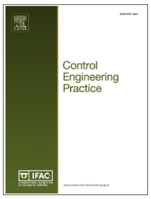
Control Engineering Practice 174 (2026) 107009

Contents lists available at ScienceDirect

Control Engineering Practice

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



Transient-aware spectral segmentation informed by energy-distribution modeling and optimized band selection for railway bogie fault diagnosis

Peng Chen [a,b], Qingsheng Wei [a], Yazheng Wang [a], Ge Xin [c], Changbo He [d,*]

[a] *College of Engineering, Shantou University, Shantou, 515063, Guangdong, PR China*
[b] *Key Laboratory of Intelligent Manufacturing Technology, Ministry of Education, Shantou, 515063, Guangdong, PR China*
[c] *School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, PR China*
[d] *School of Electrical Engineering and Automation, Anhui University, Hefei, 230601, PR China*

ARTICLE INFO

Keywords:
High-speed railway
Bogie axle bearings
Fault diagnosis
Vibration signal processing
Particle swarm optimization (PSO)
Frequency-band optimization
Nonstationary interference

ABSTRACT

Bogie axle bearings, critical to high-speed train safety and performance, operate in challenging low Signal-to-Noise Ratio (SNR) environments–particularly affected by complex, broadband interference from track impacts and braking events. Current diagnostics struggle with random pulse noise-masked fault signatures, while decomposition methods face mode mixing issues, and deep learning approaches require extensive labeled data while lacking interpretability. We introduce Transient-Aware Spectral Segmentation (TASS), which models energy distributions and optimizes frequency-band selection to suppress broadband, nonstationary interference in bogie bearing signals. Central is Amplitude-Weighted Ordinal Entropy (AWOE), which reweights ordinal patterns by amplitude to localize transient noise and overcome the amplitude insensitivity of permutation entropy, enabling an entropy-guided decomposition that fuses spectral-envelope and entropy distributions for physically interpretable, impulse-robust segmentation. In parallel, particle swarm optimization with an Enhanced Spectral Energy Ratio (ESER) automatically selects fault-sensitive bands that emphasize characteristic frequencies and harmonics while penalizing broadband transients. Across operating regimes and sensing modalities (onboard and wayside), TASS increases effective SNR, improves detection reliability, and reduces false alarms, supporting engineering assurance and condition-based maintenance.

1. Introduction

High-speed trains are highly integrated mechatronic systems in which the bogie–comprising the frame, wheelsets, and suspension–acts as a critical subsystem that governs dynamic performance and, consequently, operational safety (Wan et al., 2023). Within this architecture, bogie axle bearings operate under high rotational speeds, fluctuating radial and axial loads, sustained vibration, and environmental stressors, rendering them susceptible to progressive wear, surface fatigue (e.g., spalling), and cage or raceway damage. Because such failures can escalate rapidly and compromise service continuity and passenger safety, operators should employ continuous condition monitoring that integrates vibration, temperature, acoustic-emission, and lubricant-condition measurements with threshold alarms and trend analytics. When coupled with physics-informed models and data-driven prognostics, these data streams support predictive maintenance that enables early fault detection, optimized intervention scheduling, and targeted component replacement, thereby preserving rolling-stock integrity and availability while ensuring compliance with railway safety standards.

Bogie axle bearings–safety-critical rotating elements in rail bogies that shape running safety, ride quality, and maintenance policy–are commonly diagnosed via two complementary streams: vibration signal-processing fault diagnosis (Chen et al., 2026c, 2024b; Wu et al., 2026) and deep learning-based, data-driven monitoring (Chen et al., 2024a,c; Peng et al., 2026). The former treats measured vibration as a carrier of fault signatures and, by extracting physically interpretable time-, frequency-, and time-frequency-domain descriptors, supports expert reasoning and tends to generalize across operating regimes when labeled data are scarce (Cheng et al., 2022; Deng et al., 2021; Wei et al., 2026). By contrast, CNN-, RNN-, and Transformer-based models (Chen et al., 2025c; He et al., 2026; Shao et al., 2025) learn hierarchical representations directly from raw or minimally processed signals and, therefore, capture subtle, nonlinear, and nonstationary patterns that hand-crafted features may miss (Chen et al., 2026a; He et al., 2025b; Pan et al., 2024). However, these gains entail trade-offs: effective training typically

---

* Corresponding author.
E-mail addresses: pengchen@alu.uestc.edu.cn, dr.pengchen@foxmail.com (P. Chen), wei2311295448@163.com (Q. Wei), yazhengwang123@outlook.com (Y. Wang), ge.xin@bjtu.edu.cn (G. Xin), changbh@ahu.edu.cn (C. He).

https://doi.org/10.1016/j.conengprac.2026.107009
Received 19 December 2025; Received in revised form 7 March 2026; Accepted 10 April 2026

requires large, balanced labeled datasets; the resulting predictors often operate as "black boxes" that limit engineering assurance; and real-world label acquisition is hampered by fault rarity, variable loads and speeds, and environmental noise. As a result, method choice hinges on data availability and quality, the required level of interpretability and auditability, and the operational constraints of the target bogie system. In practice, hybrid strategies that fuse physics-guided features with learned representations are increasingly adopted to balance transparency, accuracy, and deployability.

In the safety-critical context of railway rolling stock, and specifically within the domain of bogie axle bearing fault diagnosis, the detection and interpretation of damage-indicative vibration signatures remain a key challenge that continues to drive innovation in condition monitoring and prognostics. Because defects arising from fatigue spalling, cracks, and abrasive wear typically excite structural resonances, they produce sparse, high-amplitude transients in the time domain and, via cyclostationary amplitude modulation, yield distinct fault-characteristic frequencies–such as cage frequency, ball-pass inner/outer race, and ball-spin–in the spectral domain. However, under realistic operating conditions marked by variable speed, fluctuating load, complex transmission paths, and pervasive ambient noise, the signal-to-noise ratio is often low, and thus the isolation of these diagnostic components becomes both technically demanding and operationally consequential. A notable methodological milestone was the introduction of Spectral Kurtosis (SK) and its Kurtogram representation, which quantify impulsiveness across frequency and thereby inform the selection of an optimal demodulation band for envelope analysis, as originally articulated by Antoni and Randall (2006), Park et al. (2025), Ugwiri et al. (2021). Nevertheless, despite this conceptual elegance, contemporary SK-driven bandwidth partitioning strategies remain vulnerable to exogenous impulsive events and nonstationary interferences that can mimic or obscure fault modulations; consequently, they may degrade diagnostic robustness, inflate false-alarm and missed-detection rates, and ultimately impede reliable decision-making in practical deployments.

Although spectral-kurtosis- (SK-) based bandwidth selection remains a widely adopted strategy for envelope demodulation in rotating machinery, it exhibits pronounced sensitivity to exogenous disturbances and nonstationary operating regimes, which, in dynamic environments, increases the risk of both false positives and missed detections (Xin et al., 2025). To mitigate these limitations while preserving interpretability, Zhang et al. (Hou et al., 2023) introduce the fault characteristic signal-to-noise ratio (FCSNR) to quantitatively assess periodic fault components in high-speed train wheelset bearings and, by integrating FCSNR with SK, deconvolution, and mode decomposition, demonstrate more reliable diagnosis under noisy, real-world conditions. In a complementary direction, Chen et al. (2025b) propose a short-term Markov transition frequency peak rate method that more precisely tracks temporal state transitions and flags abnormal signatures, thereby reducing the influence of transient noise and improving demodulation band selection accuracy under complex operations. Extending adaptivity to renewable energy assets, Chen et al. (2025a) develop a tailored wind-turbine bearing framework that combines a Dynamic Markov Transition Frequency with Adaptive Peak Rates (DMTF-APR) model and a Multi-Period Weighted Average Framework (MPWAF), thus enhancing robustness across variable loads and speeds. Furthermore, Yang et al. (2024) advance blind deconvolution by proposing harmonic noise kurtosis-time characteristic blind deconvolution (HTBD), whose objective function jointly optimizes Harmonic Noise Kurtosis (HNK) and the Time Characteristic Energy Ratio (TCER), which in turn better handles periodicity and sparsity while accelerating convergence and strengthening interference resistance. From a sparsity-driven perspective, Wang et al. (2024b) present generalized nonlinear hybrid-norm parallel sparse filtering (GNHNPSF), which leverages a Sigmoid nonlinearity to suppress impulsive anomalies and hybrid-norm constraints to stabilize feature extraction, thereby improving diagnostic accuracy amid heavy noise. Finally, Wang et al. (2024a) propose the LFIgram method, which uses the log envelope autocorrelation slice bispectrum (LEAB) and an LEAB feature index (LFI) to adaptively select resonant bands and extract coupled features, consequently overcoming demodulation limitations due to fault coupling and random slip and achieving high accuracy in both simulations and experiments.

In parallel with advances in demodulation and sparse filtering, recent studies have increasingly leveraged entropy-driven representations to construct generalized, data-efficient diagnostic frameworks for rotating machinery. Wang et al. (2025c) established a generalized fault diagnosis architecture grounded in refined time-shifted multi-scale phase entropy (RTSMPhE), which effectively captures phase-modulated fault signatures across variable operating conditions while enabling extension to multi-channel analysis. Complementing this, the same research group introduced composite multi-scale phase reverse permutation entropy (CMPRPE) (Zheng et al., 2023), a symmetry-aware entropy metric that enhances sensitivity to subtle asymmetries in vibration dynamics; when coupled with a whale optimization algorithm-tuned support vector machine (WOA-SVM), CMPRPE achieves high classification fidelity even under low signal-to-noise ratios. To further address spectral uncertainty, cumulative spectrum distribution entropy (CSDEn) was proposed by Wang et al. (2024c) as a global spectral regularity measure that quantifies energy dispersion across frequency bins, demonstrating superior robustness in distinguishing crack, spall, and breakage faults compared to DispEn, PerEn, and other classical entropies. Recognizing the challenge of limited labeled data, particularly in aerospace applications, another line of work developed a bonobo-inspired evolutionary optimization framework for aero-engine bearing diagnosis under extreme sample scarcity (Wang et al., 2024d), integrating hybrid mating strategies with classifiers such as LSSVM, CNN, LSTM, and time-series transformers (TST) to maximize discriminative power from minimal training instances. Extending into the visual domain, few-shot fault diagnosis via multi-scale perception multi-level feature fusion image quadrant entropy (MPSFFIQE) (Wang et al., 2025a) transforms time-series signals into angular field images and extracts quadrant-wise entropy features across scales, achieving 100% accuracy with only five samples per class on aero-engine datasets. Most recently, this paradigm was generalized to heterogeneous sensing through multi-modal multi-scale multi-level fusion quadrant entropy (MMMFQE) (Wang et al., 2025b), which synergistically fuses vibration, acoustic, and thermal modalities into a unified entropy tensor, thereby improving cross-domain generalization and diagnostic confidence. Despite these significant strides in entropy-based modeling, interpretability, and data efficiency, existing approaches remain largely unprepared for environments dominated by nonstationary, broadband transient interferences–such as those induced by track irregularities, braking events, or wheel-rail impacts in high-speed railways–where impulsive noise can severely corrupt entropy estimation and mislead classifier decisions.

Motivated by the inherent limitations of current diagnostic techniques–which remain acutely sensitive to track-induced impacts and braking transients–this research introduces a rigorously formulated framework to mitigate these vulnerabilities. Existing approaches often lack robustness in regimes dominated by multivariate random pulse noise and low signal-to-noise ratios; furthermore, canonical decompositions (e.g., wavelet-, EEMD-, and VMD-based pipelines) are prone to mode mixing, boundary effects, and heavy reliance on expert-selected parameters, while purely data-driven deep models demand large labeled corpora and frequently lack transparency, thereby complicating certification and deployment under nonstationary operating conditions. To overcome these constraints, we propose Transient-Aware Spectral Segmentation (TASS), informed by energy-distribution modeling and complemented by optimized frequency-band selection, specifically engineered to suppress the broadband, nonstationary interference produced by transient disturbances in railway bogie bearing signals. Specifically, we first develop the Amplitude-Weighted Ordinal Entropy (AWOE) algorithm, which reweights ordinal patterns by amplitude to accurately localize transient noise in the frequency domain, thereby remedying the amplitude insensitivity–and associated blind spots–of conventional permutation entropy. Building on this characterization, we introduce an entropy-guided decomposition that fuses spectral-envelope information with entropy-derived distributions, yielding a physically interpretable segmentation that is demonstrably robust to impulsive disturbances and suitable for subsequent feature construction. In parallel, we employ Particle Swarm Optimization (PSO) driven by a novel Enhanced Spectral Energy Ratio (ESER) criterion to autonomously identify frequency bands that concentrate defect-related energy around characteristic frequencies and their harmonics while penalizing broadband transient energy, thus enabling targeted feature extraction and systematically suppressing spurious contributions. Collectively, the proposed framework enhances effective SNR, improves detection reliability, and reduces false alarms across operating regimes and sensing modalities (onboard and wayside), while preserving interpretability for engineering assurance and facilitating integration with condition-based maintenance policies.

The principal contributions of this research are threefold:

1. Proposed a transient-aware spectral segmentation framework (TASS) informed by energy-distribution modeling to suppress broadband, nonstationary transient interference in bogie bearing signals, thereby improving robustness under low-SNR, nonstationary operating conditions while preserving physical interpretability.
2. Developed Amplitude-Weighted Ordinal Entropy (AWOE) to accurately localize transient noise in the frequency domain, overcoming the amplitude insensitivity of conventional permutation entropy and enabling an entropy-guided spectral decomposition that fuses spectral-envelope and entropy-derived information.
3. Introduced an optimized band-selection scheme using particle swarm optimization (PSO) driven by an Enhanced Spectral Energy Ratio (ESER) criterion, which autonomously identifies fault-sensitive frequency bands, enhances effective SNR and detection reliability, and reduces false alarms across operating regimes and sensing modalities.

This study is organized as follows: Section 2 introduces the theoretical foundations, including permutation entropy, that underpin the subsequent analyses; Section 3 then presents the Transient-Aware Spectral Segmentation (TASS) framework, comprising entropy-guided decomposition and evolutionary optimization of frequency bands; and Section 4 evaluates the approach through two case studies designed to capture distinct operating conditions and random transient fluctuation characteristics. Specifically, Case Study I (Section 4.1) employs a railway bogie dataset to assess performance under weak transient impulses embedded in strong background noise in high-speed railway bearing vibration signals, whereas Case Study II (Section 4.2) utilizes a second railway bogie dataset to further examine the method's effectiveness and generalizability under varying experimental conditions. Collectively, these components substantiate the practical applicability of the proposed framework while maintaining coherence with its theoretical basis.

2. Preliminaries

To facilitate the interpretation of the mathematical formulations utilized throughout this work –encompassing the fundamental Permutation Entropy framework and the proposed TASS methodology–a comprehensive list of symbols and abbreviations is provided in Table 1.

2.1. Permutation entropy

Permutation Entropy (PE) (Bandt & Pompe, 2002) is a widely adopted nonlinear complexity measure based on the relative ordering of samples rather than absolute amplitudes. Being model-free, PE effectively detects subtle regime shifts and nonlinear behaviors, making it valuable for mechanical fault diagnosis. By mapping signal segments to permutation symbols in a reconstructed phase space, it captures local dynamics efficiently. This ordinal approach is robust against noise and computationally fast–essential for low SNR conditions. Thus, PE serves as a principled baseline for the amplitude-sensitive extensions introduced later in this work.

Formally, consider a discrete time series $y = y_1, y_2, \dots, y_N$ of length N. The computation of PE proceeds through three steps:

1. **Phase-space reconstruction**: For an embedding dimension $m \in \mathbb{N},\ m \geq 2$ and a time delay $\tau \in \mathbb{N},\ \tau \geq 1$, construct the sequence of m-dimensional delay vectors

$$\mathbf{Y}_t^{(m,\tau)} = \left[y_t,\ y_{t+\tau},\ \dots,\ y_{t+(m-1)\tau}\right]^\top, \quad t = 1, 2, \dots, M. \tag{1}$$

where $M = N - (m-1)\tau$ is the number of admissible vectors.

2. **Ordinal symbolization**: For each vector $\mathbf{Y}_t^{(m,\tau)}$, determine the permutation $\pi_t \in S_m$ (the symmetric group of order m) that sorts its components in ascending order. To ensure a well-defined mapping in the presence of ties–common in quantized or noisy data–we adopt a deterministic tie-breaking scheme. A convenient approach is index-priority tie-breaking or, equivalently, an infinitesimal jitter model:

$$\tilde{y}_k = y_k + \epsilon\, k, \quad 0 < \epsilon \ll 1, \tag{2}$$

so that $\pi_t = \operatorname{argsort}\left(\tilde{y}_t, \tilde{y}_{t+\tau}, \dots, \tilde{y}_{t+(m-1)\tau}\right)$.

3. **Entropy calculation**: Let $p(\pi)$ denote the empirical probability of observing permutation π across the M vectors:

$$p(\pi) = \frac{1}{M} \sum_{t=1}^{M} \mathbf{1}\{\pi_t = \pi\}, \qquad \sum_{\pi \in S_m} p(\pi) = 1. \tag{3}$$

The permutation entropy is then

$$H_{\mathrm{PE}}(y; m, \tau) = -\sum_{\pi \in S_m} p(\pi)\, \log p(\pi), \tag{4}$$

where the logarithm base can be chosen as e or 2 (natural or bits). For comparability across m, the normalized PE is

$$H^*_{\mathrm{PE}}(y; m, \tau) = \frac{H_{\mathrm{PE}}(y; m, \tau)}{\log(m!)} \in [0, 1]. \tag{5}$$

Impulsive components–appear as abrupt, short-duration amplitude excursions that perturb the ordinal structure of embedded vectors. These perturbations systematically alter the frequency of certain permutations while suppressing those associated with more regular, deterministic dynamics, thereby shifting the empirical distribution $p(\pi)$. As a result, PE responds to changes in temporal organization rather than energy alone, which enables the localization of weak impulsive events embedded in heavy background noise through temporal or windowed PE tracking (Azami & Escudero, 2016). Concretely, a windowed estimate centered at time t with window length L can be defined as

$$H_{\mathrm{PE}}(t; m, \tau, L) = -\sum_{\pi \in S_m} p_t(\pi; L)\, \log p_t(\pi; L), \tag{6}$$

where $p_t(\pi; L)$ is computed using the subset of delay vectors whose indices fall within the window, thereby yielding a time-resolved complexity profile that accentuates transient disruptions.

3. Transient-aware spectral segmentation (TASS): Entropy-guided decomposition and evolutionary frequency band optimization

To address the challenges posed by low Signal-to-Noise Ratio (SNR) environments–and in particular the complex, broadband interference induced by transient disturbances–this research proposes Transient-Aware Spectral Segmentation (TASS), which is informed by energy-distribution modeling and complemented by optimized frequency band selection. Specifically, Section 3.1 develops the Amplitude-Weighted Ordinal Entropy (AWOE) algorithm, which accurately localizes transient noise in the frequency domain and thereby overcomes the amplitude insensitivity (and associated blind spots) inherent to conventional permutation entropy. Building upon this characterization, Section 3.2 introduces

Table 1

List of mathematical symbols and abbreviations.

Sym.	Desc.	Sym.	Desc.	Sym.	Desc.
1. Preliminaries					
y	Discrete time series	N	Time series length	m	Embedding dimension
τ	Time delay	$\mathbf{Y}_t^{(m,\tau)}$	Delay embedding vector	M	Count of delay vectors
S_m	Symmetric group	π_t	Ordinal pattern	$\tilde{y}_k$	Jittered time series
ϵ	Jitter parameter	$p(\pi)$	Permutation probability	H_{PE}	Permutation Entropy
H_{PE}^*	Normalized PE	L	Window length	$p_t(\pi; L)$	Local permutation prob.
2. AWOE-driven adaptive noise-band identification and mitigation					
$S(t,f)$	STFT matrix	f_s	Sampling frequency	f_k	Frequency bin index
T	Total time points	$\tilde{s}(t,f_k)$	Standardized magnitude	μ_k	Mean spectral magnitude
σ_k	Std. dev. of magnitude	ε	Stabilizer constant	$Y_t^{m,\tau}(f_k)$	Embedded spectral vector
$\mathcal{C}_t$	Amplitude contribution	α_{AWOE}	AWOE tuning parameter	$\mathcal{A}_t$	Local energy term
$\mathcal{V}_t$	Peak-to-peak variability	$\mathcal{D}_t$	Local gradient energy	$y_t^{(i)}$	ith vector component
γ	Power weighting param.	$P(\pi)$	Weighted probability	$\Psi(f_k)$	AWOE profile
$\bar{\Psi}(f)$	Normalized AWOE profile	$\Psi_{\mathrm{min/max}}$	Min/Max AWOE values	$\mathcal{L}_\eta$	Integrated entropy functional
η	Regularization parameter	$[f_s, f_e]$	Anomalous interval bounds	μ_Ψ	Mean entropy profile
ϵ_{AAPE}	Threshold parameter	$\Delta_f^{\pm}$	Difference operators		
3. Morphological-guided spectral-envelope decomposition					
W_{opt}	Optimal window length	$\mathbb{W}$	Candidate window sizes	$\mathcal{E}(f)$	Morphological envelope
$\mathbb{B}(W)$	Window band boundaries	$\mathbb{I}_{[f_a,f_b]}$	Band indicator function	GISE	Gini Index of Sq. Env.
$x_\beta(t)$	Time-domain component	W_0	Initial window length	ξ_i	Spectral peak locations
N_p	Spectral peak count	N_f	Total frequency bins	$P_1(f_k)$	Unilateral power spectrum
Δ	Dilation half-width	$\mathcal{S}(f_k)$	Smoothed envelope	W_s	Smoothing window size
Γ	Smoothing half-width	$\mathcal{R}_i$	Fixed protected boundaries	$\mathcal{D}$	Adaptive boundary candidates
N_R	Protected region count	τ_e	Expansion threshold	τ_m	Merge threshold
$\mathcal{C}_k$	Boundary cluster	$\mathcal{B}_{\mathrm{final}}$	Fused boundaries	$\mathbb{B}$	Final frequency bands
$\mathbf{SE}$	Squared Envelope sequence				
4. PSO-Guided ESER Optimization					
$\mathbf{Y}$	Swarm population matrix	n	Particle count	D	Search space dimension
$\mathbf{Y}_i$	Particle position	$\mathbf{V}_i$	Particle velocity	l	Current iteration
$\mathbf{P}_i$	Personal best position	$\mathbf{P}_g$	Global best position	ω	Inertia weight
c_1, c_2	Learning factors	r_1, r_2	Random numbers $[0,1]$	ESER	Enhanced Spectral Energy Ratio
$C_{\mathrm{es}}(n)$	Envelope spectrum	f_{ch}	Fault characteristic freq.	h	Harmonic order
TF_h	Harmonic freq. band				

an entropy-guided decomposition framework that fuses spectral envelope information with entropy-derived distributions to achieve spectral segmentation with clear physical interpretability as well as strong robustness to disturbances. Finally, Section 3.3 employs particle swarm optimization (PSO) guided by a novel Enhanced Spectral Energy Ratio (ESER) criterion to autonomously identify frequency bands that are most susceptible to fault signatures, thus enabling targeted feature extraction while systematically suppressing spurious contributions. Collectively, this coordinated strategy substantially improves the robustness and accuracy of bearing fault diagnosis under low-SNR conditions, and the comprehensive architectural design of the proposed approach is illustrated in Fig. 1.

3.1. AWOE-driven adaptive noise-band identification and mitigation

This subsection delineates the development of a frequency-domain transient-noise localization technique that pivots on a novel amplitude-aware entropy measure and, consequently, augments the separability of fault-induced transients from structured noise. Conventional spectral band-selection methods frequently fail to distinguish deterministic or semi-deterministic transient signatures from broadband interference. By contrast, our approach slices the time-frequency representation into individual frequency bins and quantifies the temporal irregularity of each bin's amplitude trajectory; in doing so, it constructs a frequency-resolved contamination map that explicitly guides subsequent denoising and decomposition.

The broadband nature of transient interference poses a significant challenge in frequency-domain analysis because its energy spreads across a wide range of frequencies and often masks fault-related spectral components. Unlike narrowband interference, such transients cannot be mitigated through simple bandpass filtering without distorting the underlying signal structure. It is therefore essential to quantitatively assess the influence of transient disturbances on the spectral content and to identify–rather than blindly suppress–frequency bands corrupted by these nonstationary artifacts.

To address this challenge, we propose a time-frequency analysis framework centered on Amplitude-Weighted Ordinal Entropy (AWOE), an enhanced entropy measure that integrates ordinal dynamics with amplitude information for robust transient characterization. Conventional Permutation Entropy (PE) relies solely on the rank-order structure of phase-space vectors, rendering it insensitive to variations in magnitude–an inherent limitation when separating deterministic transients from random fluctuations. In contrast, AWOE improves pattern discrimination by weighting permutation probabilities with a composite amplitude-aware contribution metric, thereby preserving critical energy-related information while retaining the noise robustness of ordinal analysis.

The methodology begins with a Short-Time Fourier Transform (STFT), yielding a complex-valued representation $S(t,f)$, where t denotes time and f frequency; only positive frequencies (0 to $f_s/2$) are retained. For each frequency bin f_k, we extract the temporal evolution of the spectral magnitude $\{|S(t,f_k)|\}_{t=1}^{T}$ and treat it as an independent one-dimensional time series. To ensure cross-frequency comparability and mitigate scale disparities, we standardize each sequence to zero mean and unit variance:

$$\begin{aligned}\tilde{s}(t,f_k) &= \frac{|S(t,f_k)| - \mu_k}{\sigma_k + \varepsilon},\\ \mu_k &= \frac{1}{T}\sum_{t=1}^{T} |S(t,f_k)|,\end{aligned} \tag{7}$$

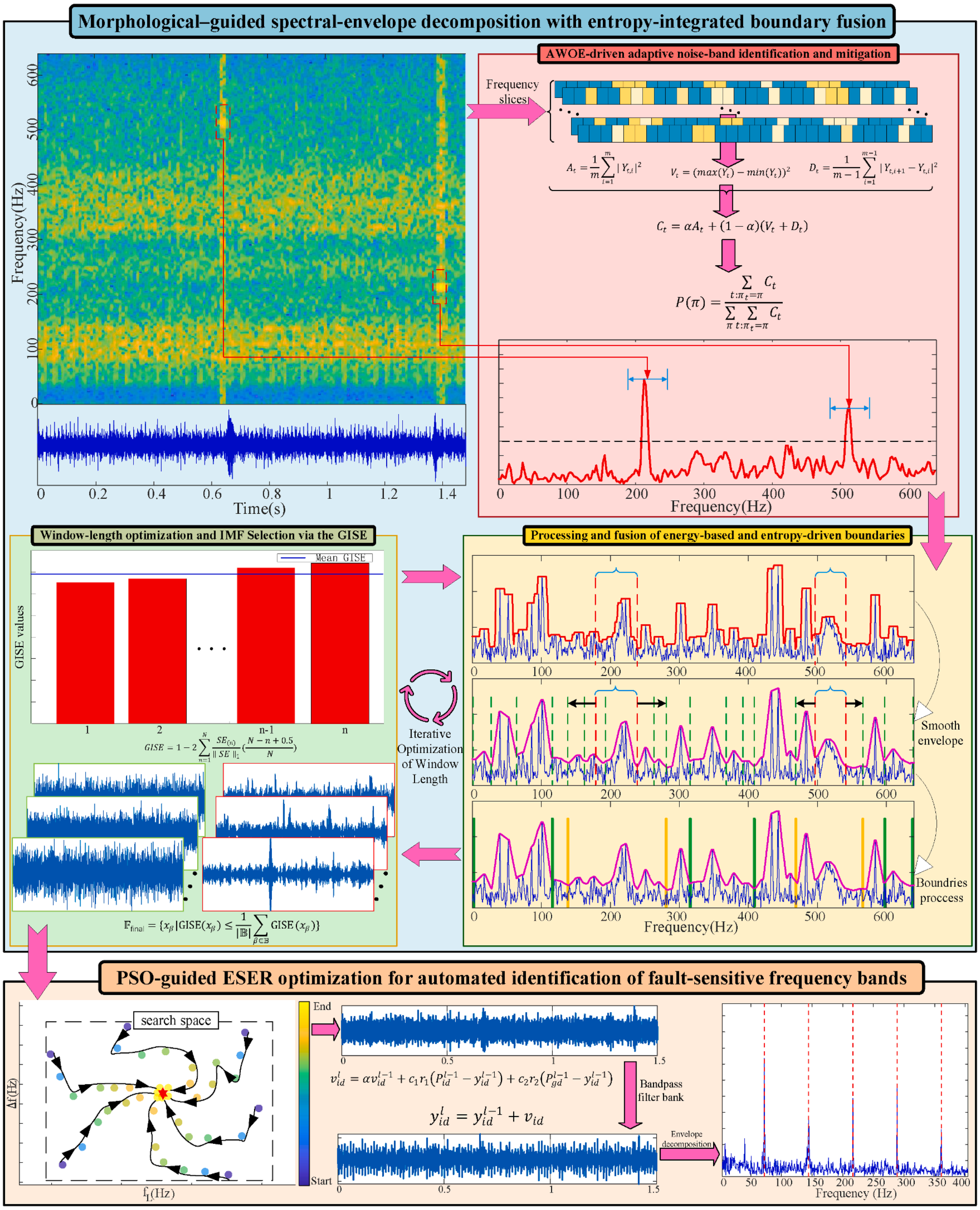


Fig. 1. Overall architecture of the TASS framework: (i) Morphological-guided spectral-envelope decomposition with entropy-integrated boundary fusion, (ii) ESER-PSO band optimization for targeted feature extraction under random disturbances.

$$\sigma_k^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(|S(t,f_k)| - \mu_k\right)^2,$$

with a small numerical stabilizer $\varepsilon > 0$. Phase-space reconstruction is then performed for each frequency bin via delay embedding:

$$Y_t^{m,\tau}(f_k) = \left[\tilde{s}(t,f_k),\ \tilde{s}(t+\tau,f_k),\ \ldots,\ \tilde{s}(t+(m-1)\tau,f_k)\right]^{\top}, \tag{8}$$

where m and τ denote the embedding dimension and delay, respectively.

AWOE is computed independently for each $\tilde{s}(t,f_k)$, producing a frequency-resolved entropy profile $\Psi(f_k)$ that serves as a quantitative indicator of temporal irregularity and impulsive content. The core innovation lies in an amplitude-sensitive contribution function. Given the embedded vector $Y_t^{m,\tau}$, we define

$$C_t = \alpha\underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(y_t^{(i)}\right)^2}_{\mathcal{A}_t} + (1-\alpha)\left[\underbrace{\left[\max(Y_t^{m,\tau}) - \min(Y_t^{m,\tau})\right]^2}_{\mathcal{V}_t} + \underbrace{\frac{1}{m-1}\sum_{i=1}^{m-1}\left(y_t^{(i+1)} - y_t^{(i)}\right)^2}_{\mathcal{D}_t}\right], \tag{9}$$

where $y_t^{(i)}$ represents the ith scalar component of the embedded vector, $\mathcal{A}_t$ measures local mean-squared amplitude (instantaneous energy), $\mathcal{V}_t$

captures peak-to-peak variability (sensitivity to abrupt jumps), and $\mathcal{D}_t$ quantifies local gradient energy (rapid transitions). The tuning parameter $\alpha \in [0, 1]$ balances energy preservation against transient sensitivity which is 0.2 in this paper.

Let π_t denote the ordinal pattern induced by $Y_t^{m,\tau}$. We define an amplitude-weighted pattern probability using a parameter $\gamma > 0$:

$$P(\pi) = \frac{\sum_{t:\, \pi_t=\pi} C_t^{\gamma}}{\sum_{\pi'} \sum_{t:\, \pi_t=\pi'} C_t^{\gamma}}, \quad \sum_{\pi} P(\pi) = 1, \tag{10}$$

and compute a normalized AWOE value

$$\Psi(f_k) = -\frac{1}{\log(m!)} \sum_{\pi} P(\pi) \log P(\pi) \ \in [0, 1], \tag{11}$$

which facilitates cross-frequency comparison. We further apply min-max normalization to obtain

$$\begin{aligned} \bar{\Psi}(f) &= \frac{\Psi(f) - \Psi_{\min}}{\Psi_{\max} - \Psi_{\min} + \varepsilon}, \\ \Psi_{\min} &= \min_f \Psi(f), \qquad \Psi_{\max} = \max_f \Psi(f). \end{aligned} \tag{12}$$

Lower values of $\bar{\Psi}(f)$ indicate heightened temporal irregularity and are indicative of transient contamination; thus, $\bar{\Psi}(f)$ provides a sensitive indicator of spectral regions affected by transients and, in turn, guides the selection of cleaner bands for diagnostic analysis.

To achieve automated and precise identification of contaminated frequency bands, we define an integrated entropy functional over an interval $[f_a, f_b]$:

$$\mathcal{L}_{\eta}(f_a, f_b) = \int_{f_a}^{f_b} \bar{\Psi}(f)\, df \ + \ \eta\, (f_b - f_a), \tag{13}$$

where $\eta \geq 0$ explicitly regularizes bandwidth and promotes compact, low-regularity regions. The optimal anomalous interval is then obtained by

$$\begin{aligned} &[f_s, f_e] = \arg \min_{f_a < f_b} \ \mathcal{L}_{\eta}(f_a, f_b), \\ &\text{subject to} \quad \frac{d}{df} \bar{\Psi}(f_a) < 0, \quad \frac{d}{df} \bar{\Psi}(f_b) > 0. \end{aligned} \tag{14}$$

which enforces that detected bands are bounded by descending and ascending slopes of $\bar{\Psi}(f)$, respectively–thereby isolating local minima (valleys) that correspond to high transient activity.

In practice, the optimization is implemented in two stages. First, all local minima of $\bar{\Psi}(f)$ are identified. A local minimum at frequency f_k is classified as an anomalous minimum if it satisfies the threshold criterion:

$$\bar{\Psi}(f_k) < \mu_{\Psi} - \epsilon_{\mathrm{AAPE}} \tag{15}$$

where $\mu_{\Psi} = \mathrm{mean}(\bar{\Psi}(f))$ and ϵ_{AAPE} is a tunable sensitivity parameter with a value of 0.25 based on experience in this paper that controls the detection threshold. This step filters out statistically significant dips in entropy associated with transient disturbances. Second, starting from each anomalous minimum, we expand the interval bidirectionally using discrete gradient tests until the slope constraints in Eq. (14) are violated:

$$\Delta_f^- \bar{\Psi}(f_s) \geq 0 \quad \text{and} \quad \Delta_f^+ \bar{\Psi}(f_e) \leq 0, \tag{16}$$

where Δ_f^- and Δ_f^+ denote backward and forward finite differences, respectively. This adaptive expansion tightly encloses regions of sustained irregularity while avoiding over-segmentation and retaining sensitivity to narrowband transients. The resulting segments $\{[f_{s_i}, f_{e_i}]\}$ delineate spectral zones most affected by transient noise and, as such, provide actionable guidance for subsequent conditioning–either by suppressing corrupted bands or by enhancing cleaner regions to facilitate reliable feature extraction.

3.2. Morphological-guided spectral-envelope decomposition with entropy-integrated boundary fusion

Conventional nonstationary decomposition methods face limitations in practical fault diagnosis. For instance, Empirical Wavelet Transform (EWT) (Gilles, 2013) relies on automated Fourier spectrum segmentation, where boundary detection—often based on local maxima or thresholds—is sensitive to noise, leading to over-partitioning or physically inconsistent modes. Similarly, Fourier Decomposition Method (FDM) yields quasi-periodic components prone to modal aliasing and lacking clear instantaneous frequency (IF) interpretation. To address these limitations, we propose a spectral decomposition framework inspired by morphological filtering–a technique widely used in image processing. Unlike linear filters, morphological filters (MF) operate by applying a structuring element to extract rank-ordered values (e.g., maximum or minimum) within a sliding window, providing robustness against impulsive noise. We extend this concept to the frequency domain by using morphological dilation to construct a smoothed spectral envelope that captures dominant energy trends while suppressing noise-induced fluctuations. Stable segmentation boundaries are identified at persistent valleys between adjacent spectral lobes, ensuring physically meaningful bandwidths.

The goal is to obtain components with interpretable and stable IF behavior. To this end, the optimal envelope window W_{opt} is selected by a global criterion that promotes sparsity and temporal localization in the component envelopes.

$$\begin{aligned} W_{\mathrm{opt}} &= \arg \max_{W \in \mathbb{W}} \left\{ \frac{1}{|\mathbb{B}(W)|} \sum_{\beta \in \mathbb{B}(W)} \Big(\mathrm{GISE}(x_{\beta}) - \mu_{\mathrm{GISE}}\Big)^2 \right\}, \\ &\quad \text{where } X(f) = \mathcal{F}\{x(t)\}, \\ &\quad x_{\beta}(t) = \Re\left\{ \mathcal{F}^{-1}\Big(X(f) \cdot \mathbb{I}_{[f_a, f_b]}(f) \Big) \right\}, \\ &\quad \beta \equiv [f_a, f_b] \in \mathbb{B}(W). \end{aligned} \tag{17}$$

Here, $\mathbb{W}$ denotes the set of candidate window sizes for MF-based envelope estimation; $\mathbb{B}(W)$ is the set of band boundaries derived from $\mathcal{E}_W(f)$; $\mathbb{I}_{[f_a, f_b]}$ is the passband indicator; and $\mathrm{GISE}(\cdot)$ denotes the Gini Index of the Squared Envelope (see Section 3.2.2). By maximizing the average GISE across reconstructed components, the method encourages spectrally coherent bands that yield temporally concentrated envelopes–thereby improving mode separability, mitigating aliasing, and enhancing the physical interpretability of the resulting intrinsic mode functions.

3.2.1. Processing and fusion of energy-based and entropy-driven boundaries

For reproducibility and robustness in deployment, we detail the end-to-end implementation of the proposed decomposition pipeline, which integrates spectral envelope analysis with transient-noise characterization. The workflow comprises two principal stages: (i) adaptive spectral-envelope construction with optimal window selection and (ii) fusion of energy-based boundary candidates (from $\mathcal{E}_W$) with entropy-driven candidates (from AWOE-based transient localization).

We first determine a data-adaptive initial window length from the spacing of spectral peaks. Let ξ_i denote spectral-peak locations and N_p the number of peaks. Then,

$$W_0 \ = \ \frac{1}{N_p - 1} \sum_{i=1}^{N_p - 1} \left| \xi_{i+1} - \xi_i \right|. \tag{18}$$

We construct a discrete candidate set $\mathbb{W}$ by expanding around W_0 with even increments, capped relative to the frequency-bin count N_f:

$$\mathbb{W} \ = \ \left\{ W \in \mathbb{Z}^+ \ \middle| \ W = W_0 + 2k, \ \ 0 \leq k \leq \left\lfloor \frac{N_f}{20} \right\rfloor \right\}. \tag{19}$$

For each candidate $W_c \in \mathbb{W}$, we extract an envelope over the unilateral power spectrum $P_1(f_k)$:

$$\mathcal{E}(f_k) = \max_{j \in [k-\Delta, k+\Delta]} P_1(f_j), \ \ \Delta = \frac{W_c - 1}{2}. \tag{20}$$

We then smooth $\mathcal{E}(f_k)$ with an odd-length moving average,

$$\mathcal{S}(f_k) = \frac{1}{W_s} \sum_{m=-\Gamma}^{\Gamma} \mathcal{E}(f_{k+m}), \tag{21}$$

where $\Gamma = \dfrac{W_s - 1}{2}$ and $W_s = \text{odd\_floor}(1.75\,W_c)$ denotes the largest odd integer less than or equal to $1.75\,W_c$, which further reduces residual fluctuations while preserving lobe structure. Local maxima of $\mathcal{S}$ are identified and adjacent minima define preliminary energy-based boundary candidates.

To curb over-segmentation and improve physical consistency, we fuse these energy-based candidates with entropy-driven boundary proposals derived from AWOE (i.e., amplitude-weighted ordinal entropy) analysis of transient contamination. Two boundary classes are considered:

- **Class I (Fixed Boundaries):** User-specified noise or interference regions $\mathcal{R}_i = [f_{s_i}, f_{e_i}]$.
- **Class II (Adaptive Boundaries):** Algorithm-detected candidates $\mathcal{D} = \{d_1, \ldots, d_M\}$ from $\mathcal{S}(f)$ and AWOE.

Boundary refinement proceeds in three stages.

Stage 1: Protected-Zone Enforcement.

Discard adaptive candidates that fall within any protected region:

$$\mathcal{D}_{\text{stage1}} = \left\{ d_j \in \mathcal{D} \;\middle|\; d_j \notin \bigcup_{i=1}^{N_R} [f_{s_i}, f_{e_i}] \right\}. \tag{22}$$

Stage 2: Boundary Expansion.

Expand the edges of each protected region to absorb nearby adaptive candidates within distance τ_e, thereby ensuring conservative coverage of user-declared interference:

$$f'_{s_i} = \min\Big(\{f_{s_i}\} \cup \Big\{ d_j \in \mathcal{D}_{\text{stage1}} \;\Big|\; |d_j - f_{s_i}| < \tau_e \Big\} \Big), \tag{23}$$

$$f'_{e_i} = \max\Big(\{f_{e_i}\} \cup \Big\{ d_j \in \mathcal{D}_{\text{stage1}} \;\Big|\; |d_j - f_{e_i}| < \tau_e \Big\} \Big), \tag{24}$$

with $\tau_e = 0.07\,N_f$ chosen to reflect a moderate physically meaningful expansion scale relative to spectral resolution.

Stage 3: Adaptive Band Formation.

Cluster the remaining adaptive candidates using a merge threshold τ_m, applied after sorting the set in ascending order; close points are aggregated to form a single representative boundary per cluster. Specifically, let $\{d_{(1)}, d_{(2)}, \ldots, d_{(L)}\}$ denote the sorted distances. A contiguous subset C_k is formed whenever adjacent elements satisfy

$$d_{(m)} - d_{(m-1)} \le \tau_m, \tag{25}$$

where $\tau_m = 0.02\,N_f$. Each resulting cluster C_k contributes one boundary $b_k = \max(C_k)$, and the finalized boundary set is

$$\mathcal{B}_{\text{final}} = \left(\bigcup_{i=1}^{N_R} \{f'_{s_i}, f'_{e_i}\} \right) \cup \left(\bigcup_{k=1}^{N_C} \{b_k\} \right), \tag{26}$$

from which we derive the usable bands by excluding intervals wholly contained within protected zones:

$$\mathbb{B} = \left\{ [b_m, b_{m+1}] \;\middle|\; b_m, b_{m+1} \in \mathcal{B}_{\text{final}},\ [b_m, b_{m+1}] \not\subset \mathcal{R}_i \right\}. \tag{27}$$

3.2.2. *Window-length optimization and IMF selection via the gini index of the squared envelope*

Through the computational process detailed in Section 3.2.1, adaptively optimized spectral boundary points are obtained; however, despite this data-driven formulation, the approach exhibits several practical limitations. In particular, the final boundary configuration is highly sensitive to the sliding-window length, which can lead to substantial variations in the signal decomposition results across different window sizes. Moreover, some of the resulting IMFs do not exhibit the periodic impulsive characteristics that are typically associated with bearing fault signatures. Owing to the energy partitioning inherent to spectral decomposition, these components may contain only a small fraction of the total signal energy, thereby allowing residual transient noise to become relatively dominant. Consequently, such noise manifests as high-amplitude, pulse-like artifacts in the time domain and, as a result, can mislead subsequent fault diagnosis if not carefully controlled.

To address these issues and to achieve robust window-length selection during the iterative enumeration process, we adopt the Gini Index of the Squared Envelope (GISE) (Ma et al., 2025) as a quantitative criterion for evaluating the effectiveness of spectral segmentation under varying window configurations. While bearing faults generate periodic transient impulses in vibration signals, these fault signatures are often obscured by transmission-path effects and broadband background noise. Crucially, when residual pulse-like noise is concentrated within specific frequency bands, it produces pronounced impulsive behavior in the corresponding time-domain sub-signals. This noise-induced impulsiveness can be effectively quantified using the GISE, wherein a higher value indicates stronger impulsiveness and, by implication, a higher likelihood of contamination by residual transients rather than genuine fault-induced periodicity.

To mitigate the influence of such artifacts, our objective is to avoid frequency bands contaminated by substantial residual pulse noise. We therefore preferentially select decomposition schemes associated with lower GISE values, thereby enhancing the physical interpretability and diagnostic reliability of the extracted components. The Gini Index (GI) (Miao et al., 2022), originally conceived as a measure of inequality in economics, is here repurposed as a robust sparsity metric. For a sorted sequence $\vec{x} = [|x_1|, |x_2|, \ldots, |x_N|]$ with $0 \le |x_1| \le \ldots \le |x_N|$, the GI is defined as

$$\text{GI} = 1 - 2\sum_{n=1}^{N} \frac{|x_n|}{\|\vec{x}\|_1 + \varepsilon} \left(\frac{N - n + 0.5}{N} \right), \tag{28}$$

where $\|\ldots\|_1$ is the L_1-norm (sum of absolute values). In the fault-diagnosis context, we compute the Gini Index of the Squared Envelope (GISE). For a signal $\mathbf{s} = (s_1, s_2, \ldots, s_N)$ with squared envelope sequence $\mathbf{SE} = (\text{SE}_1, \text{SE}_2, \ldots, \text{SE}_N)$, the GISE is given by

$$\text{GISE} = 1 - 2\sum_{n=1}^{N} \frac{\text{SE}_n}{\|\mathbf{SE}\|_1 + \varepsilon} \left(\frac{N - n + 0.5}{N} \right), \tag{29}$$

which retains the desirable scale invariance while improving numerical robustness through ε.

During the enumeration and iterative evaluation of candidate window lengths, different spectral partitioning schemes are derived according to Eq. (27). While the ultimate objective is to select the component with the lowest GISE value, the optimization process adopts a complementary strategy: for each candidate window configuration, we maximize the variance of GISE values across all decomposed components, as formalized in Eq. (17). This enhances the discriminability between noise-dominated subcomponents (exhibiting high GISE) and those containing cleaner, structurally meaningful information (with low GISE). By amplifying this contrast, the proposed strategy not only improves the separation performance but also establishes a more distinct and reliable threshold for component selection. Subsequently, each frequency band $\beta \in \mathbb{B}$ is transformed into the corresponding time-domain component $x_\beta[n]$ via band-pass filtering followed by the inverse Fourier transform. The final retained components are then determined according to

$$\mathbb{F}_{\text{final}} = \left\{ x_\beta \;\middle|\; \text{GISE}(x_\beta) \le \frac{1}{|\mathbb{B}|} \sum_{\beta \in \mathbb{B}} \text{GISE}(x_\beta) \right\}, \tag{30}$$

which implements a principled, data-adaptive threshold anchored to the mean GISE across bands and, consequently, favors components with lower impulsiveness consistent with fault-periodic content rather than residual transient contamination.

3.3. *PSO-guided ESER optimization for automated identification of fault-sensitive frequency bands*

This section proposes an intelligent optimization strategy to accurately identify the most fault-revealing frequency band from the IMFs selected in Eq. (30). To this end, we harness the global search capability of Particle Swarm Optimization (PSO) (Allaoui et al., 2025), and we direct it using a novel fitness function, the Enhanced Spectral Energy Ratio

(ESER) (He et al., 2025a). This metric is specifically crafted to be robust against both impulsive and cyclostationary noise, thereby ensuring that the selected band maximizes the energy concentration at the fault characteristic frequency and its harmonics in the envelope spectrum.

PSO is a population-based stochastic optimization algorithm widely adopted across engineering domains due to its conceptual simplicity and efficient global search characteristics. In the PSO framework, a swarm consists of n particles, collectively denoted as $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n]$, where each particle i is characterized by its position $\mathbf{Y}_i = [y_{i1}, y_{i2}, \dots, y_{id}, \dots, y_{iD}]$ and velocity $\mathbf{V}_i = [v_{i1}, v_{i2}, \dots, v_{id}, \dots, v_{iD}]$ in a D-dimensional solution space. The position vector $\mathbf{Y}_i$ represents a candidate solution to the optimization problem, whereas the velocity vector $\mathbf{V}_i$ governs the direction and magnitude of the search update. During each iteration l, the state of every particle is updated by referencing two key attractors: the individual best position $\mathbf{P}_i = [P_{i1}, P_{i2}, \dots, P_{id}, \dots, P_{iD}]$ achieved by particle i, and the global best position $\mathbf{P}_g = [P_{g1}, P_{g2}, \dots, P_{gd}, \dots, P_{gD}]$ discovered by the swarm. The update equations for velocity and position are formulated as

$$v_{id}^{l} = \alpha\, v_{id}^{l-1} + c_1 r_1 \left(P_{id}^{l-1} - y_{id}^{l-1}\right) + c_2 r_2 \left(P_{gd}^{l-1} - y_{id}^{l-1}\right), \tag{31}$$

$$y_{id}^{l} \;=\; y_{id}^{l-1} + v_{id}^{l}, \tag{32}$$

where r_1 and r_2 are independent random numbers drawn uniformly from $[0, 1]$ (optionally sampled per dimension to enhance exploration).

In this study, the inertia weight α is set to 1, thereby maintaining momentum in the search process by modulating the impact of the forward velocity; this parameter is crucial for balancing global exploration and local exploitation. The coefficients c_1 and c_2 represent the cognitive and social learning factors, respectively, guiding each particle toward its personal best position and the global best position found by the swarm. In our implementation, both c_1 and c_2 are set to 2, following commonly adopted configurations in practice. The terms r_1 and r_2 introduce stochasticity that enhances exploratory behavior and reduces the likelihood of premature convergence. Furthermore, we utilize a swarm size of 40 particles and set the maximum number of optimization iterations to 20 rounds, which provides a favorable balance between computational efficiency and solution quality.

As discussed in Section 3.2, the vibration signal has been decomposed into physically meaningful sub-bands, with noise-contaminated components effectively suppressed; consequently, the procedure isolates frequency regions suitable for subsequent fault diagnosis. Building upon this pre-processed decomposition, the PSO algorithm is integrated with the ESER fitness function. By constraining the search space within the predefined sub-bands, the PSO-ESER framework efficiently identifies the optimal frequency band in which fault-related energy is most concentrated. Once the sensitive frequency band is determined, a band-pass filter is applied to extract the corresponding signal component; the filtered signal is then subjected to envelope demodulation to reveal the underlying impulsive signatures induced by localized bearing faults, thereby enabling precise fault identification.

The ESER metric is specifically designed to enhance resilience against both impulsive and periodic interferences by emphasizing energy concentration at the fault characteristic frequency and its harmonics. It is defined as

$$\mathrm{ESER} \;=\; \frac{\displaystyle\sum_{h=1}^{10} \sum_{n=2}^{N/2} \left|C_{\mathrm{es}}(n)\right|^2 \, \mathbb{I}\!\left[\, n \in TF_h \,\right]}{\displaystyle\sum_{n=2}^{N/2} \left|C_{\mathrm{es}}(n)\right|^2 \;+\; \varepsilon}, \tag{33}$$

where $C_{\mathrm{es}}(n)$ denotes the envelope spectrum of the filtered sub-signal; h indexes the harmonic order of the theoretical fault characteristic frequency f_{ch}; TF_h represents a narrow frequency band centered at $h \cdot f_{\mathrm{ch}}$ with lower and upper bounds (a, b) in frequency-bin indices; $\mathbb{I}[\cdot]$ is the indicator function; N is the signal length; and $\varepsilon > 0$ is a small constant added for numerical stability. The computation of ESER begins with estimating f_{ch} based on the shaft rotational speed and bearing geometric parameters. The numerator aggregates the squared envelope-spectrum magnitudes over ten harmonic bands of f_{ch}, thereby capturing fault-related energy, while the denominator normalizes by the total spectral energy across the Nyquist range. This ratio enhances sensitivity to incipient faults by amplifying periodic fault signatures and, at the same time, suppressing contributions from broadband noise and non-fault-related components. Furthermore, the use of narrow, harmonically aligned bands renders the metric inherently robust to minor speed fluctuations–common in real-world applications where precise tachometer data may be unavailable–thus reducing the risk of misdiagnosis and improving reliability in practice.

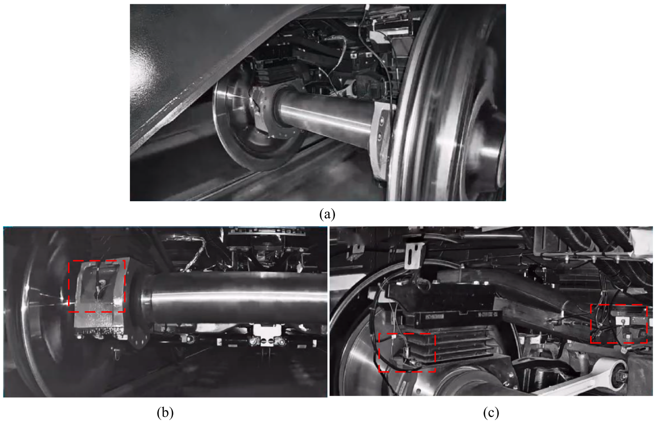

(a)
(b)
(c)

Fig. 2. (a) Instrumentation layout of the Siemens-platform bogie (SF500 series) test rig (Chen et al., 2026b). The setup utilizes a swing-arm positioning device characteristic of high-speed railway applications. The red dashed regions highlight the key sensor locations: (b) the accelerometer mounted on the axle box, and (c) the accelerometer mounted on the bogie frame side beam. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

By combining ESER as the fitness criterion with PSO operating within a reduced solution space, the proposed method achieves an advantageous trade-off between computational efficiency and diagnostic performance. The constrained search over sub-bands accelerates PSO convergence and avoids entrapment in local optima, thereby enabling accurate identification of the most sensitive frequency band for bearing fault detection. Consequently, this integrated framework is robust, adaptive, and automated, making it particularly suitable for early and reliable fault diagnosis in complex industrial environments.

4. Experimental validation and comparative analysis across railway and wind subway platforms

To rigorously validate the proposed methods, two representative case studies have been employed in this section so as to encompass distinct operating environments and random transient fluctuations characteristics. All experiments were conducted on a computer equipped with an Intel(R) Core(TM) Ultra 7 155H processor (3.80 GHz), 32.0 GB RAM, and an NVIDIA GeForce RTX 4060 Laptop GPU. First, a railway bogie dataset is utilized to assess performance under weak transient impulses embedded within strong background noise in high-speed railway bearing vibration signals (see Section 4.1). To further extend and corroborate the effectiveness and generalizability of the approach, a second railway bogie dataset is employed in Case Study II (see Section 4.2), thereby enabling a complementary evaluation.

4.1. *Case study I: Real-world validation of railway bogie system*

4.1.1. *Experimental apparatus configuration and dataset description*

To rigorously evaluate the effectiveness and robustness of the proposed methodology, a full-scale experimental validation was conducted

Table 2
Geometric parameters and characteristic frequencies of the SF500 series bogie bearing.

Parameter	Symbol	Value
Roller diameter	d	26.0 mm
Pitch diameter	D_m	183.93 mm
Contact angle	α	10°
Number of rolling elements	Z	19
Shaft rotational frequency	f_r	10.18 Hz (≈106 km/h)
Ball pass frequency outer race	BPFO	83.244 Hz

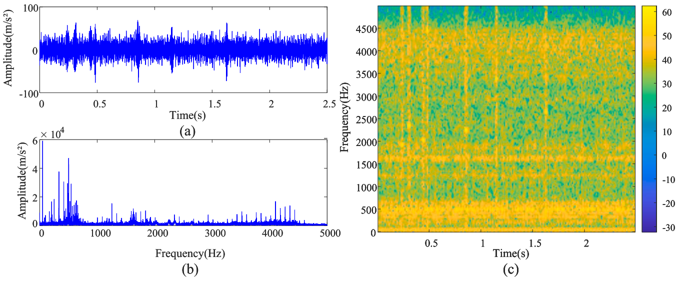


Fig. 3. Signal visualization of the Siemens-platform bogie dataset: (a) time-domain waveform exhibiting high-energy impacts; (b) frequency spectrum revealing the specific fault characteristic frequency (83.244 Hz); and (c) time-frequency representation exposing the non-stationary nature of the signal.

using a Siemens-platform bogie (SF500 series). The experimental subject features a standard wheelbase of 2500 mm and utilizes a swing-arm positioning system. The data acquisition specifically targets the bogie bearings (double-row tapered roller bearings housed within the axle box), which are the critical load-bearing components determining the bogie's operational safety.

The specific geometric and operational parameters of the tested bogie bearing are summarized in Table 2.

Vibration signals were acquired using piezoelectric accelerometers installed at two distinct locations as depicted in Fig. 2: the front cover of the axle box and the bogie frame side beam. The signals were digitized at a sampling rate of 10 kHz to ensure sufficient bandwidth for capturing the fault characteristic frequencies and their harmonics (Fig. 3).

The acquired vibration signal, which is shown in Fig. 3, reflects the operating conditions of the bogie system at the tested speed. Due to the swing-arm suspension structure, the vibration transmission path is complex, necessitating advanced signal processing to extract the specific 83.244 Hz fault signature from the background noise.

4.1.2. Implementation and diagnostic validation of the proposed TASS framework

To address weak transient impulses obscured by strong background noise in high-speed railway bearing signals, we employ an AWOE-based procedure to quantify spectral contamination. Despite their low energy, these broadband transients can generate spurious peaks or mask true fault features in the frequency domain, undermining diagnostic reliability. As shown in Fig. 4, AWOE computes slice-wise entropy across the time-frequency plane, effectively revealing anomalous energy patterns caused by transients. Mapping these entropy values onto the frequency spectrum quantifies noise pollution levels. Using the adaptive band selection framework defined by Eqs. (13) and (14), we identify frequency bands heavily contaminated by transients, yielding boundary points that guide subsequent decomposition. This enables focused analysis on cleaner signal segments for more reliable fault diagnosis.

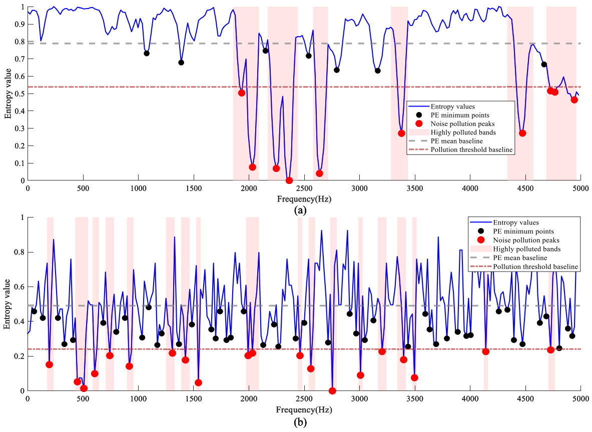


Fig. 4. (a) AWOE profile. (b) comparison with conventional Permutation Entropy (PE): AWOE reveals distinct low-entropy valleys corresponding to transient-contaminated bands, demonstrating its superior sensitivity.

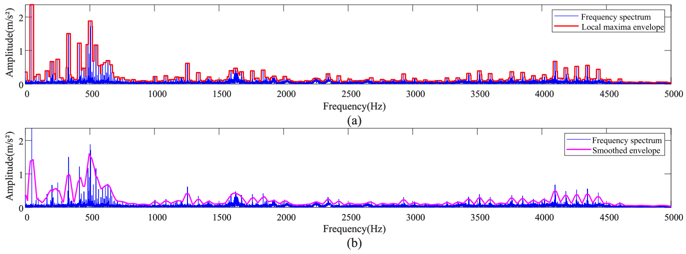


Fig. 5. Spectral envelope construction and refinement: (a) initial envelope extracted from local maxima using the empirically optimal window (width = 208), capturing dominant spectral trends while retaining residual fluctuations; (b) smoothed envelope that suppresses local irregularities and clarifies regions of energy concentration consistent with fault-related modulations.

To underscore AWOE's advantage over conventional permutation entropy (PE), Fig. 4(b) shows that PE lacks discriminative power in localizing transient noise due to its insensitivity to amplitude variations. In contrast, AWOE's amplitude-weighting scheme enhances sensitivity to low-amplitude, broadband transients, yielding a more reliable spectral contamination map for band selection.

Given the noise, fluctuations, and irrelevant components in the original Fourier spectrum, we perform spectral segmentation to extract informative sub-signals. First, a candidate window set is constructed via Eqs. (18) and (19), followed by adaptive local-extremum extraction: each window scans the spectrum to locate local maxima (Eq. (20)), forming an initial envelope. To suppress residual irregularities, dynamic smoothing is applied (Eq. (21)), producing a refined envelope that better captures the underlying energy distribution. For each window length, boundary points are placed using Eqs. (22)-(27) to generate candidate band-decomposition schemes. Each scheme is evaluated by computing GISE (Eq. (29)), and the optimal decomposition–corresponding to the window width maximizing the mean GISE across all IMFs–is selected. In this case, the optimal window width is 208, achieving a mean GISE of 0.6985. The detailed decomposition procedure is described below.

As shown in Fig. 5(a), an initial spectral envelope is built using a sliding window of length 208, which suppresses noise and ripples while preserving dominant amplitude features–such as bearing fault frequencies. This envelope is then smoothed with a wider window ($1.75 \times 208 = 364$), as depicted in Fig. 5(b), enhancing genuine spectral peaks and removing spurious spikes to yield a stable baseline for segmentation.

Local maxima of the smoothed envelope–representing core physical frequency components–are first identified. Segmentation boundaries are then placed at energy minima between adjacent maxima, ensuring that partitioning follows the signal's intrinsic energy distribution. This peak-guided valley localization strategy balances boundary rationality, noise robustness, and physical interpretability of the resulting spectral components.

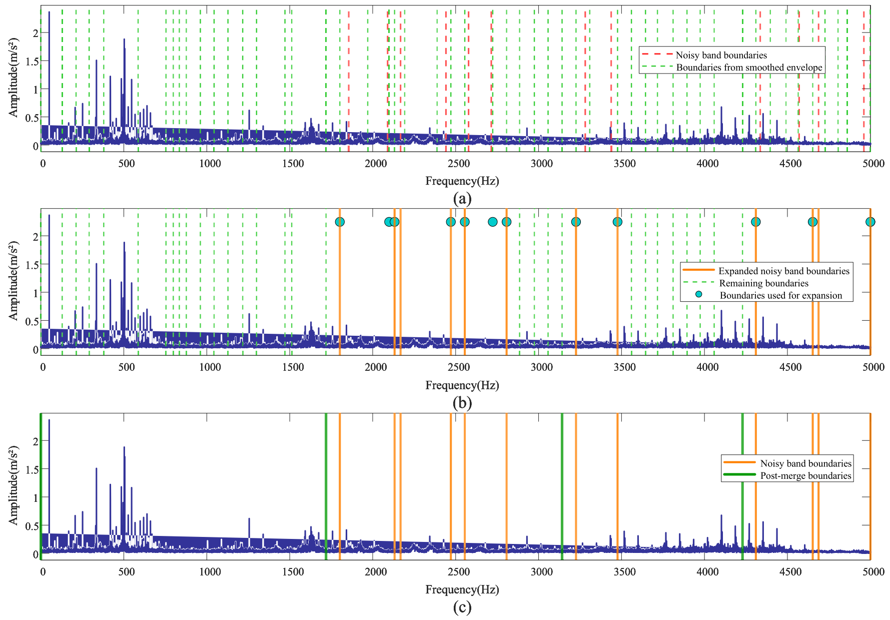


Fig. 6. Boundary construction and refinement: (a) complete set of boundary candidates, including Type I (noise-informed) and Type II (envelope-valley) boundaries; (b) processing and expansion of noise bands to establish protective isolation zones; and (c) final set of merged and physically coherent boundary points that define interpretable sub-bands suited to demodulation.

Fig. 6(a) shows two boundary types: red dashed lines (Type-I) mark noise-severely corrupted band edges, while green dashed lines (Type-II) derive from the smoothed spectral envelope. Most Type-II boundaries correctly flank dominant peaks, confirming the envelope's ability to capture global energy trends; however, some remain suboptimal due to spectral complexity.

To improve reliability, a refinement step removes any Type-II boundaries falling within Type-I-defined noise-contaminated zones (Fig. 6(b)). A protective guard band–formed by expanding a controlled buffer around each Type-I boundary–is introduced to prevent noise leakage into adjacent sub-bands during reconstruction, enhancing sub-signal purity and diagnostic relevance. Remaining Type-II boundaries are then merged via Eq. (27) to group adjacent bands likely belonging to the same fault-induced harmonic family (Fig. 6(c)), ensuring segmentation aligns with the physical modulation mechanism of bearing faults.

As shown in Fig. 7, the segmentation initially yields five meaningful sub-bands (excluding noise-dominated or spurious narrowband modes). Three are selected for fine-grained analysis based on GISE: (0, 1718.8) Hz, (2807.2, 3141.6) Hz, and (3476.0, 4229.2) Hz. To maximize fault information concentration, an iterative PSO-ESER optimization framework refines these bands. Fig. 8 shows ESER values converging monotonically to their maxima, with particles clustering tightly around the optimum–demonstrating robust convergence and validating the optimality of the selected bands.

The optimized band-pass filter bank–comprising passbands (77.5, 468.4) Hz, (2807.2, 3125.7) Hz, and (3476.0, 4228.9) Hz (Fig. 9(a))–is applied to the original signal. The filtered waveforms (Fig. 10) reveal enhanced periodic impulses, with fault-related transients previously buried in noise now clearly recovered and amplified, demonstrating robust noise suppression under low SNR.Envelope spectra of the demodulated sub-signals are coherently superimposed into a composite diagnostic spectrum (Fig. 9(b)), where distinct peaks appear at the Ball Pass Frequency of the Outer race (BPFO) and its harmonics, free from spurious interference. This confirms the method's effectiveness in extracting weak fault features under severe noise contamination.

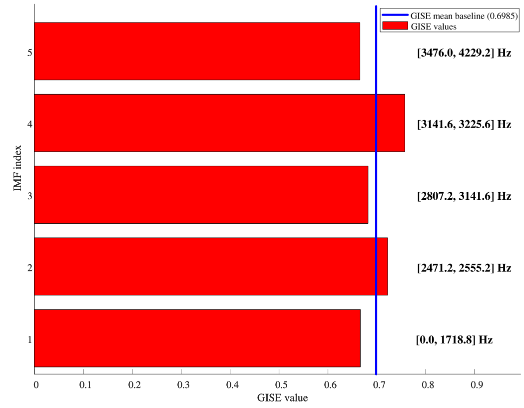


Fig. 7. GISE-based evaluation of decomposed sub-bands. Bands exhibiting lower impulsiveness (interpreted via the sparsity/inequality of the squared envelope) are prioritized for fault demodulation, whereas bands with elevated impulsiveness are de-emphasized as likely noise-dominated.

4.1.3. *Parameter justification and ablation studies*

To rigorously validate the TASS framework, we conducted systematic investigations to justify key parameter choices and quantify the necessity of each sub-module.

First, regarding the AWOE algorithm, the weighting coefficient α balances local energy against transient sensitivity, while the threshold ε determines noise detection sensitivity. Sensitivity analysis (Fig. 11) supports the selection of $\alpha = 0.2$ and $\varepsilon = 0.25$. Specifically, $\alpha = 0.2$ places the algorithm in a high-gradient region, ensuring sensitivity to low-amplitude broadband transients without over-amplifying high-energy components. Concurrently, $\varepsilon = 0.25$ corresponds to a stable plateau in the detection count (Fig. 11(b)), offering a robust statistical criterion ($\mu_\Psi - 0.25$) that minimizes false positives while reliably identifying anomalous low-entropy regions.

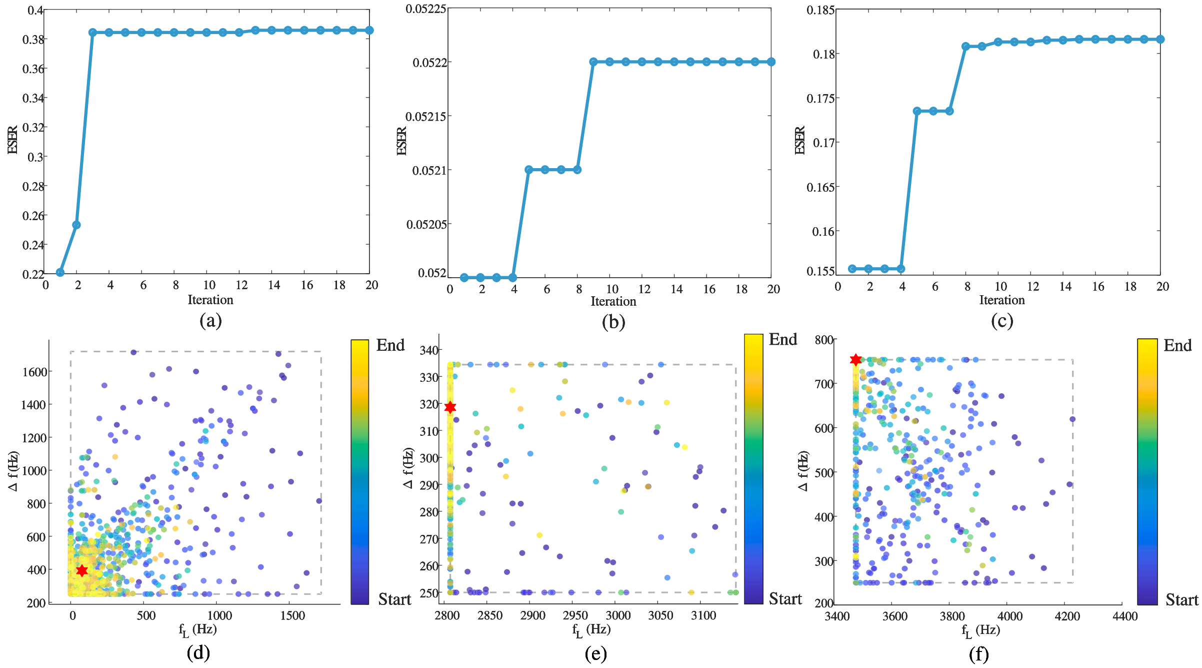


Fig. 8. PSO-driven ESER optimization within each candidate band. Curves exhibit monotonic or near-monotonic improvement of the ESER objective toward convergence: (a) (0, 1718.8) Hz, (b) (2807.2, 3141.6) Hz, (c) (3476.0, 4229.2) Hz; while conceptual swarm snapshots indicate progressive clustering around the global optimum, evidencing both stability and efficiency: (d) (0, 1718.8) Hz, (e) (2807.2, 3141.6) Hz, (f) (3476.0, 4229.2) Hz.

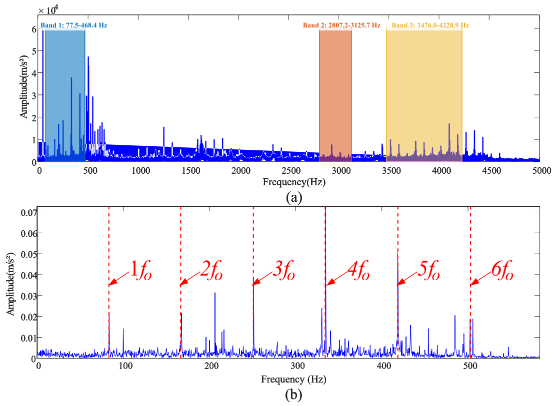


Fig. 9. Diagnostic outcomes: (a) final selected passbands identified by the PSO-ESER procedure; and (b) envelope spectra of demodulated sub-signals together with their composite diagnostic spectrum, in which distinct peaks appear at BPFO and its harmonics with minimal interference.

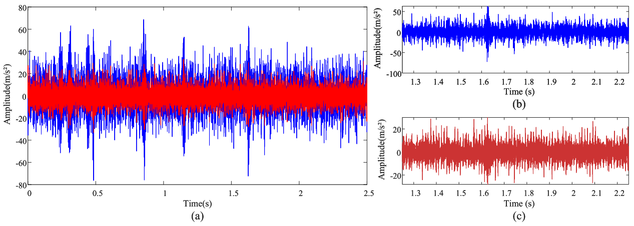


Fig. 10. Time-domain comparison: (a) raw versus filtered vibration signal, demonstrating enhanced impulsive periodicity post-filtering; (b) zoomed view of the original signal between 1.25 s and 2.25 s; and (c) corresponding zoomed view of the filtered signal over the same interval, revealing previously buried fault-induced impulses.

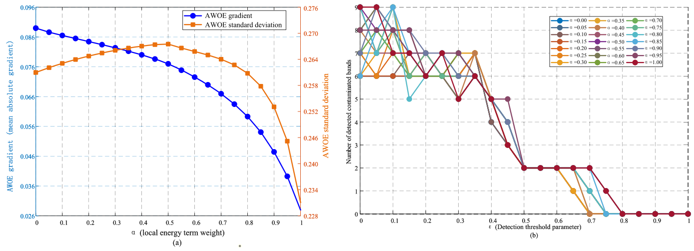


Fig. 11. Parameter sensitivity analysis for (a) weighting coefficient α and (b) threshold parameter ε.

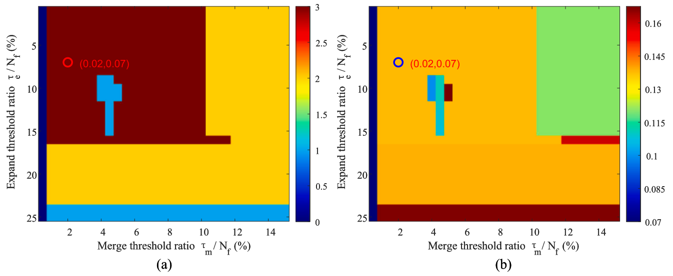


Fig. 12. Parameter sensitivity analysis for expansion threshold τ_e and merge threshold τ_m: (a) the number of selected sub-bands; (b) the mean ESER.

Second, we conducted a grid-search sensitivity analysis for the boundary expansion threshold τ_e and merge threshold τ_m over a wide range of values. As shown in Fig. 12, both the number of selected sub-bands and the mean ESER remain nearly constant when the parameters vary within a reasonable interval, confirming that the decomposition outcomes and fault feature extraction are insensitive to moderate deviations of these thresholds.

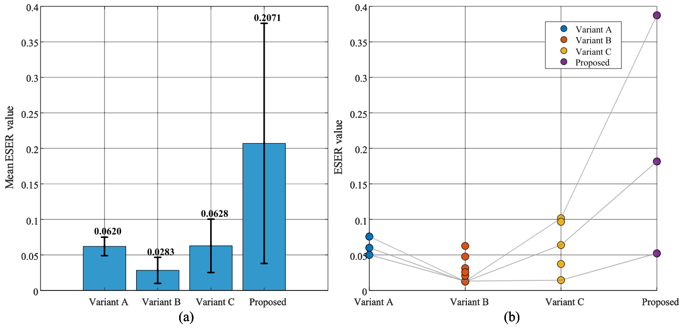


Fig. 13. Ablation results: (a) Mean ESER with std dev; (b) Band-wise ESER evolution.

Table 3
Diagnostic performance of algorithmic variants.

Variant	Mean ESER	Std Dev	Max ESER	# Bands
Var. A	0.0620	0.0130	0.0759	3
Var. B	0.0283	0.0183	0.0627	8
Var. C	0.0628	0.0376	0.1017	5
TASS	**0.2071**	**0.1690**	**0.3874**	**3**

Table 4
ESER values for intrinsic mode functions obtained via different decomposition algorithms.

	IMF1	IMF2	IMF3	IMF4	IMF5
TASS	**0.3874**	**0.0522**	**0.1816**	–	–
PSO-VMD	0.0891	0.1148	0.1768	–	–
EMD	0.0428	0.0695	0.1178	0.0507	0.0041
LMD	0.0399	0.0778	0.1161	0.0549	0.0052

Third, to assess the contribution of individual components, we performed an ablation study comparing four variants:

1. **Variant A:** PE + GISE-based IMF selection + PSO-ESER.
2. **Variant B:** PE + no GISE selection + PSO-ESER.
3. **Variant C:** AWOE + no GISE selection + PSO-ESER.
4. **Proposed:** AWOE + GISE-based IMF selection + PSO-ESER.

Diagnostic performance was evaluated using the Enhanced Spectral Energy Ratio (ESER). As summarized in Table 3 and visualized in Fig. 13, the results indicate:

1. **AWOE Superiority:** Replacing PE with AWOE (Proposed vs. Variant A) increased the mean ESER from 0.0620 to 0.2071 (a ∼234% improvement), confirming AWOE's enhanced ability to localize transients in noisy environments.
2. **GISE Utility:** GISE-based selection (Variant A vs. Variant B) significantly improved stability and reduced the number of processed bands (from 8 to 3), effectively filtering noise-dominant components.
3. **Synergistic Efficacy:** The full TASS framework achieved the highest Max ESER (0.3874) and Mean ESER (0.2071), outperforming all partial implementations. This confirms that the integration of entropy-guided boundaries and GISE-based selection is critical for robust diagnosis in low-SNR conditions.

These findings rigorously validate the proposed parameter settings and the necessity of the hierarchical TASS architecture.

4.1.4. Comparative evaluation analysis

To assess the robustness of the proposed method under strong noise, we compare it with three widely used decomposition techniques: VMD (Dragomiretskiy & Zosso, 2013), EMD (Torres et al., 2011), and LMD (Jia et al., 2019). For EMD and LMD, the number of decomposition levels is set to match that of the proposed approach (five IMFs).

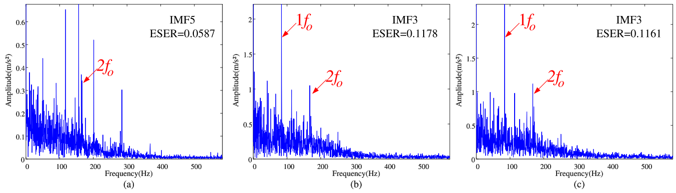


Fig. 14. Comparative envelope demodulation results across different decompositions: (a) PSO-VMD, (b) EMD, and (c) LMD.

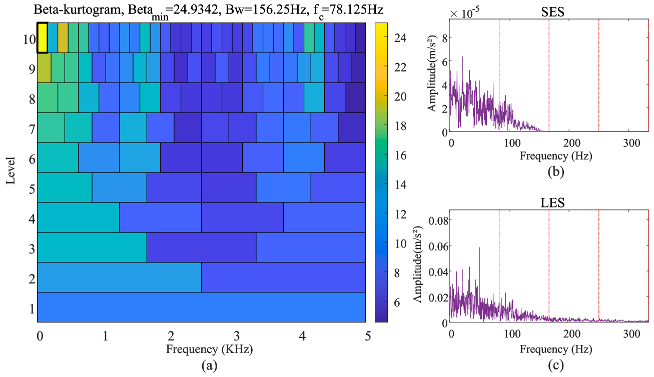


Fig. 15. Beta-kurtogram-based demodulation band selection and outcomes: (a) selected demodulation band; (b) squared envelope spectrum; and (c) logarithmic envelope spectrum. The selected band fails to exhibit BPFO-related harmonic content under the prevailing non-stationary, broadband disturbance.

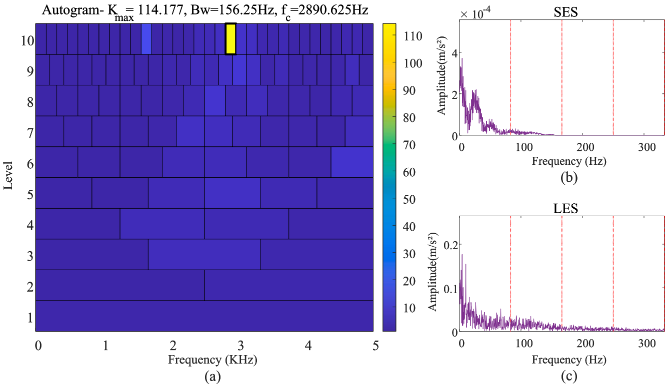


Fig. 16. Autogram-based demodulation band selection and outcomes: (a) selected demodulation band; (b) squared envelope spectrum; and (c) logarithmic envelope spectrum. Similar to the beta-kurtogram, the selected band does not reveal BPFO harmonics, indicating insufficient robustness to the observed interference.

For VMD, to ensure a fair comparison, we employ the same PSO framework as in TASS to optimize its parameters ($K \in [2, 12]$, $\alpha \in [100, 5000]$), with the objective of maximizing the mean ESER across all IMFs. The optimization yields an optimal combination of $K = 3$ and $\alpha = 3599.2$, achieving a mean ESER of 0.1269, with individual IMF ESER values of 0.0891, 0.1148, and 0.1768.

ESER is then used to quantify the fault information in each mode. As summarized in Table 4, the IMFs extracted by the proposed method exhibit the highest ESER values. For each method, the mode with the maximum ESER is selected for envelope demodulation, and the resulting spectra are shown in Fig. 14. Under strong background noise, the BPFO and its harmonics are weak or completely obscured in the spectra from VMD, EMD, and LMD, highlighting the limited robustness of conventional decomposition methods in preserving transient fault signatures in low-SNR environments, even after parameter optimization.

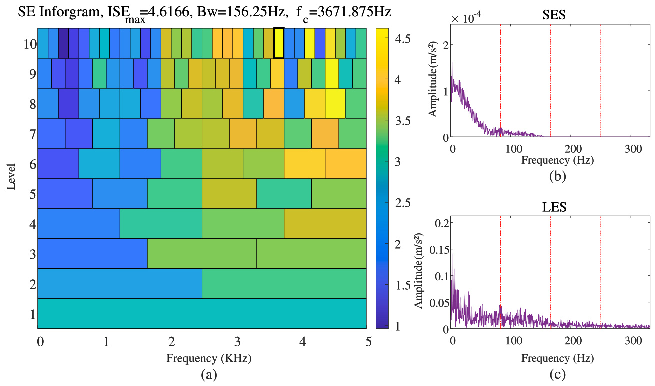


Fig. 17. Infogram-based demodulation band selection and outcomes: (a) selected demodulation band; (b) squared envelope spectrum; and (c) logarithmic envelope spectrum. The selected band fails to exhibit BPFO-related harmonic content under the prevailing non-stationary, broadband disturbance.

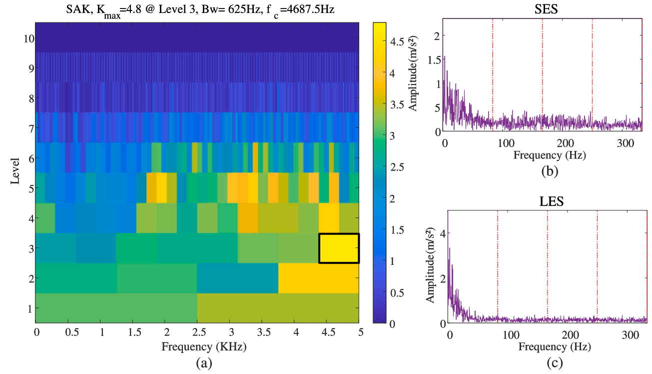


Fig. 18. SAK-based demodulation band selection and outcomes: (a) selected demodulation band; (b) squared envelope spectrum; and (c) logarithmic envelope spectrum. The selected band does not reveal BPFO harmonics, indicating insufficient robustness to the observed interference.

Furthermore, a comparative evaluation is conducted against four established band-selection methods: the beta-kurtogram, autogram, infogram, and Subband Averaging Kurtogram (SAK) (Wang et al., 2020). The optimal frequency bands identified by these approaches are $[0, 156.25]$ Hz, $[2812.5, 2968.75]$ Hz, $[3593.75, 3750]$ Hz, and $[4375, 5000]$ Hz, respectively. However, the resulting envelope spectra exhibit no significant peaks at the Ball Pass Frequency Outer race (BPFO) or its harmonics (see Figs. 15–18), indicating ineffective fault feature enhancement and failure in fault detection under these settings.This outcome, when considered alongside the preceding analysis, further highlights the superiority of the proposed method in achieving an optimal balance between frequency resolution and modulation sensitivity. By adaptively selecting bandwidths that align with the physical characteristics of fault-induced transients–while explicitly excluding noise-dominated regions through entropy-guided constraints–the proposed framework enables robust and accurate fault diagnosis even in complex, high-noise environments.

To assess the practical feasibility, we also compared the computational efficiency of these methods. The average execution times (over 10 independent runs) are listed in Table 5. Although the proposed TASS framework involves multiple processing stages and thus requires slightly more computation, its runtime remains acceptable for periodic condition monitoring. The faster methods fail to extract the fault signature, confirming that diagnostic accuracy is the primary consideration in this low-SNR scenario.

Table 5
Computational time comparison of different band-selection methods (averaged over 5 runs).

Method	Execution time (s)
TASS	3.122
Beta-kurtogram	1.612
Autogram	1.605
Infogram	0.718
SAK	0.771

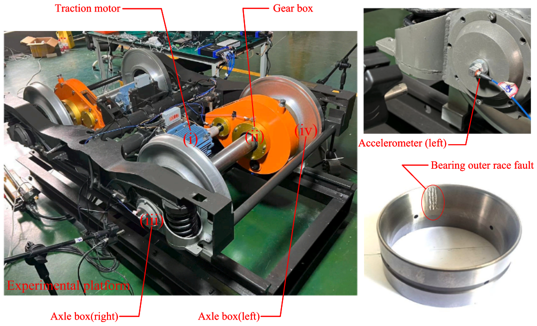


Fig. 19. Experimental platform of the subway train bogie transmission system (BJTU-RAO), illustrating the deployment of accelerometers on the axle boxes, traction motor, and gearbox.

4.2. Case study II: Validation on a subway train bogie transmission system

4.2.1. Experimental apparatus configuration and dataset description

To further corroborate the generalization capability and robustness of the proposed TASS framework under complex mechanical transmission conditions, this study employs the BJTU-RAO bogie dataset collected by the State Key Laboratory of Advanced Rail Autonomous Operation at Beijing Jiaotong University (Ding et al., 2024). The dataset is acquired from a high-fidelity experimental platform designed as a 1:2 scale isotropic model of a subway train bogie, capturing the coupled dynamic behaviors of the traction motor, driving gearbox, and axle boxes inherent to rail transit systems.

The powertrain configuration consists of a three-phase asynchronous motor driving a wheelset through a helical gearbox, with an electro-hydraulic device applied to simulate realistic operational loads. The transmission chain utilizes a reduction gear ratio of 107/16 (approximately 6.69). Comprehensive monitoring is achieved through a 24-channel sensor array recording vibration, current, and acoustic signals at a high sampling frequency of $f_s = 64$ kHz. Specifically, vibration data is acquired using tri-axial accelerometers mounted directly on the axle box end covers; the precise installation location of these accelerometers is indicated in the experimental setup diagram shown in Fig. 19.

For this case study, we focus on the diagnosis of the left axle box bearing (Model: HRB 352213) under a specific high-speed, heavy-load operational regime–Working Condition 6–characterized by a motor input frequency of 60 Hz and a lateral load of +10 kN. Based on the transmission kinematics, the axle shaft rotates at a fundamental frequency of $f_r = 8.972$ Hz. Given the bearing geometry and fault location (a localized outer race defect, labeled `LA2`), the theoretical ball pass frequency on the outer race (BPFO) is calculated as $8.071 \times f_r$, yielding approximately 72.41 Hz–a key diagnostic indicator for outer race faults in this configuration.

As illustrated in Fig. 20, the vibration signals are subject to complex modulation and attenuation due to the transmission path. The acquired signal for the left axle box exhibits non-stationary interference and broadband noise typical of bogie transmission systems, providing a rigorous test for the proposed transient-aware spectral segmentation method.

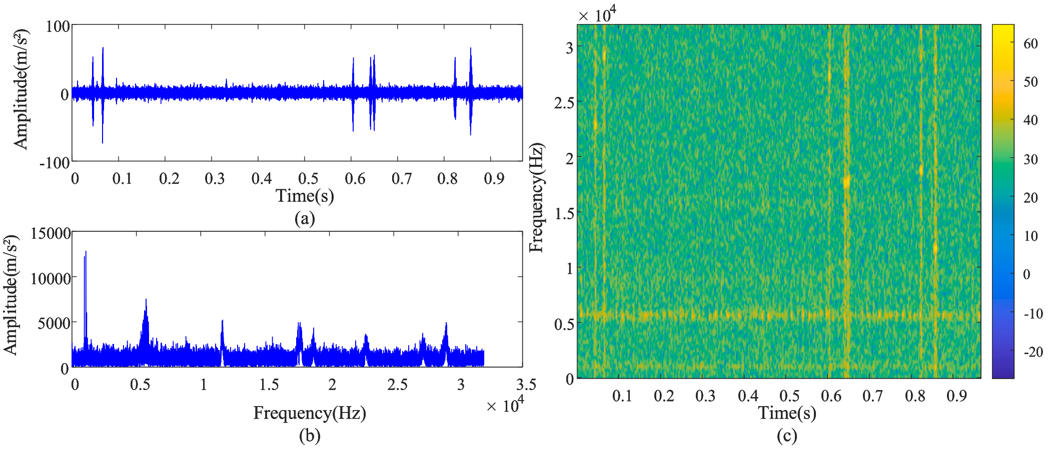


Fig. 20. Signal visualization: (a) time-domain waveform exhibiting intermittent transients and overlapping events; (b) frequency spectrum revealing broadband disturbances alongside localized clusters; and (c) time-frequency representation exposing non-stationary, broadband interference superimposed on weak periodic content.

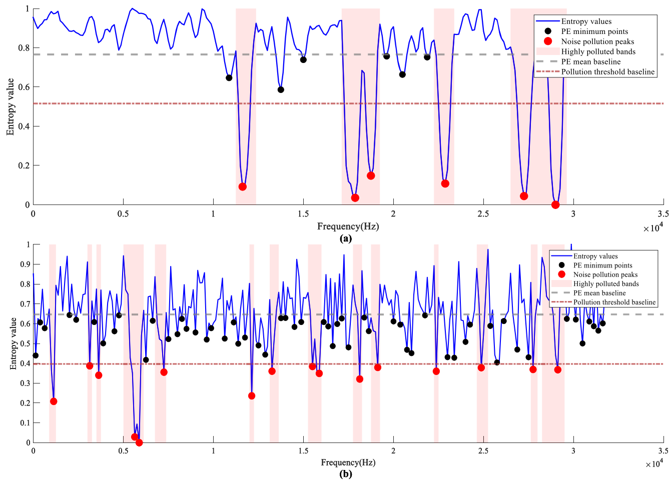


Fig. 21. (a) Frequency-resolved AWOE distribution. (b) Comparison between PE and AWOE for transient noise localization.

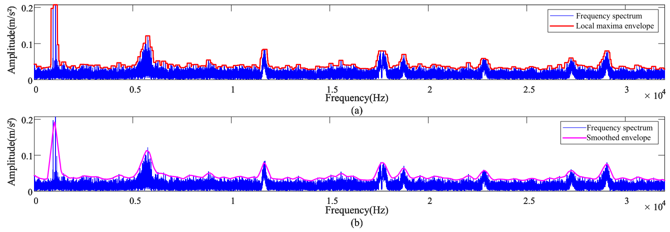


Fig. 22. (a) Spectral envelope obtained using the optimal window length 201; (b) Smoothed spectral envelope computed with a window width of 351, highlighting dominant spectral peaks while suppressing spurious ripples.

4.2.2. Implementation and diagnostic validation of the proposed TASS framework

In response to the transient noises, we propose the utilization of AWOE to quantify the degree of spectral contamination caused by these noises in the original vibration signals. As illustrated in Fig. 21(a), this modified permutation entropy captures subtle energy changes induced by such noises, identifying abnormal variations within specific bands. The distribution of these entropy values across the spectrum reflects the extent of noise pollution, and boundary points are identified to mark severely contaminated bands, guiding subsequent decomposition tasks.In contrast, conventional Permutation Entropy (PE), as illustrated in Fig. 21(b), exhibits limited sensitivity to amplitude variations and consequently fails to distinguish noise-affected frequency bands as effectively as AWOE.

Subsequently, spectral segmentation is performed to extract sub-signals across different frequency bands. A candidate window set is constructed, and an adaptive local-extremum extraction technique is applied. Each window scans the spectrum to identify local maxima, forming an initial envelope curve. This envelope is then dynamically smoothed to better reflect the underlying energy distribution. Boundary points are deployed on these envelopes to generate multiple decomposition schemes. The optimal decomposition corresponds to a window width of 201, yielding a mean GISE value of 0.7201 among all extracted IMFs.

As illustrated in Fig. 22(a), an initial spectral envelope is constructed using a sliding window of length 201, suppressing noise and secondary ripples while preserving dominant frequency components. The envelope is further smoothed using a window width of 351, enhancing meaningful spectral peaks and eliminating spurious spikes, as shown in Fig. 22(b). Segmentation boundaries are established at energy minima between adjacent maxima, guided by the physical energy distribution.

As depicted in Fig. 23(a), two types of boundary points are identified: Type-I boundaries denote bands corrupted by noise, and Type-II boundaries are derived from the smoothed envelope. Most Type-II boundaries are accurately positioned around dominant spectral peaks. To improve reliability, Type-II boundaries within noise-contaminated bands are removed. Furthermore, a protective guard band is introduced around each Type-I boundary. As shown in Fig. 23(b), this prevents noise infiltration. Remaining Type-II boundaries are merged to group adjacent sub-bands likely belonging to the same fault-related harmonic family, as shown in Fig. 23(c).

The final boundary-point configuration scheme involves extracting frequency-band signals outside the noise spectrum and discarding those that can be considered false modes. The frequency bands ultimately selected for GISE analysis are illustrated in Fig. 24. Based on the GISE values of each IMF, we select the IMFs below the GISE mean for subsequent processing. The subsequent analysis is focused on four pre-identified sub-bands: [0, 9927.2] Hz, [23375.5, 25783.7] Hz, [25783.7, 26380.4] Hz and [29624.8, 32000] Hz. Within these frequency intervals, a refined search is conducted to identify the optimal resonant band most suitable for envelope demodulation in bearing fault detection. To achieve this, the PSO algorithm is employed in conjunction with ESER.

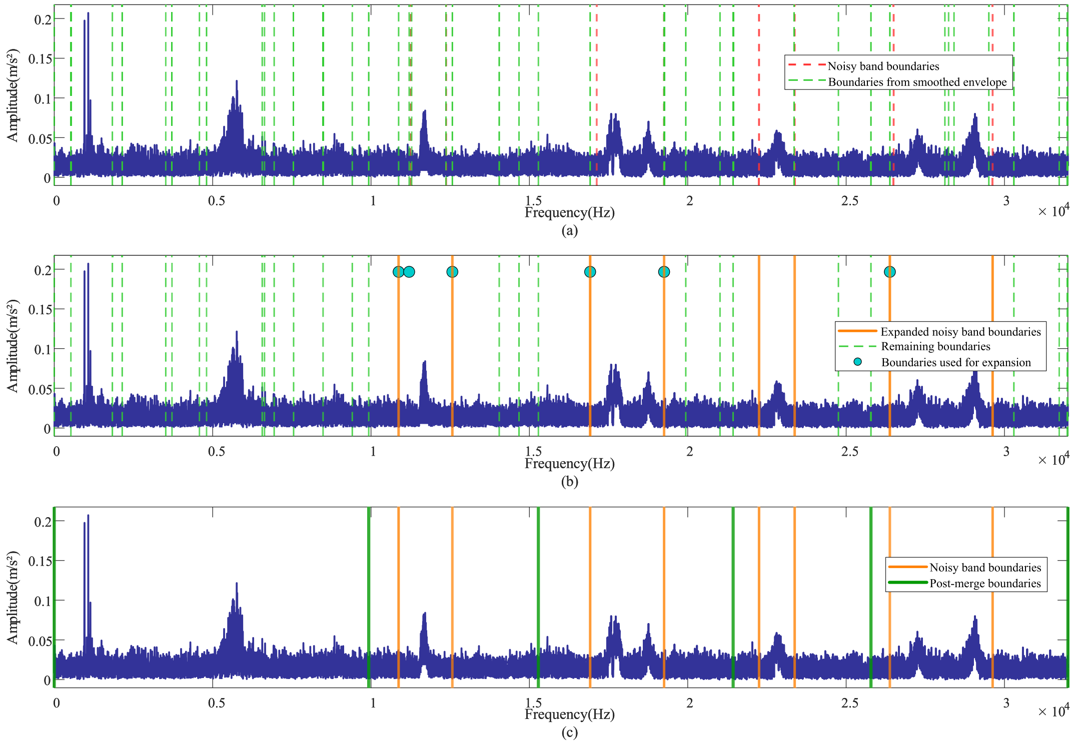


Fig. 23. (a) Total boundary set: Type-I (red, noise-dominated) and Type-II (green, envelope-derived); (b) Noise-band processing with protective guard bands; (c) Boundary refinement and final boundary configuration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

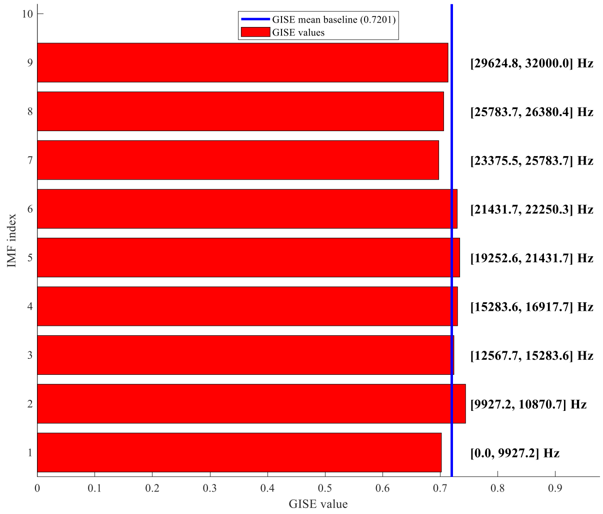


Fig. 24. GISE-based assessment of the decomposed components, with frequency bands ultimately selected for further analysis highlighted.

Within each sub-band, PSO iteratively searches for the center frequency and bandwidth that maximize the ESER value, thereby adaptively identifying the most sensitive frequency region. As illustrated in Fig. 25, the evolutionary behavior of the optimization process is visualized. The ESER value of each sub-band exhibits a consistent upward trend over successive iterations and eventually converges to a stable maximum, indicating effective convergence of the algorithm. Furthermore, in the vicinity of the optimal solution (marked by red asterisk), the particles (denoted by yellow dots) gradually cluster, demonstrating a clear aggregation behavior around the global best.

The final selected frequency bands are [5505.5, 5868] Hz, [23620, 24376.8] Hz, [25783.7, 26110.9] Hz and [30431.5, 31557.3] Hz as shown in Fig. 26(a), where the four optimized sub-bands are highlighted. Based on these results, an optimal band-pass filter bank is constructed. Subsequently, the filtered sub-signals from all bands are individually subjected to envelope demodulation, and their envelope spectra are coherently superimposed to form a composite diagnostic spectrum. The resulting envelope spectrum is presented in Fig. 26(b). Distinct peaks are clearly observed at the ball pass frequency of the outer race (BPFO) and its higher-order harmonics, with negligible spectral interference from unrelated components. The prominent alignment of these peaks with the theoretical fault characteristic frequencies provides strong evidence of an incipient outer-race defect. The raw vibration signal is then filtered through this bank, and the resulting time-domain waveforms exhibit significantly enhanced periodic impulsive characteristics. Notably, even in regions where low-amplitude transient noise previously obscured the fault signatures, the underlying fault-induced impulses are effectively recovered and amplified, which can be seen in Fig. 27.

4.2.3. Parameter justification and ablation studies

To ensure the reliability of the proposed method for subway bogie applications, we performed parameter sensitivity analysis and ablation studies following the same experimental design as in Section 4.1.3. The sensitivity analysis, shown in Fig. 28, confirms that the chosen parameter values ($\alpha = 0.2, \varepsilon = 0.25$) yield an optimal trade-off between transient sensitivity and noise robustness for the subway dataset.

To further validate the robustness of the chosen thresholds on the subway bogie dataset, we performed the same grid-search sensitivity analysis for τ_e and τ_m. As illustrated in Fig. 29, both the number of selected bands and the mean ESER exhibit negligible variation over a wide parameter range, demonstrating that the method's performance is stable and not critically dependent on precise threshold tuning.

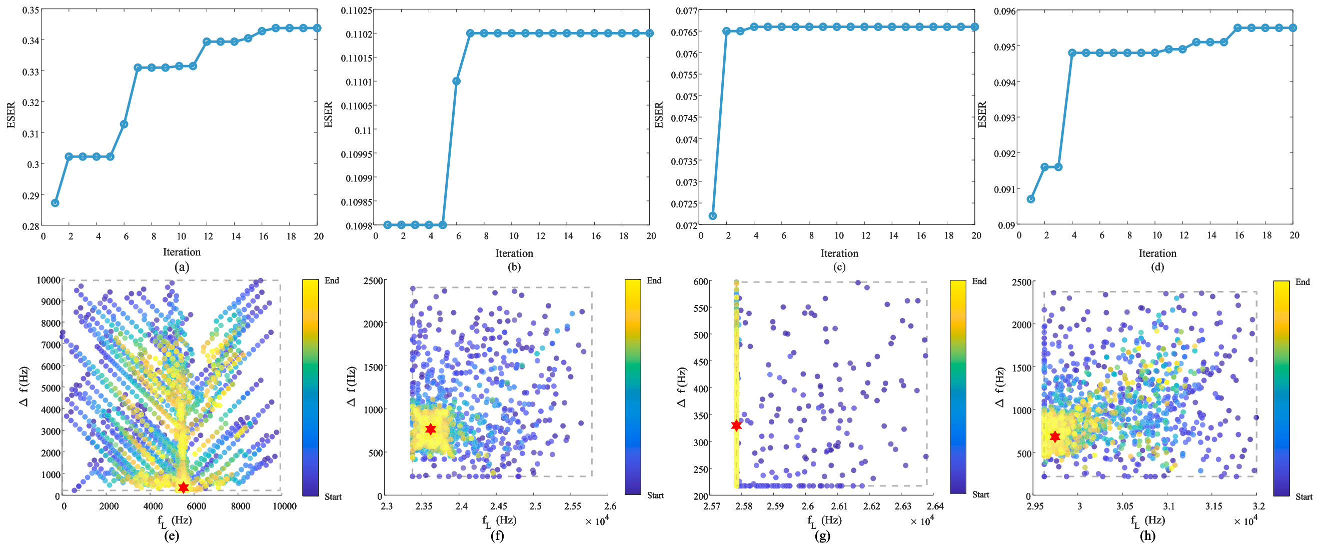


Fig. 25. PSO-driven ESER optimization within each candidate band. Curves exhibit monotonic or near-monotonic improvement of the ESER objective toward convergence: (a) [0, 9927.2] Hz, (b) [23375.5, 25783.7] Hz, (c) [25783.7, 26380.4] Hz, (d) [29624.8, 32000] Hz; while conceptual swarm snapshots indicate progressive clustering around the global optimum, evidencing both stability and efficiency: (e) [0, 9927.2] Hz, (f) [23375.5, 25783.7] Hz, (g) [25783.7, 26380.4] Hz, (h) [29624.8, 32000] Hz.

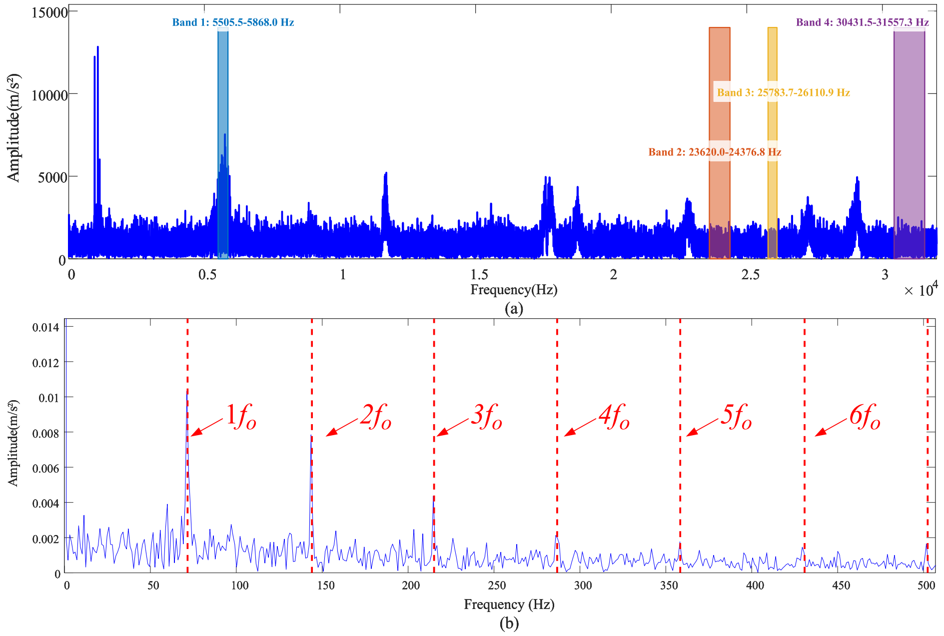


Fig. 26. (a) Final selected demodulation bands; (b) Composite envelope spectrum formed by coherent superposition of band-specific demodulation results.

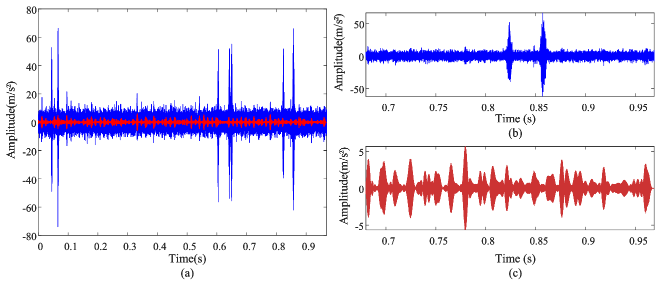


Fig. 27. (a) Comparison of vibration signals before and after band-pass filtering; (b) Zoomed original waveform (0.68 s - 0.93 s); (c) Zoomed filtered waveform (0.68 s - 0.93 s) highlighting recovered impulsive events.

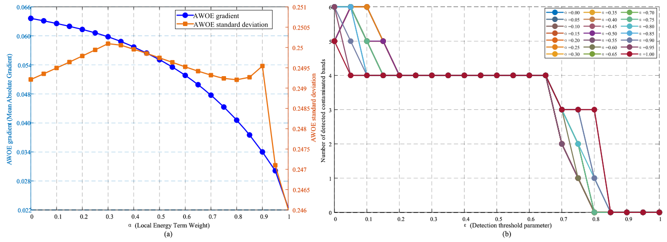


Fig. 28. Parameter sensitivity analysis: (a) Effect of α on AWOE spectral statistics. (b) Effect of ε on the count of detected noise-contaminated bands.

The ablation study results, summarized in Table 6 and visualized in Fig. 30, consistently demonstrate the superiority of the pro-

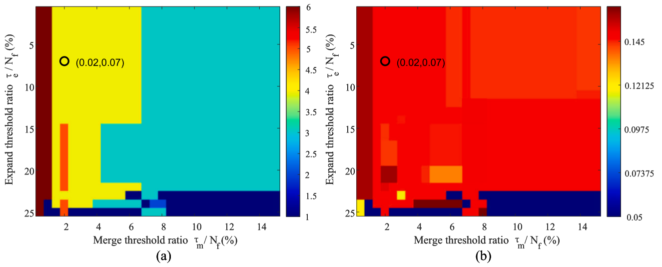


Fig. 29. Parameter sensitivity analysis for expansion threshold τ_e and merge threshold τ_m: (a) the number of selected sub-bands; (b) the mean ESER.

Table 6
Diagnostic performance of algorithmic variants.

Variant	Mean ESER	Std Dev	Max ESER	# Bands
Var. A	0.0963	0.0196	0.1254	4
Var. B	0.1032	0.0269	0.1854	19
Var. C	0.0977	0.0516	0.2332	9
TASS	**0.1563**	**0.1249**	**0.3425**	**4**

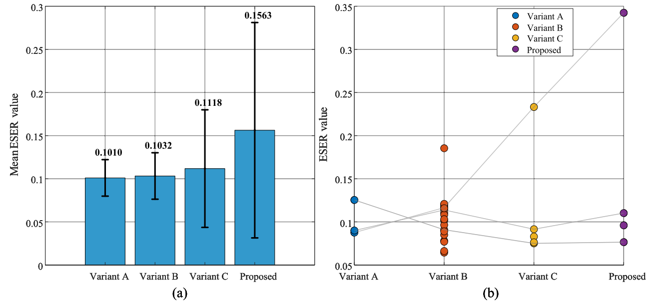


Fig. 30. Ablation study results: (a) Mean ESER values with standard deviations. (b) ESER of individual frequency bands across variants.

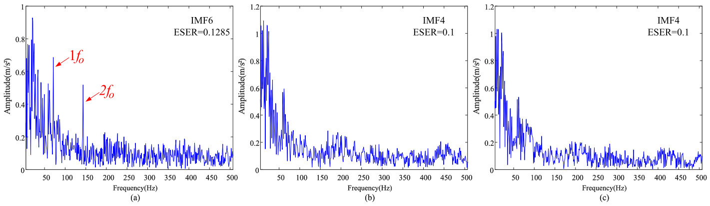


Fig. 31. Envelope demodulation of the best ESER mode for (a) PSO-VMD, (b) EMD, and (c) LMD under strong noise conditions for the subway bogie dataset.

posed full framework. The full TASS framework achieves the highest mean ESER (0.1563) and substantial improvements over all variants, with the synergistic combination of AWOE and GISE proving essential for optimal diagnostic performance. Notably, variants without GISE selection either extracted an excessive number of bands or exhibited high variability, confirming the necessity of both components.

These results collectively validate the parameter selections and the necessity of each sub-module in the TASS framework for subway bogie fault diagnosis applications.

4.2.4. *Comparative evaluation analysis*

For comparative evaluation of feature-extraction performance, three conventional signal-decomposition methods–EMD, LMD and PSO-VMD(K = 2, α = 5000),–are tested. As shown in Table 7, the ESER values of the IMFs obtained by these methods are generally low. Even the IMF with the highest ESER, selected as the most fault-informative component, yields an envelope spectrum in which harmonic structures at the Ball Pass Frequency of the Outer race (BPFO) are either weak or entirely absent, as illustrated in Fig. 31. This observation indicates that fault signatures are severely degraded or lost during decomposition, rendering these methods unreliable under such challenging conditions.

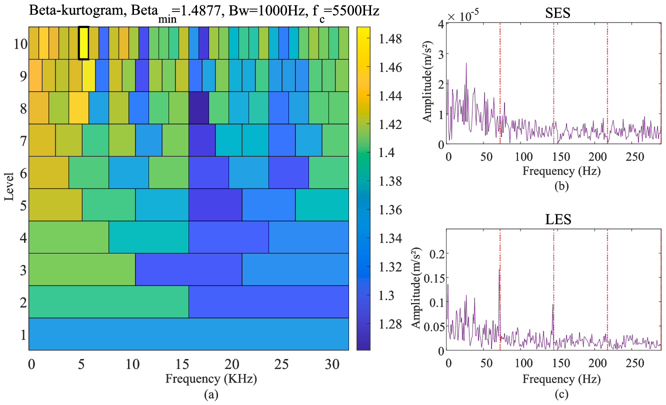


Fig. 32. Beta-Kurtogram-based demodulation for subway bogie: (a) selected band via the Beta-Kurtogram; (b) squared envelope spectrum; (c) logarithmic envelope spectrum.

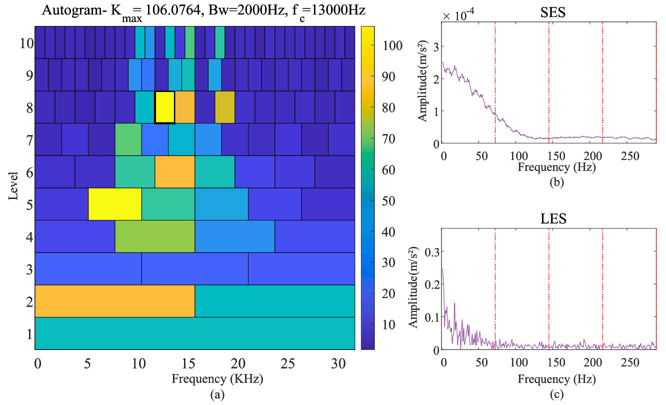


Fig. 33. Autogram-based demodulation for subway bogie: (a) selected band via the Autogram; (b) squared envelope spectrum; (c) logarithmic envelope spectrum.

Furthermore, the proposed method is compared with four representative band-selection techniques–namely, the Beta-Kurtogram, Autogram, Infogram, and Subband Averaging Kurtogram (SAK)–all of which are based on binary-tree filter bank decomposition, as illustrated in Figs. 32(a)–35(a). The optimal frequency bands identified by these methods are [5000, 6000] Hz, [12000, 14000] Hz, [28000, 29000] Hz and [24000, 32000]Hz respectively.

Although these bands exhibit partial overlap with the sensitive band selected by the proposed method, all four techniques employ deep decomposition structures, resulting in excessively narrow sub-bands. Consequently, the demodulated signals within these narrow bands lack sufficient bandwidth to preserve the amplitude modulation (AM) structure induced by localized faults. As a result, critical modulation information is lost, and no significant peaks are observed at the Ball Pass Frequency Outer race (BPFO) or its harmonics in the corresponding envelope spectra–shown in Figs. 32(b)–(c)–35(b)–(c). This outcome highlights the inherent trade-off in binary-tree-based methods between high frequency resolution and effective retention of fault-induced modulation. In contrast, the proposed adaptive band-selection strategy avoids excessive decomposition by intelligently balancing spectral localization and modulation fidelity, thereby preserving fault-relevant features and

Table 7
ESER values for various decomposition algorithms.

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
TASS	**0.3425**	**0.1102**	**0.096**	**0.0765**	–	–	–	–	–
PSO-VMD	0.1385	0.1141	–	–	–	–	–	–	–
EMD	0.04	0.036	0.084	0.1	0.075	0.0338	0.0008	0.0001	0.0001
LMD	0.0427	0.028	0.085	0.102	0.052	–	–	–	–

Table 8
Computational time comparison of different band-selection methods (averaged over 5 runs).

Method	Execution time (s)
TASS	9.532
Beta-kurtogram	6.346
Autogram	2.523
Infogram	1.081
SAK	1.327

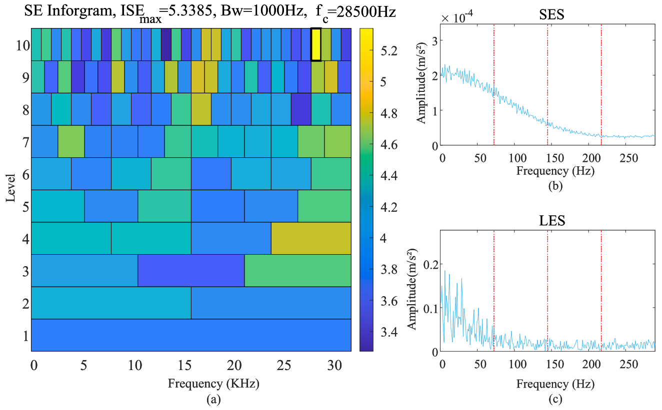


Fig. 34. Infogram-based demodulation for subway bogie: (a) selected band via the infogram; (b) squared envelope spectrum; (c) logarithmic envelope spectrum.

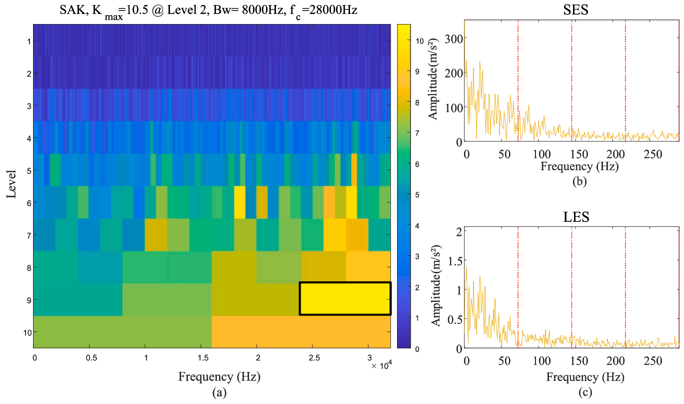


Fig. 35. SAK-based demodulation for subway bogie: (a) selected band via the SAK; (b) squared envelope spectrum; (c) logarithmic envelope spectrum.

enabling more robust fault diagnosis under challenging conditions. The runtime of all methods is shown in the Table 8

5. Conclusion

This study presented Transient-Aware Spectral Segmentation (TASS) combining Amplitude-Weighted Ordinal Entropy (AWOE) for transient localization with PSO-guided band selection via an Enhanced Spectral Energy Ratio (ESER) to deliver physically interpretable, impulse-robust spectral segmentation that suppresses broadband, nonstationary interference while emphasizing defect-related energy in bogie axle bearing signals. Empirically, the framework increases effective SNR, improves detection reliability, and reduces false alarms across onboard and wayside sensing, while preserving transparency suitable for engineering assurance and condition-based maintenance. Future work will prioritize real-time deployment on adaptive/self-supervised parameter updates to manage concept drift across lines and climates, and multimodal fusion with domain adaptation for cross-fleet generalization under limited labels. In addition, we will incorporate uncertainty quantification with calibrated decision thresholds and conduct large-scale field trials with standards-aligned validation to support certification and robust operational adoption.

CRediT authorship contribution statement

Peng Chen: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization; **Qingsheng Wei:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization; **Yazheng Wang:** Visualization, Validation, Resources, Investigation, Data curation; **Ge Xin:** Supervision, Project administration, Funding acquisition, Conceptualization; **Changbo He:** Supervision, Project administration, Funding acquisition, Formal analysis.

Code availability

To promote reproducibility and facilitate future research, the code implementations developed in this study will be made publicly available through the Smart Sensing and Sustainable Diagnosis Prognostics Lab (S3DP-LAB)'s digital archival platform upon publication. The platform provides comprehensive version control and long-term preservation capabilities, ensuring sustained accessibility and reusability of our research outputs. Researchers can access the complete codebase, including implementation details, documentation, and example usage through our institutional repository at https://drpengchen.vip.cpolar.cn/research-work/.

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.

Acknowledgment

This research was supported by the National Natural Science Foundation of China (Grants 52105111, 52375078, U23B20104, 52305085), and the Guangdong Basic and Applied Basic Research Foundation (Grant 2025A1515012256).

References

Allaoui, M., Belhaouari, S. B., Hedjam, R., Bouanane, K., & Kherfi, M. L. (2025). t-SNE-PSO: Optimizing t-SNE using particle swarm optimization. *Expert Systems with Applications, 269*, 126398.

Antoni, J., & Randall, R. B. (2006). The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mechanical Systems and Signal Processing*, *20*(2), 308–331.

Azami, H., & Escudero, J. (2016). Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation. *Computer Methods and Programs in Biomedicine*, *128*, 40–51.

Bandt, C., & Pompe, B. (2002). Permutation entropy: A natural complexity measure for time series. *Physical Review Letters*, *88*(17), 174102.

Chen, P., Gao, J., Wu, Y., He, C., Xin, G., Fan, S., & Qi, J. (2026a). Semi-supervised transfer graph representation learning with few-shot adaptation for gearbox diagnostics under extraneous transient noise. *Structural Health Monitoring*, (p. 14759217251414344).

Chen, P., Ma, J., Gao, J., Xin, G., & He, C. (2026b). Markovian spectral transition modeling with temporal dependencies for railway bogie axle bearing diagnostics in nonstationary transient environments. *Nonlinear Dynamics*, *114*, 244.

Chen, P., Ma, Z., Xu, C., Jin, Y., & Zhou, C. (2024a). Self-supervised transfer learning for remote wear evaluation in machine tool elements with imaging transmission attenuation. *IEEE Internet of Things Journal*, *11*, 23045–23054.

Chen, P., Wu, Y., Fan, S., He, C., Jin, Y., Qi, J., & Zhou, C. (2025a). Adaptive signal regime for identifying transient shifts: A novel approach toward fault diagnosis in wind turbine systems. *Ocean Engineering*, *325*, 120798.

Chen, P., Wu, Y., Ma, J., Zhang, R., Xin, G., & He, C. (2026c). Vibration-weighted maximum correlated kurtosis deconvolution and latent cyclic pattern discovery for fault diagnosis of high-speed rail bogies. *Journal of Sound and Vibration*, (p. 119657).

Chen, P., Wu, Y., Xu, C., Huang, C.-G., Zhang, M., & Yuan, J. (2025b). Interference suppression of nonstationary signals for bearing diagnosis under transient noise measurements. *IEEE Transactions on Reliability*, *74*(3), 4047–4061.

Chen, P., Wu, Y., Xu, C., Jin, Y., & Zhou, C. (2024b). Markov modeling of signal condition transitions for bearing diagnostics under external interference conditions. *IEEE Transactions on Instrumentation and Measurement*, *73* (3518308).

Chen, P., Zhang, R., Fan, S., Guo, J., & Yang, X. (2024c). Step-wise contrastive representation learning for diagnosing unknown defective categories in planetary gearboxes. *Knowledge-Based Systems*, *309*, 112863.

Chen, P., Zhang, R., He, C., Jin, Y., Fan, S., Qi, J., Zhou, C., & Zhang, C. (2025c). Progressive contrastive representation learning for defect diagnosis in aluminum disk substrates with a bio-inspired vision sensor. *Expert Systems with Applications*, (p. 128305).

Cheng, Y., Wang, S., Chen, B., Mei, G., Zhang, W., Peng, H., & Tian, G. (2022). An improved envelope spectrum via candidate fault frequency optimization-gram for bearing fault diagnosis. *Journal of Sound and Vibration*, *523*, 116746.

Deng, F., Qiang, Y., Yang, S., Hao, R., & Liu, Y. (2021). Sparse representation of parametric dictionary based on fault impact matching for wheelset bearing fault diagnosis. *ISA Transactions*, *110*, 368–378.

Ding, A., Qin, Y., Wang, B., Guo, L., Jia, L., & Cheng, X. (2024). Evolvable graph neural network for system-level incremental fault diagnosis of train transmission systems. *Mechanical Systems and Signal Processing*, *210*, 111175.

Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. *IEEE Transactions on Signal Processing*, *62*(3), 531–544.

Gilles, J. (2013). Empirical wavelet transform. *IEEE Transactions on Signal Processing*, *61*(16), 3999–4010.

He, C., Ma, P., Zhi, Y., Hu, S., Wang, J., Qiao, Z., & Lu, S. (2025a). Weak fault feature extraction in non-gaussian noise interference based on adaptive recombination empirical wavelet transform incorporated by sparse coding shrinkage. *Measurement Science and Technology*, *36*(3), 036119.

He, C., Meng, Q., Xu, X., Chen, P., Liang, P., & Cao, J. (2026). An improved lightweight residual network model deployed on the edge device for the unsupervised cross-domain fault diagnosis. *Expert Systems with Applications*, *296*, 129106.

He, D., Wu, J., Jin, Z., Huang, C., Wei, Z., & Yi, C. (2025b). AGFCN: A bearing fault diagnosis method for high-speed train bogie under complex working conditions. *Reliability Engineering & System Safety*, *258*, 110907.

Hou, D., Qi, H., Wang, C., Peng, C., Luo, H., & Han, D. (2023). High-speed train wheel set bearing fault diagnosis and prognostics: Evaluation of signal processing methods under multi-source interference. *Structural Health Monitoring*, *22*(4), 2280–2304.

Jia, L., Zhang, Q., Zheng, X., Yao, P., He, X., & Wei, X. (2019). The empirical optimal envelope and its application to local mean decomposition. *Digital Signal Processing*, *87*, 166–177.

Ma, C., Zhang, W., Shi, M., Zou, X., Xu, Y., & Zhang, K. (2025). Feature identification based on cepstrum-assisted frequency slice function for bearing fault diagnosis. *Measurement*, *246*, 116753.

Miao, Y., Wang, J., Zhang, B., & Li, H. (2022). Practical framework of gini index in the application of machinery fault feature extraction. *Mechanical Systems and Signal Processing*, *165*, 108333.

Pan, T., Wang, T., Chen, J., Xie, J., & Cao, S. (2024). A global and joint knowledge distillation method with gradient-modulated dynamic parameter adaption for EMU bogie bearing fault diagnosis. *Measurement*, *235*, 114927.

Park, J., Yoo, J., Kim, T., Kim, M., Ha, J. M., & Youn, B. D. (2025). Spectral kurtosis attention network (SKAN): Synergizing signal processing and deep learning for fault diagnosis of rolling element bearings. *Expert Systems with Applications*, (p. 128581).

Peng, Y., Shao, H., Xiao, Y., Wang, J., & Liu, B. (2026). Dual-stage interpretable domain generalization fault diagnosis: Integrating prior knowledge and gradient-weighted class activation mapping. *Engineering Applications of Artificial Intelligence*, *166*, 113655.

Shao, H., Ming, Y., Liu, Y., & Liu, B. (2025). Small sample gearbox fault diagnosis based on improved deep forest in noisy environments. *Nondestructive Testing and Evaluation*, *40*(8), 3935–3956.

Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. In *2011 IEEE International conference on acoustics, speech and signal processing (ICASSP)* (pp. 4144–4147). IEEE.

Ugwiri, M. A., Carratú, M., Paciello, V., & Liguori, C. (2021). Benefits of enhanced techniques combining negentropy, spectral correlation and kurtogram for bearing fault diagnosis. *Measurement*, *185*, 110013.

Wan, W., Chen, J., & Xie, J. (2023). Mim-graph: A multi-sensor network approach for fault diagnosis of hsr bogie bearings at the iot edge via mutual information maximization. *ISA Transactions*, *139*, 574–585.

Wang, H., Yan, C., Liu, Y., Li, S., & Meng, J. (2024a). The LFIgram: A targeted method of optimal demodulation band selection for compound faults diagnosis of rolling bearing. *IEEE Sensors Journal*, *24*(5), 6687–6699.

Wang, J., Jia, M., Zhang, Z., Jiang, X., Han, B., & Bao, H. (2024b). Generalized nonlinear hybrid-norm parallel sparse filtering for bearing fault diagnosis under complex interference. *Measurement Science and Technology*, *35*(9), 096109.

Wang, L., Liu, Z., Cao, H., & Zhang, X. (2020). Subband averaging kurtogram with dual-tree complex wavelet packet transform for rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing*, *142*, 106755.

Wang, S., Li, Y., Noman, K., Wang, D., Feng, K., Liu, Z., & Deng, Z. (2024c). Cumulative spectrum distribution entropy for rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing*, *206*, 110905.

Wang, Z., Liang, P., Bai, R., Liu, Y., Zhao, J., Yao, L., Zhang, J., & Chu, F. (2025a). Few-shot fault diagnosis for machinery using multi-scale perception multi-level feature fusion image quadrant entropy. *Advanced Engineering Informatics*, *63*, 102972.

Wang, Z., Liu, Y., Bai, R., Chen, H., Li, J., Chen, X., Yao, L., Zhao, J., & Chu, F. (2025b). Multi-modal multi-scale multi-level fusion quadrant entropy for mechanical fault diagnosis. *Expert Systems with Applications*, *281*, 127715.

Wang, Z., Luo, Q., Chen, H., Zhao, J., Yao, L., Zhang, J., & Chu, F. (2024d). A high-accuracy intelligent fault diagnosis method for aero-engine bearings with limited samples. *Computers in Industry*, *159*, 104099.

Wang, Z., Zhang, M., Chen, H., Li, J., Li, G., Zhao, J., Yao, L., Zhang, J., & Chu, F. (2025c). A generalized fault diagnosis framework for rotating machinery based on phase entropy. *Reliability Engineering & System Safety*, *256*, 110745.

Wei, Q., Chen, P., Gao, J., Zhang, R., He, C., & Qi, J. (2026). EKBD-MK: Entropy-kurtosis bilateral discernment with maximum kurtosis blind deconvolution for fault diagnosis in wind turbine systems. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, *12*(4), 041102.

Wu, D., Wang, Y., Chen, P., Guo, C., Li, R., Xin, G., & He, C. (2026). Cross-correlation heat-maps synthesis with walsh-fourier transformation in articulated robotic systems diagnostics. *Measurement*, (p. 120862).

Xin, G., Chen, Y., Li, L., Chen, C., Liu, Z., & Antoni, J. (2025). Complex symplectic geometry mode decomposition and a novel time–frequency fault feature extraction method. *IEEE Transactions on Instrumentation and Measurement*, *74*, 1–10.

Yang, J., Sun, R., Yao, D., Wang, J., & Wei, M. (2024). Research on rolling bearing fault diagnosis method based on harmonic noise kurtosis-time characteristic blind deconvolution. *IEEE Transactions on Instrumentation and Measurement*, *73*, 1–15.

Zheng, J., Chen, Y., Pan, H., & Tong, J. (2023). Composite multi-scale phase reverse permutation entropy and its application to fault diagnosis of rolling bearing. *Nonlinear Dynamics*, *111*(1), 459–479.