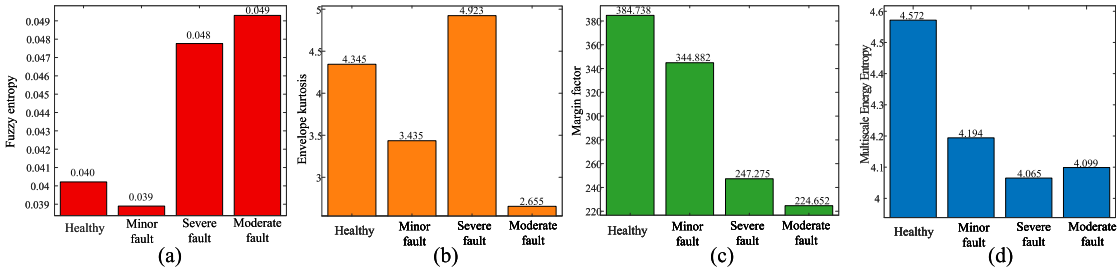

Fig. 13. Comparative assessment of diagnostic indices for healthy versus faulty samples at 3000 r/min. Subpanels: (a) Fuzzy Entropy; (b) Envelope Kurtosis; (c) Margin Factor; (d) CHS–WFT-derived Multi-Scale Energy Entropy.

Table 2
t-test results of feature separability across varying speeds.

Speed (r/min)	Comparison (vs. Normal)	*P*-Value	Significance
3000	Minor fault	6.71×10^{-3}	**
	Severe fault	4.08×10^{-4}	***
	Moderate fault	4.30×10^{-4}	***
3500	Minor fault	4.42×10^{-2}	*
	Severe fault	9.48×10^{-4}	***
	Moderate fault	5.11×10^{-3}	**
4000	Minor fault	7.93×10^{-3}	**
	Severe fault	1.58×10^{-5}	***
	Moderate fault	1.56×10^{-3}	**

* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

Margin Factor [30]—together with the CHS–WFT-derived Multi-Scale Energy Entropy (MEE). Specifically, steady-state signals from the 3000 r/min stage (panels (a_2)~(d_2) in Fig. 6), as well as those acquired at 3500 r/min and 4000 r/min, were evaluated using all four indices under consistent preprocessing and parameter settings. The corresponding outcomes are summarized in Figs. 13, 14, and 15, each comprising four subpanels that map one-to-one to the indices under comparison. Taken together, these results show that the three traditional indices do not reliably distinguish healthy from faulty samples across rotational speeds, thereby underscoring the advantages of our Fourier–Walsh joint methodology.

To evaluate discriminative capability beyond simple averages, Fig. 16 presents the frame-wise distribution of these indicators for the 3000 r/min condition. As visualized in the scatter plots, the traditional indices fail to establish a valid diagnostic gradient, suffering from severe class entanglement: Fuzzy Entropy (Fig. 16a) fails to distinguish the "Healthy" state from the "Minor Fault" state, as the two distributions are virtually identical in their mean levels and overlap significantly, rendering the method blind to early-stage degradation. Regarding Envelope Kurtosis (Fig. 16b), the diagnostic trend is logically incoherent; the "Healthy" baseline registers higher kurtosis values than the "Minor Fault" condition, violating the physical expectation that fault-induced impulses should increase kurtosis. Similarly, the Margin Factor (Fig. 16c) struggles to separate the "Healthy" baseline from the "Minor Fault" condition, where substantial overlap indicates a high probability of false negatives.

In contrast, the improved discriminative capability observed in the results originates from our proposed Cross-correlation Heatmap Synthesis with Walsh–Fourier Transformation (CHS–WFT). The framework jointly encodes frequency-domain periodicity (via the Fourier transform) and time-domain transient signatures (via the Walsh–Hadamard transform) and, crucially, constructs cross-correlation heatmaps that render explicit the interactions between these dual-domain representations. By capturing such cross-domain relationships, CHS–WFT extracts stable, fault-relevant patterns that persist across rotational speeds; moreover, we operationalize this information using MEE, which quantifies the uniformity of energy distribution over frequency transient pairs in the heatmaps and thereby enables reliable state differentiation. Collectively, these findings confirm that the core advantage lies in the CHS–WFT paradigm itself: its multi-domain fusion together with heatmap-based feature visualization overcomes the limitations of single-attribute indices and provides a more comprehensive basis for diagnosing articulated robotic manipulator conditions.

To provide a consolidated quantitative benchmark, Table 3 summarizes the average diagnostic indicator values for the proposed CHS-WFT method against the three conventional baseline methods across all tested rotational speeds. The evaluation criterion is the ability to maintain a monotonic severity gradient, which is essential for automated decision-making and prognosis.

As evidenced in Table 3, conventional indices exhibit stochastic behavior that compromises diagnostic reliability. For instance, at 3000 r/min, the Margin Factor drops from Healthy to Moderate but anomalously rises at the Severe stage, breaking the degradation trend. Similarly, Envelope Kurtosis and Fuzzy Entropy fluctuate without a clear correlation to severity, often assigning higher values to faulty states than to the baseline.

In distinct contrast, the proposed CHS-WFT method yields MEE values that maintain a strict monotonic decrease across all three speed

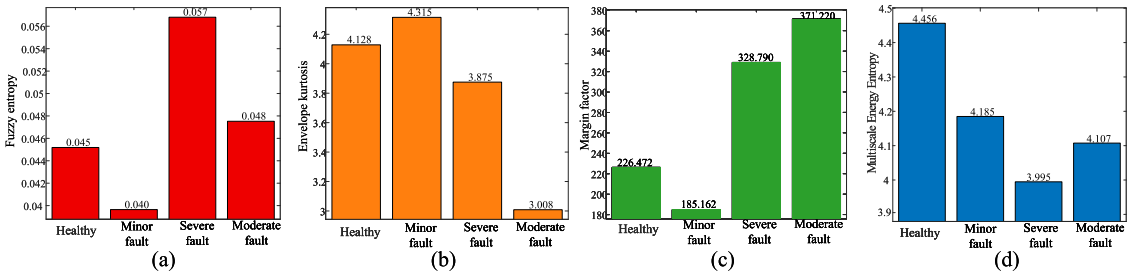

Fig. 14. Comparative assessment of diagnostic indices for healthy versus faulty samples at 3500 r/min. Subpanels: (a) Fuzzy Entropy; (b) Envelope Kurtosis; (c) Margin Factor; (d) CHS–WFT-derived Multi-Scale Energy Entropy.

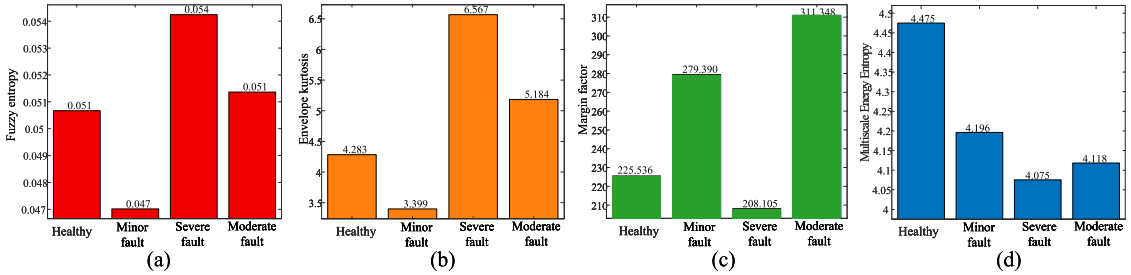

Fig. 15. Comparative assessment of diagnostic indices for healthy versus faulty samples at 4000 r/min. Subpanels: (a) Fuzzy Entropy; (b) Envelope Kurtosis; (c) Margin Factor; (d) CHS–WFT-derived Multi-Scale Energy Entropy.

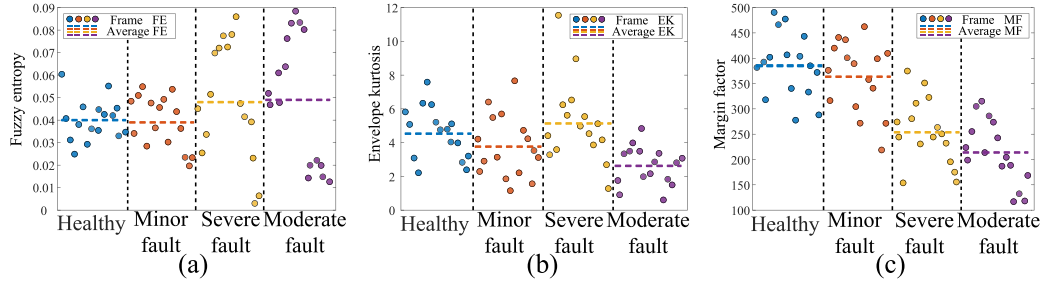

Fig. 16. Frame-wise distribution of comparative diagnostic indicators at 3000 r/min: (a) Fuzzy Entropy, (b) Envelope Kurtosis, and (c) Margin Factor.

Table 3
Quantitative comparison of average diagnostic indicator values across methods and speeds.

Speed (r/min)	Method	Healthy	Minor	Moderate	Severe	Diagnostic Trend Stability
3000	Fuzzy Entropy	0.040	0.039	0.049	0.048	Unstable (Fluctuating)
	Envelope Kurtosis	4.345	3.435	2.655	4.923	Failed (No clear trend)
	Margin Factor	384.738	344.882	224.652	247.275	Unstable (Inverted at Severe)
	CHS-WFT	**4.572**	**4.194**	**4.099**	**4.065**	**Stable Monotonic** (↘)
3500	Fuzzy Entropy	0.045	0.040	0.048	0.057	Unstable (Fluctuating)
	Envelope Kurtosis	4.128	4.315	3.008	3.875	Failed (No clear trend)
	Margin Factor	226.472	185.162	371.220	328.790	Failed (Inverted at Moderate)
	CHS-WFT	**4.456**	**4.185**	**4.107**	**3.995**	**Stable Monotonic** (↘)
4000	Fuzzy Entropy	0.051	0.047	0.051	0.054	Unstable (Fluctuating)
	Envelope Kurtosis	4.283	3.399	5.184	6.567	Failed (Inverted Trend)
	Margin Factor	225.536	279.390	311.348	208.105	Unstable (Fluctuating)
	CHS-WFT	**4.475**	**4.196**	**4.118**	**4.075**	**Stable Monotonic** (↘)

regimes. This validates the method's robustness: by effectively decoupling the fault modulation from the speed-dependent carrier, CHS-WFT maps the physical degradation process to a reliable, linearly separable prognostic indicator.

4.6. Robustness analysis and mechanism verification under noise injection

To rigorously evaluate the framework's resilience to environmental interference, a systematic noise injection study was conducted. Gaussian white noise was superimposed onto the raw vibration signals of all four health states to create synthetic datasets with SNR of 10 dB, 5 dB, and 0 dB. This progression allows for a granular assessment of how signal degradation impacts both the steady-state extraction and the core diagnostic engine.

Fig. 17 illustrates the impact of noise on the time-domain waveforms of the severe fault condition. Panels (a)–(c) depict the full-length severe fault signals across the decreasing SNR range. As the noise level increases, the amplitude fluctuations associated with the robot's motion start and stop events – critical for automated segmentation – become

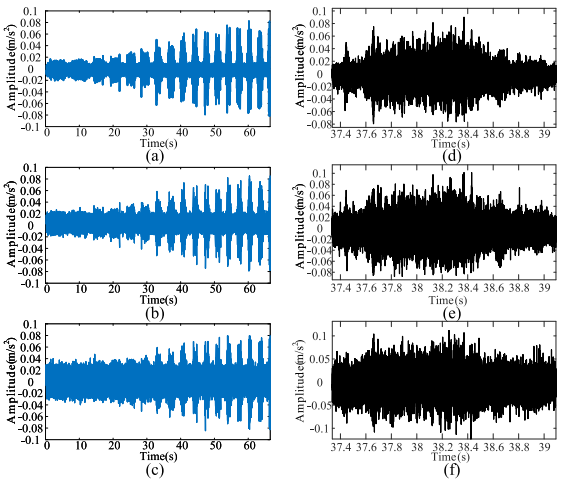

Fig. 17. Time-domain waveforms of the severe fault condition under varying noise levels. (a)–(c) Full-length signals at SNR = 10 dB, 5 dB, and 0 dB, respectively; (d)–(f) Extracted steady-state signals (3000 r/min) at SNR = 10 dB, 5 dB, and 0 dB, respectively.

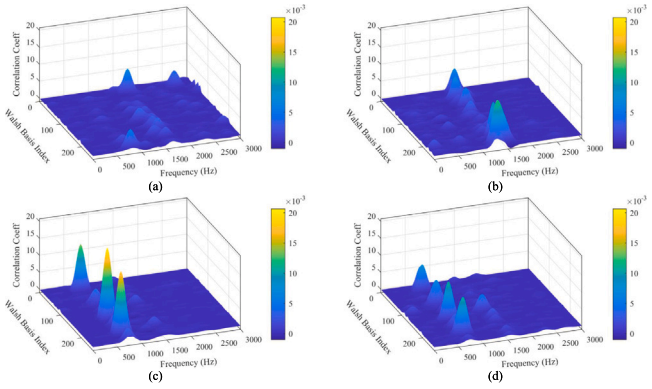

Fig. 18. Cross-correlation Heatmaps at SNR = 10 dB. (a) Healthy; (b) Minor fault; (c) Severe fault; (d) Moderate fault.

progressively submerged. Panels (d)–(f) present the extracted steady-state signals corresponding to the 3000 r/min interval. At 10 dB and 5 dB, the adaptive segmentation algorithm remains robust, successfully isolating the target interval. However, at 0 dB (Fig. 17(c) and (f)), the broadband noise floor completely masks the transient energy changes. Consequently, the time stamps obtained from the stable, higher-SNR experiments were adopted for the 0 dB case to ensure the correct extraction of the steady-state phase shown in Fig. 17(f).

Following stable phase extraction, Cross-correlation Heatmaps were generated for all four health conditions to verify feature preservation. The evolution of these features under decreasing SNR is presented in Figs. 18 through 20. In each figure, subplots (a)–(d) correspond to the Healthy, Minor, Severe, and Moderate fault states, respectively.

At 10 dB (Fig. 18) and 5 dB (Fig. 19), the distinction between health states is sharp: the Healthy heatmaps (a) remain topologically flat, while the faulty states (b-d) exhibit distinct interaction peaks. Crucially, even at 0 dB (Fig. 20), where the time-domain signal is severely degraded, this contrast is preserved. The Healthy heatmap (Fig. 20(a)) does not hallucinate false features from the Gaussian noise, while the Severe fault (Fig. 20(c)) retains its characteristic high-intensity clusters. This demonstrates that the Walsh–Fourier cross-correlation acts as a matched filter, effectively suppressing incoherent noise while preserving periodic-transient dependencies.

Finally, the quantitative robustness is validated in Fig. 21, which presents the frame-wise Multi-Scale Energy Entropy (MEE) distributions

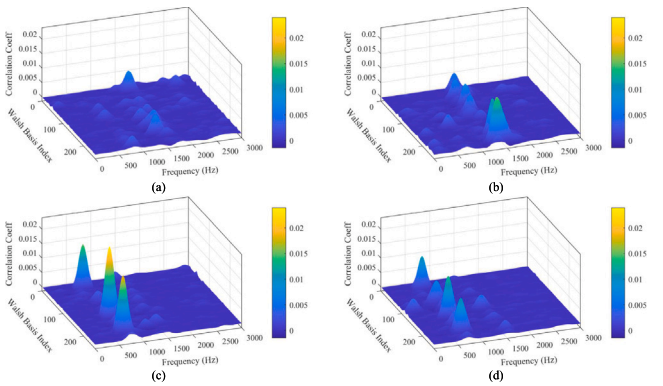

Fig. 19. Cross-correlation Heatmaps at SNR = 5 dB. (a) Healthy; (b) Minor fault; (c) Severe fault; (d) Moderate fault.

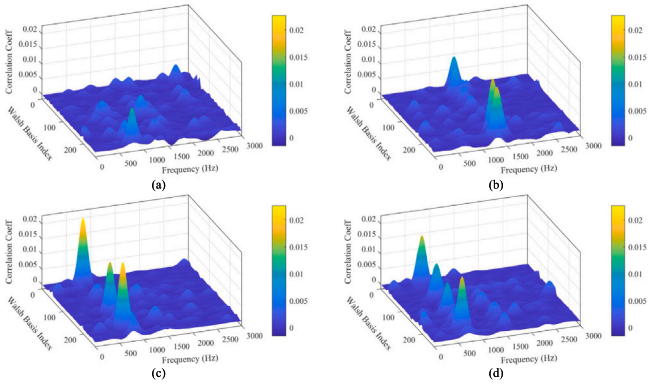

Fig. 20. Cross-correlation Heatmaps at SNR = 0 dB. (a) Healthy; (b) Minor fault; (c) Severe fault; (d) Moderate fault.

across the three evaluated noise levels: (a) 10 dB, (b) 5 dB, and (c) 0 dB. A consistent diagnostic topology is observable across all panels. The Healthy state (blue markers) consistently maintains the highest entropy values, reflecting the spatially uniform energy distribution of the background noise in the absence of fault features. Conversely, the Severe fault condition (yellow markers) exhibits the lowest entropy, confirming that the sharp, high-intensity interaction peaks preserved in the heatmaps result in highly concentrated energy distributions. The Minor and Moderate faults occupy distinct intermediate bands. Most critically, even at 0 dB (Panel c), the intra-class clustering remains tight and the inter-class separation margins are preserved. This confirms that the monotonic severity gradient (Healthy > Minor > Moderate > Severe) is not merely an artifact of clean data but a robust feature that survives severe signal degradation.

5. Conclusion

CHS-WFT offers a speed-aware, cross-signal diagnostic framework that remains stable under nonstationary dynamics and wide speed variations that typically degrade vibration-based methods for harmonic drives. By fusing Fourier's harmonic sensitivity with Walsh's responsiveness to abrupt, piecewise transients in a joint cross-correlation heatmap, the method reveals inter-signal dependencies that are obscured in Fourier-only analyses. Multi-scale energy entropy computed on this representation yields compact, discriminative features that separate healthy from faulty states despite large speed fluctuations, reducing dependence on labeled data and detailed operating context. Collectively, these capabilities enable earlier, more reliable fault detection to support condition-based maintenance in articulated robotic systems. Despite the promising results, this study relies on certain

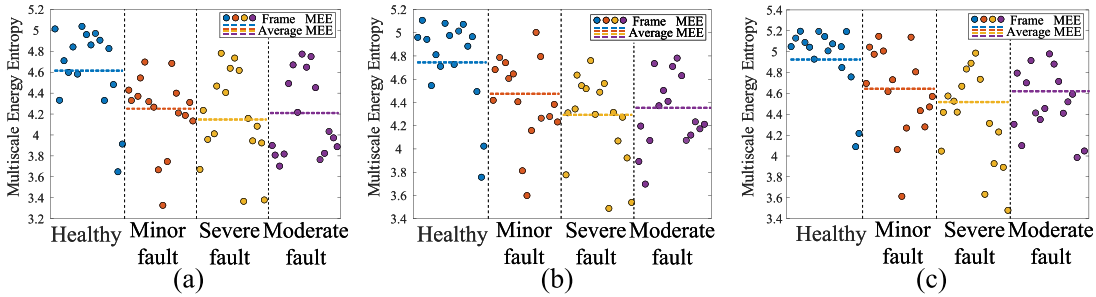

Fig. 21. Frame-wise Multi-Scale Energy Entropy (MEE, $S = 3$) distributions derived from interaction heatmaps. (a) SNR= 10 dB; (b) SNR= 5 dB; (c) SNR= 0 dB.

assumptions and limitations that merit consideration. First, the experimental validation was conducted primarily under variable rotational speeds with a constant load; the influence of varying load torque on the CHS-WFT feature distribution has not yet been explicitly characterized. Second, while the dataset was sufficient to establish statistical significance, the volume of data remains relatively limited. Future investigations will aim to expand the dataset diversity and examine the method's robustness under coupled speed-load variations to further generalize these findings. The future works include rigorous benchmarking across drives and regimes, real-time edge deployment with bounded latency/memory, and label-efficient learning on CHS-WFT representations. We will add multimodal fusion, extend to component-level diagnosis and RUL, establish invariance/robustness guarantees with calibrated uncertainty, and release datasets and protocols for reproducible adoption.

CRediT authorship contribution statement

Dong Wu: Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis. **Yazheng Wang:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Peng Chen:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition. **Chunhua Guo:** Validation, Resources, Methodology, Formal analysis. **Renpu Li:** Project administration, Methodology, Investigation, Funding acquisition, Formal analysis. **Ge Xin:** Supervision, Project administration, Investigation, Funding acquisition. **Changbo He:** Supervision, Software, Resources, Project administration, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] S. Park, J. Kim, W. Kwon, J. Kang, N. Kim, Mechanical fault detection and classification in variable reluctance resolver based on Lissajous curve indices, Measurement (2025) 118631.
[2] Z. Zhao, S. Sun, W. Xu, C. Shen, D. Wang, Dynamics modeling and fault diagnosis of flexible thin-walled elliptical bearings in harmonic reducers, Measurement 238 (2024) 115378.
[3] P. Chen, R. Zhang, S. Fan, J. Guo, X. Yang, Step-wise contrastive representation learning for diagnosing unknown defective categories in planetary gearboxes, Knowl.-Based Syst. 309 (2024) 112863.
[4] G. Zhang, Y. Tao, J. Wang, K. Feng, X. Han, A motor current signal-based fault diagnosis method for harmonic drive of industrial robot under time-varying speed conditions, IEEE Trans. Instrum. Meas. 74 (2025) 1–10.
[5] P. Chen, J. Gao, Y. Wu, C. He, G. Xin, S. Fan, J. Qi, Semi-supervised transfer graph representation learning with few-shot adaptation for gearbox diagnostics under extraneous transient noise, Struct. Health Monit. (2026) 14759217251414344.
[6] S. Yang, Y. Xiang, Z. Long, X. Ma, Q. Ding, J. Jia, Fault diagnosis of harmonic drives based on an SDP-ConvNeXt joint methodology, IEEE Trans. Instrum. Meas. 72 (2023) 1–8.
[7] P. Chen, J. Ma, C. He, Y. Jin, S. Fan, Progressive contrastive representation learning for defect diagnosis in aluminum disk substrates with a bio-inspired vision sensor, Expert Syst. Appl. (2025).
[8] P. Chen, Z. Ma, C. Xu, Y. Jin, C. Zhou, Self-supervised transfer learning for remote wear evaluation in machine tool elements with imaging transmission attenuation, IEEE Internet Things J. 11 (2024) 23045–23054.
[9] J. Chen, K. Wen, J. Xia, R. Huang, Z. Chen, W. Li, Knowledge embedded autoencoder network for harmonic drive fault diagnosis under few-shot industrial scenarios, IEEE Internet Things J. 11 (13) (2024) 22915–22925.
[10] H. Bilal, M.S. Obaidat, M. Shamrooz Aslam, J. Zhang, B. Yin, K. Mahmood, Online fault diagnosis of industrial robot using IoRT and hybrid deep learning techniques: An experimental approach, IEEE Internet Things J. 11 (19) (2024) 31422–31437.
[11] P. Chen, Y. Wu, C. Xu, C.-G. Huang, M. Zhang, J. Yuan, Interference suppression of nonstationary signals for bearing diagnosis under transient noise measurements, IEEE Trans. Reliab. (2025) 1–15.
[12] C. Xu, P. Chen, J. Gao, Y. Jin, M. Rao, Semi-supervised transfer learning preserving spatial homogeneity for gearbox diagnostics in extraneous transient noise, Nondestruct. Test. Eval. (2025) 1–29.
[13] P. Chen, Y. Wu, J. Ma, R. Zhang, G. Xin, C. He, Vibration-weighted maximum correlated kurtosis deconvolution and latent cyclic pattern discovery for fault diagnosis of high-speed rail bogies, J. Sound Vib. (2026) 119657.
[14] P. Chen, Y. Wu, C. Xu, Y. Jin, C. Zhou, Markov modeling of signal condition transitions for bearing diagnostics under external interference conditions, IEEE Trans. Instrum. Meas. 73 (3518308) (2024).
[15] Y. Wang, P. Chen, Q. Wei, J. Qi, C. He, C. Zhou, Multi-channel fusion scale transformed signals with magnetic leakage for damage detection in steel wire ropes, Nondestruct. Test. Eval. (2025) 1–26.
[16] P. Chen, Y. Wu, S. Fan, C. He, Y. Jin, J. Qi, C. Zhou, Adaptive signal regime for identifying transient shifts: A novel approach toward fault diagnosis in wind turbine systems, Ocean Eng. 325 (2025) 120798.
[17] B. Wang, Q. Liu, B. Pang, J. Xu, Dynamic modeling of flexible thin-walled ellipse bearing outer ring fault considering cam disturbance, Nonlinear Dynam. 113 (23) (2025) 32237–32263.
[18] W. He, J. Mao, Y. Wang, Z. Li, H. Xie, H. Shao, X. Zhao, Unified diagnostic and matching framework of fault and quality for robotic grinding system, IEEE Trans. Instrum. Meas. 73 (2024) 1–12.
[19] X. Wang, H. Jiang, M. Mu, Y. Dong, A trackable multi-domain collaborative generative adversarial network for rotating machinery fault diagnosis, Mech. Syst. Signal Process. 224 (2025) 111950.
[20] Y. Zhang, Y. Xu, T. Wang, Z. Yu, Z. Ma, Towards harmonic reducer performance degradation assessment via acoustic emission and micro-vibration information fusion, J. Vib. Eng. Technol. 13 (6) (2025) 431.
[21] Y. Li, S. Wang, Y. Yang, Z. Deng, Multiscale symbolic fuzzy entropy: An entropy denoising method for weak feature extraction of rotating machinery, Mech. Syst. Signal Process. 162 (2022) 108052.
[22] M. Lopez-Ramirez, C. Rodriguez-Donate, L.M. Ledesma-Carrillo, F.J. Villalobos-Pina, J.U. Munoz-Minjares, E. Cabal-Yepez, Walsh–Hadamard domain-based intelligent online fault diagnosis of broken rotor bars in induction motors, IEEE Trans. Instrum. Meas. 71 (2022) 1–11.
[23] C. Pan, Gibbs phenomenon removal and digital filtering directly through the fast Fourier transform, IEEE Trans. Signal Process. 49 (2) (2001) 444–448.
[24] J.-M. Lina, M. Mayrand, Complex daubechies wavelets, Appl. Comput. Harmon. Anal. 2 (3) (1995) 219–229.

[25] A. Halidou, Y. Mohamadou, A.A.A. Ari, E.J.G. Zacko, Review of wavelet denoising algorithms, Multimedia Tools Appl. 82 (27) (2023) 41539–41569.
[26] P.A. Schirmer, I. Mporas, Energy disaggregation from low sampling frequency measurements using multi-layer zero crossing rate, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2020, pp. 3777–3781.
[27] A. Shen, Y. Li, K. Noman, D. Wang, Z. Peng, K. Feng, Multiscale fluctuation-based symbolic dynamic entropy: a novel entropy method for fault diagnosis of rotating machinery, Struct. Health Monit. 24 (1) (2025) 402–420.
[28] F. Chen, L. Zhang, W. Liu, T. Zhang, Z. Zhao, W. Wang, D. Chen, B. Wang, A fault diagnosis method of rotating machinery based on improved multiscale attention entropy and random forests, Nonlinear Dynam. 112 (2) (2024) 1191–1220.
[29] L. Hua, X. Wu, T. Liu, S. Li, The methodology of modified frequency band envelope kurtosis for bearing fault diagnosis, IEEE Trans. Ind. Informat. 19 (3) (2022) 2856–2865.
[30] W. Teng, W. Wang, H. Ma, Y. Liu, Z. Ma, H. Mu, Adaptive fault detection of the bearing in wind turbine generators using parameterless empirical wavelet transform and margin factor, J. Vib. Control 25 (6) (2019) 1263–1278.