

Article

Physics-Guided Compositional Diagnosis of Unseen Compound Faults in Variable-Speed Induction Motors

Taehong Min *  and Joonghyeok Lee

Institute for Advanced Engineering, 175-28, Goan-ro 51 Beon-gil, Baegam-myeon, Cheoin-gu, Yongin-si 17180, Gyeonggi-do, Republic of Korea; gravity@iae.re.kr

* Correspondence: thmin@iae.re.kr; Tel.: +82-31-330-7856

Abstract

Compound faults in electric motors are difficult to diagnose because simultaneous-fault recordings are scarce and fault multiplicity is unknown at inference. We propose a compound-sample-free, cardinality-free framework for induction motors running continuously varying speed profiles at several load levels. Synchronized key-phase, triaxial vibration and three-phase current signals are converted to order-domain representations by anti-aliased computed order tracking, augmented by a band-pass envelope order spectrum and normalized with a scale-invariant, noise-floor-removed representation. Nine fault primitives are evaluated by modality-specific experts, combined through physics-regularized routing, and decoded by maximum a posteriori inference over 19 feasible machine states. Six leave-one-speed-load-combination-out folds and three seeds evaluate every held-out recording. Exact condition accuracy counts a recording as correct only when the predicted set of fault primitives matches the true set exactly; exact compound recovery applies the same criterion to the unseen compound recordings, requiring both constituent primitives and nothing else. Without a fault-count prior, the pipeline reaches 76.7% and 45.7%, against 62.5% and 11.1% for a conventional order-domain front end and 59.2–62.0% and 0.0–2.5% for three re-implemented baselines. A source-domain modality-selection control reduces compound recovery from 59.3% to 27.8%. Physically aligned representation, sensing specialization and cardinality-aware decoding are therefore critical to compositional motor-fault diagnosis.

Keywords: compound fault diagnosis; induction motor; computed order tracking; multimodal sensor fusion; zero-shot learning; condition monitoring; variable-speed operation



Academic Editors: Israel Zamudio-Ramirez and Jose Alfonso Antonino-Daviu

Received: 19 August 2026

Revised: 2 September 2026

Accepted: 8 September 2026

Published: 10 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Induction motors are the dominant prime mover in industrial drive systems, and unplanned motor failures propagate quickly into process downtime and secondary damage. Modern condition-monitoring installations therefore acquire heterogeneous signals—most commonly triaxial vibration and stator current—because mechanical and electrical degradation modes leave signatures in different physical channels [1–5]. Data-driven diagnosis of such signals has developed rapidly [6–8], yet two practical circumstances remain poorly addressed: operation at continuously varying speed and load, and the simultaneous presence of more than one fault.

Variable operation creates a representation problem. Any component tied to shaft rotation moves in absolute frequency as the speed changes, so a fixed-time spectral representation encodes the operating condition at least as strongly as the machine condition.

Computed order tracking (COT) addresses this by resampling the measured signal at uniform shaft-angle increments, which maps speed-synchronous components onto fixed rotational orders [9,10]. COT is standard practice, but its correctness depends on a detail that is frequently omitted in machine-learning pipelines: resampling to a fixed number of samples per revolution is a decimation whose rate follows the instantaneous speed, and it therefore requires a speed-tracking anti-alias filter. Section 3.4 shows that, for the benchmark used here, omitting this filter causes more than 80% of the vibration energy at 1000 rpm to fold back into the analyzed order band at speed-dependent positions.

The second circumstance is the scarcity of compound-fault data. Conventional multi-class formulations treat every combination of faults as an independent category, which requires an observation of each combination of interest. The number of possible combinations grows combinatorially, and deliberately inducing several simultaneous faults is expensive and sometimes unsafe. Zero-shot and generalized zero-shot compound-fault diagnosis therefore transfer information from single faults through semantic, attribute, physical-information, or multi-output representations [11–25]. Recent generalized formulations explicitly distinguish single from compound conditions rather than assuming that a test sample is compound [21,24,25]. The remaining deployment question addressed here is stricter: can the number and identity of coexisting faults be inferred jointly from variable-speed motor data without semantic side information, generated compound samples, target-domain data, or a fault-count input to the decoder?

A third issue concerns multimodal fusion. Vibration and stator current are not uniformly informative across fault mechanisms. Rolling-element defects excite structural resonances and are naturally observed in vibration [26], whereas supply-voltage imbalance and inter-turn winding faults act directly on the electrical circuit [2,3]. Universal early or latent fusion therefore risks interference rather than guaranteed complementarity [27,28]. The question that matters for compound diagnosis is not merely whether fusion helps on average, but which sensing mechanism should provide the evidence for each constituent of a fault combination that was never observed in training.

This paper takes these three issues together. Motor conditions are represented by nine atomic fault primitives; evidence for each primitive is produced by a modality-specific expert; the modality assignment follows fault physics but is refined by source-domain data; and the final diagnosis is obtained by maximum a posteriori (MAP) inference over the set of physically feasible machine states, which contains the healthy state, the nine single-fault states and the nine admissible compound states. Because the healthy and single-fault states are candidates, the decoder decides the number of coexisting faults itself. No compound recording, real or synthetic, is used for training.

The principal empirical finding is that isolated-fault performance is a poor proxy for compound-fault generalization, and that the gap becomes much larger when the fault-count prior is removed. Under the conventional two-fault protocol, the proposed model recovers 67.9% of compound pairs exactly. When the fault count is not supplied, the same representation and routing decoded over feasible machine states recover 45.7% of unseen compound recordings, whereas the same model with conventional front end B0 recovers 11.1% and the three comparison baselines recover 0.0–2.5%.

Each of the three obstacles is met by one specific component of the proposed pipeline, and the correspondence is one-to-one. Order-domain aliasing under varying speed is removed by the anti-aliased computed order tracking front end of Section 3.4, which low-pass filters the record once per occupied speed bin before angular resampling so that content above the angular Nyquist frequency can no longer fold into the retained order band. The scarcity of compound recordings is addressed by the compositional formulation of Section 3.3, in which training uses healthy and single-fault recordings only

and a compound condition is expressed as a set of atomic primitives rather than as a class of its own. The unknown number of coexisting faults is handled by the cardinality-free MAP decoder of Section 3.8, which selects among 19 physically feasible machine states and therefore returns the fault count as an output. The uneven informativeness of the sensing modalities is handled by the physics-regularized routing of Section 3.7, which assigns the evidence for each primitive to the modality in which its physical mechanism is expressed instead of learning a single global fusion rule.

The contributions are:

1. Compound-fault diagnosis is formulated as compound-sample-free and cardinality-free inference over nine atomic fault primitives, and is evaluated on every recording of the held-out operating condition—healthy, single and compound alike.
2. An anti-aliased key-phase-guided COT front end is combined with a band-pass envelope order spectrum and a scale-invariant, noise-floor-removed spectral representation. Each component is ablated.
3. A physics-regularized primitive-specific routing of vibration, phase-current and Clarke-current evidence is proposed, and is compared against uniform weighting, freely learned weighting, and a source-domain-only modality selection control that uses no target and no compound data.
4. Evaluation uses leakage-resistant leave-one-speed-load-combination-out folds, three seeds, source-only probability calibration, and a decoder whose ontology is systematically weakened from 19 feasible states to 46 states to none.
5. Negative and limiting results are reported explicitly, including the compositions that remain unsolved and the operating-envelope extrapolation limits.

2. Related Work

Condition monitoring of induction motors: Motor current signature analysis and vibration monitoring are the two established routes to motor fault detection, and their relative sensitivity to broken rotor bars, eccentricity, bearing damage and stator winding faults has been reviewed repeatedly [1–5]. Eccentricity in particular admits several competing indicators whose observability depends on the sensing channel [29]. Data-driven diagnosis has since become dominant [6–8], but most studies still evaluate a single sensing modality at a small number of steady operating points.

Order-domain analysis under variable speed: Computed order tracking resamples a measured signal at uniform shaft-angle increments so that speed-synchronous components map onto fixed rotational orders [9,10]. Angular resampling driven by a tachometer or key-phase reference is standard [10,30], and tachless variants estimate the instantaneous phase from the signal itself [31,32]. The angular resampling step is a speed-dependent decimation, and the need to band-limit before it has been discussed in the order-tracking literature [33]; it is nevertheless frequently omitted in learning-based pipelines. The present work treats the anti-alias stage as part of the method and quantifies its effect on cross-condition generalization.

Bearing diagnostics: Envelope demodulation of a high-frequency resonance band is the standard route to rolling-element defect information [26], with the band commonly selected by spectral kurtosis [34,35] and with cyclostationary formulations clarifying what the envelope spectrum does and does not represent [36]. Benchmarks built on this practice are widely used [37].

Compound-fault diagnosis without compound training data: A substantial body of literature addresses compound faults that are absent from training. Attribute- and semantic-embedding formulations transfer knowledge from single faults through a human-defined description space [11–14]; label-description spaces [15], fault-semantics learning [16,17],

semantic construction [18], broad zero-shot screening [19], semantic autoencoders [20], physical-information semantics [21], cross-domain transfer [22], and multi-output decomposition [23] have all been used for unseen compound conditions. More recent generalized zero-shot formulations explicitly model the coexistence of seen single faults and unseen compound faults by two-stage semantic decision rules [24] or semantic graph embedding with multi-stage fusion [25]. The formulation adopted here differs in using no semantic side information, no generated compound features, and no target-domain data, while making the fault count an output of a physically feasible-state decoder under variable-speed and variable-load operation.

Discarding semantic side information has consequences that go beyond convenience. An attribute space has to be written down by an engineer, and the choice of attributes, their granularity and their assumed independence are all judgments that are difficult to audit and that differ between authors for the same machine. Recent surveys of zero-shot diagnosis organize the field precisely along the axis of what auxiliary description is assumed [38], and recent compound-fault formulations continue to depend on a hand-built attribute or semantic space, whether it is refined by a pre-classifier and physical-information semantics [21], weighted adaptively inside a semantic autoencoder [20], generated by a diffusion model [39] or transferred across devices as fault attributes [40]. Any bias introduced at that stage propagates into every downstream prediction, and it does so silently, because the attribute space is usually reported as part of the problem definition rather than as a modeling choice. A description that is accurate for one machine also transfers poorly: attributes calibrated on one rating, mounting or sensor layout have to be rewritten for the next installation, which is precisely the situation in which cross-domain generalization is claimed. The formulation used here replaces the attribute space with two objects that are already available in an industrial setting and can be checked independently of the model: the physical mechanism that determines which sensing modality carries the evidence for each primitive, and the maintenance history that determines which combinations of primitives are physically admissible. Neither has to be tuned on data, and both can be inspected and corrected by a maintenance engineer without retraining.

Multi-label formulation and expert-based fusion: Treating coexisting faults as a multi-label rather than a multi-class problem follows established practice in multi-label learning, including its evaluation measures [41,42]. Combining several specialized predictors through a gating function dates back to mixtures of local experts [43] and has recently been applied to multimodal machinery diagnosis, both to weight modalities and to tolerate missing ones [28]. Multimodal domain-generalization methods fuse vibration, current and acoustic channels and disentangle modality and domain factors for unseen working conditions [27], and adversarial and transfer formulations remain common baselines for cross-condition generalization [44,45]. Those methods seek a globally transferable fused representation; the question addressed here is narrower and, for compositional diagnosis, more consequential: which sensing mechanism should supply the evidence for each individual fault primitive. Recent multimodal work on induction motors makes the same point from the opposite direction: correlation-driven fusion of vibration, current, voltage, torque and speed raises isolated-fault accuracy substantially, but the fusion weights are derived from statistical dependence between channels rather than from the fault mechanism, and the evaluation is confined to single faults [46].

Probability calibration: Because the decoder compares likelihoods across candidate machine states, the per-primitive probabilities must be comparable. Temperature scaling is used for this purpose [47]; here, it is fitted on source recordings only.

3. Materials and Methods

3.1. Dataset and Data Audit

The MCC5-THU benchmark [48] provides synchronized key-phase, torque, triaxial vibration and three-phase current measurements from a 2.2 kW induction motor sampled at 12.8 kHz. The speed-circulation subset used here contains 90 s recordings acquired under six speed–load combinations, $\{20, 40\} \text{ N}\cdot\text{m} \times \{1000, 2000, 3000\} \text{ rpm}$, for 24 machine conditions: healthy, 14 single-fault labels covering nine fault mechanisms with two severity levels where available, and nine compound conditions. The curated recording layout and the leave-one-speed–load–combination–out split are shown in Figure 1; compound-condition cells are evaluation-only, and one complete operating-condition column is withheld in each fold.

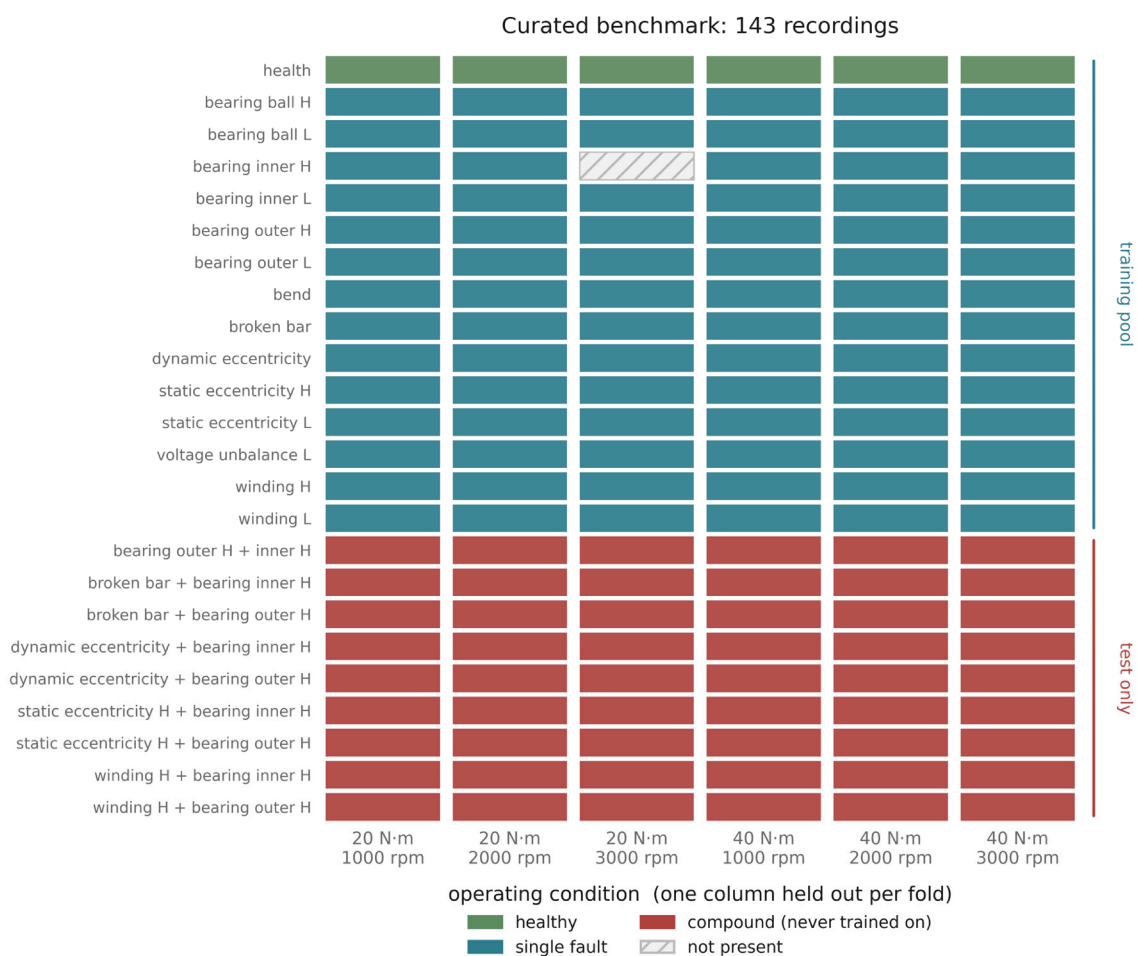


Figure 1. Curated benchmark. Twenty-four machine conditions across six speed–load combinations give 143 recordings after one anomalous duplicate is excluded; the missing cell is marked. Compound conditions (red) are never used for training, and one complete operating-condition column is withheld in each fold. H and L denote the severe and slight fault severity grades, respectively.

The nominal 24×6 layout would contain 144 recordings. The audit found one missing cell (bearing inner H at 20 N·m/3000 rpm) and one duplicated cell (two distinct bearing outer H recordings at 20 N·m/3000 rpm, one of which has a markedly different speed trajectory). The recording with the earlier timestamp is retained and the anomalous duplicate is excluded, giving a curated benchmark of 143 recordings. No relabeling was performed. One nomenclature caveat is stated explicitly: the condition distributed as dynamic eccentricity is described in the dataset documentation as rotor unbalance induced

by an added imbalance mass. The distributed label is retained throughout, and Section 4.4 compares all three sensing modalities for this primitive rather than assuming its assignment.

The consequences of this label mismatch can be bounded rather than merely noted. If the excitation is a rotating unbalance force, its evidence is a first-order radial vibration component; if it were a true dynamic eccentricity, the dominant signature would be a rotor-slot and sideband pattern in the current. The two hypotheses therefore imply different routing assignments, and the difference is measurable: Section 4.4 reassigns this primitive to vibration and to phase current and compares both against the distributed Clarke-current assignment, and both reassignments lower exact recovery from 0.593 to 0.537. Because the same primitive appears in two of the nine compound compositions, the effect of a wrong assignment is confined to those two compositions and to the routing weight of one of nine primitives; the remaining eight primitives and seven compositions are unaffected. What the mismatch does limit is interpretation rather than performance. Statements made in this paper about the primitive labeled dynamic eccentricity should be read as statements about the recorded condition, which is an added rotor imbalance mass, and not as evidence about air-gap eccentricity in general. For the same reason, the recovery rates of the two compositions containing this primitive should not be transferred to installations in which eccentricity arises from bearing wear or misalignment.

The present study reanalyzes the publicly available MCC5-THU recordings; no new experimental measurements were acquired. The test rig, sensing chain, and fault preparation described below refer to the original benchmark acquisition documented in the data descriptor [48]. The rig comprises a 2.2 kW three-phase asynchronous motor (QABP-90L2, two pole pairs, and a 0.5 mm single-side air gap), a torque sensor, a two-stage parallel gearbox with the planetary stage removed, and a magnetic powder brake acting as the load. The motor is equipped with 6205-2Z/C3 bearings (SKF, Gothenburg, Sweden) at the drive and fan ends. Eight channels were acquired synchronously at 12.8 kHz: the motor key-phase pulse; input-shaft torque (S2001 sensor, 100 mV/(N·m), 0.5% full scale); triaxial acceleration at the motor drive end in the horizontal radial, axial and vertical radial directions (TES001V accelerometers, 100 mV/g); and the three phase currents (i30s current clamps; Fluke Corporation, Everett, WA, USA; 100 mV/A). Each run lasts 90 s. Faults were introduced physically: raceway notches cut by laser to 0.2 mm width for the slight grade and 0.6 mm for the severe grade at 0.5 mm depth, a shunt resistor across two winding taps producing circulating currents of 5% and 10% of rated phase current, radial offsets of 0.125 mm and 0.250 mm for static eccentricity, a 0.2 mm permanent shaft bend, three consecutive rotor bars removed with compensating counterweights, and an imbalance mass attached to the rotor shaft. The bearing geometry gives dimensionless defect frequencies of 3.585 orders for the outer race, 5.415 orders for the inner race and 2.357 orders for the rolling element; these values are used in Section 4.6. The recordings were acquired under laboratory conditions, with temperature controlled within ± 2 °C, as reported in the original data descriptor [48].

The speed-circulation subset is not a set of steady operating points, and the nominal speed label denotes the upper plateau of a prescribed speed profile rather than a constant. Each 90 s run follows a continuous run-up to the nominal speed, a plateau, a step down to an intermediate level, a second plateau, a step back up and a run-down, as the measured shaft speed in Figure 2a shows. Measured over all 13,720 analysis segments, the median segment speed within the nominal 1000 rpm recordings spans 414–991 rpm, within the 2000 rpm recordings 624–1990 rpm and within the 3000 rpm recordings 766–3000 rpm; that is, 33% to 69% of the upper plateau. The mean absolute speed rate is 25 rpm/s at the 1000 rpm setting and 64 rpm/s at the 3000 rpm setting, reaching 435 rpm/s during the transitions. Only 63% to 72% of the segments are classified as steady by the criterion

of Section 3.9, so between 28% and 37% of the evaluation material is acquired while the machine is accelerating or decelerating; Section 4.6 reports the diagnosis performance separately for these regimes. The load is held at its nominal value within a run and is varied only between runs, so the benchmark used here is non-stationary in speed and multi-level, not continuously varying, in load. The complementary torque-circulation subset of the same benchmark, in which torque varies continuously at fixed speed, is left for future work.

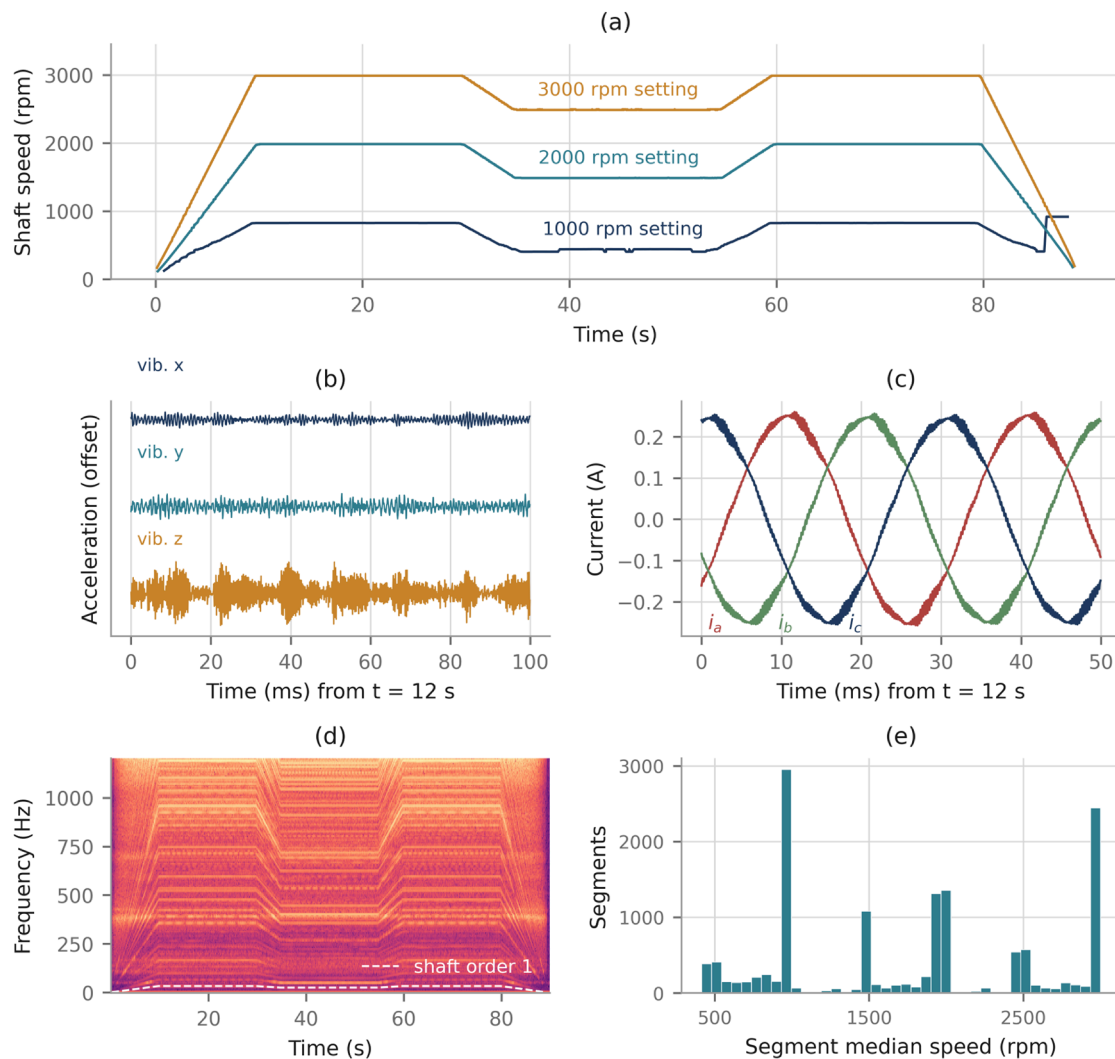


Figure 2. Raw measurements of the speed-circulation subset, taken from the compound outer-race plus inner-race bearing condition at 20 N·m. (a) Shaft speed recovered from the key-phase edges over three complete 90 s acquisitions, one for each nominal speed setting; every run is a continuous run-up, plateau, step and run-down cycle rather than a steady operating point. (b) Raw triaxial vibration and (c) raw three-phase current over a short window of the 2000 rpm acquisition, taken at $t = 12$ s while the machine is accelerating. (d) Short-time spectrum of the horizontal radial vibration of the same acquisition, with the first shaft order overlaid; the speed-swept harmonic families are what order tracking is required to stationarize. (e) Distribution of the median shaft speed of all 13,720 analysis segments in the curated benchmark.

Table 1 summarizes the composition of the curated subset: the number of recordings and of analysis segments for the healthy condition, for each single fault at each severity grade and for each compound composition, together with their distribution over the six speed-load combinations.

Table 1. Composition of the curated MCC5-THU speed-circulation subset, giving for every machine condition its severity grade, the number of 90 s recordings and the number of 16-revolution analysis segments. Severity grades follow the data descriptor. The severe inner-race condition has five recordings because one operating-condition cell is missing from the benchmark, and the compound outer + inner condition yields eight fewer segments because one of its recordings admits only 88 valid segments. Except for the missing severe inner-race cell, each condition is represented by one recording at each of the six speed-load combinations.

Category	Machine Condition	Severity	Recordings	Segments
Healthy	Healthy	-	6	576
Single	Bearing ball defect	slight/severe	6/6	576/576
Single	Bearing inner-race defect	slight/severe	6/5	576/480
Single	Bearing outer-race defect	slight/severe	6/6	576/576
Single	Static eccentricity	slight/severe	6/6	576/576
Single	Stator winding fault	slight/severe	6/6	576/576
Single	Voltage unbalance	slight	6	576
Single	Bent shaft	-	6	576
Single	Broken rotor bar	-	6	576
Single	Dynamic eccentricity (rotor unbalance)	-	6	576
Compound	Bearing outer + inner race	severe	6	568
Compound	Broken bar + bearing inner	severe	6	576
Compound	Broken bar + bearing outer	severe	6	576
Compound	Dynamic ecc. + bearing inner	severe	6	576
Compound	Dynamic ecc. + bearing outer	severe	6	576
Compound	Static ecc. + bearing inner	severe	6	576
Compound	Static ecc. + bearing outer	severe	6	576
Compound	Winding fault + bearing inner	severe	6	576
Compound	Winding fault + bearing outer	severe	6	576

3.2. Research Pipeline

Figure 3 summarizes the pipeline. Key-phase edges drive a speed-binned, anti-aliased angular resampling; the vibration branch additionally carries a band-pass envelope order spectrum. Primitive-wise evidence from the three experts is combined by physics-regularized routing, and the diagnosis is decoded by MAP inference over the 19 physically feasible machine states, so that the number of coexisting faults is an output. Training uses healthy and single-fault recordings only.

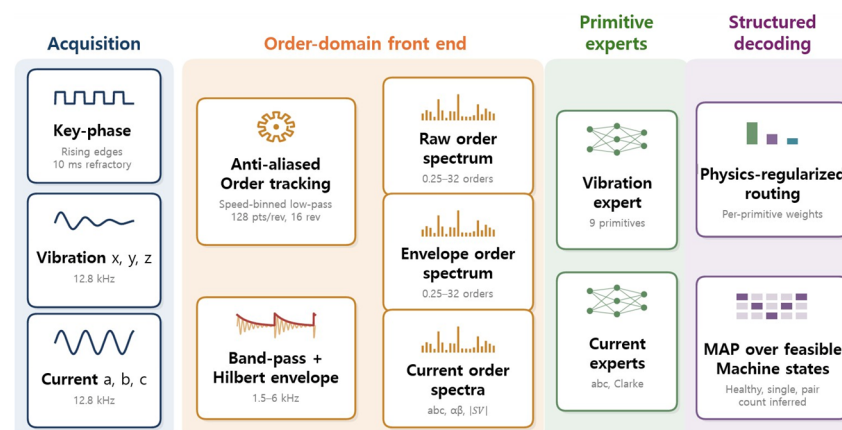


Figure 3. Processing pipeline. Key-phase edges drive a speed-binned, anti-aliased angular resampling; the vibration branch additionally carries a band-pass envelope order spectrum. Primitive-wise evidence from the three experts is combined by physics-regularized routing, and the diagnosis is decoded by maximum a posteriori inference over the 19 physically feasible machine states, so that the number of coexisting faults is an output. Training uses healthy and single-fault recordings only.

3.3. Problem Formulation

Let a recording be described by a multi-hot vector $y \in \{0,1\}^9$ over the primitive vocabulary $P = \{\text{bearing ball, bearing inner race, bearing outer race, bent shaft, broken rotor bar, dynamic eccentricity, static eccentricity, voltage unbalance, winding fault}\}$. Healthy recordings have $\|y\|_0 = 0$, single faults $\|y\|_0 = 1$ and compound faults $\|y\|_0 = 2$.

Training uses $D_{\text{train}} = D_{\text{healthy}} \cup D_{\text{single}}$ only, so $D_{\text{compound}} \cap D_{\text{train}} = \emptyset$. Neither real nor synthetic compound samples are used. At inference, the fault cardinality, $\|y\|_0$, is not given. The decoder must return a member of the feasible state set defined in Section 3.8, which includes the healthy and single-fault states.

The formulation places no upper bound on the number of coexisting primitives. The decoder of Section 3.8 searches a set of admissible machine states, and states of three or more simultaneous primitives can be added to that set without any change to the experts, to the routing, or to the training procedure, because training never sees a compound recording in the first place. What limits the present study is the benchmark: every compound recording in the MCC5-THU speed-circulation subset contains exactly two primitives, so no measurement is available against which a triple or higher-order combination could be scored. Multi-fault inference beyond two primitives is therefore available in principle but is not validated here, and the experimental conclusions of this paper are claimed only for two-fault combinations.

3.4. Anti-Aliased Key-Phase-Guided Order Tracking

Shaft revolutions are located from the rising edges of the key-phase signal using a 50% amplitude threshold and a 10 ms refractory interval. Each revolution is resampled to 128 angular samples, and 16 consecutive revolutions with a hop of 8 revolutions form one segment of 2048 angular samples. Segments are accepted when the median speed lies between 200 and 3600 rpm and the intra-segment speed ratio does not exceed 1.65.

Resampling a 12.8 kHz record to 128 samples per revolution is a decimation by a factor of $12,800/(128 f_{\text{shaft}})$, i.e., by 6, 3 and 2 at 1000, 2000 and 3000 rpm, respectively. Interpolating directly from the time samples therefore folds all content above the angular Nyquist frequency, $64 f_{\text{shaft}}$, back into the analyzed band. Figure 4 quantifies this on the benchmark.

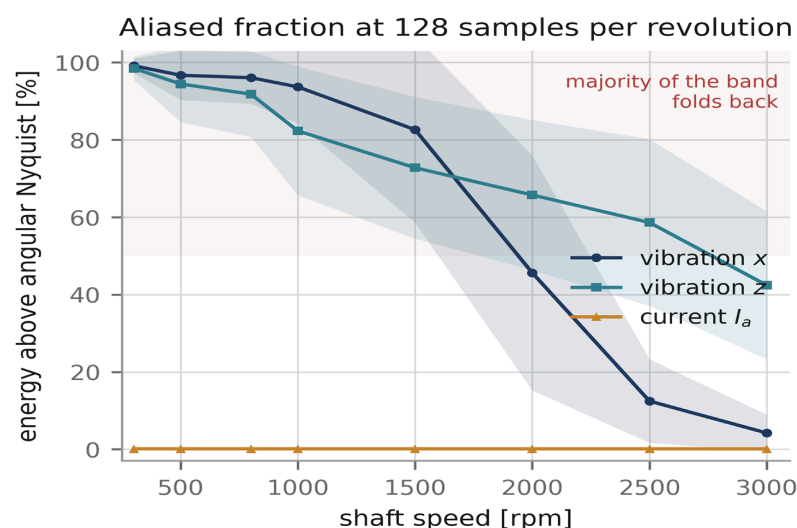


Figure 4. Fraction of signal energy that lies above the angular Nyquist frequency and, therefore, folds back into the analyzed order band when a record is resampled to 128 samples per revolution without anti-aliasing, plotted against shaft speed. Each curve is the mean over all 143 audited recordings of the benchmark, and the shaded band is plus or minus one standard deviation across those recordings. The two vibration channels lose the majority of their band to fold-back below about 1500 rpm, whereas the phase current is unaffected at every speed because its energy is concentrated near the supply frequency.

Averaged over all 143 audited recordings of the benchmark, the fraction of vibration energy above the angular Nyquist frequency is 96.7% (x) and 94.5% (z) at 500 rpm, 93.7% and 82.3% at 1000 rpm, and still 4.2% and 42.4% at 3000 rpm, whereas for the phase current, it never exceeds 0.14% at any speed. The two vibration axes also differ from each other: the horizontal radial channel clears the band by 2500 rpm, while the vertical channel still folds back 58.6% of its energy at that speed, because the structural resonance it excites lies higher in frequency. Aliasing is thus a vibration-specific problem whose severity depends on both the speed and the channel, and the fold-back positions move with speed, reintroducing exactly the operating-condition dependence that order normalization is meant to remove. The proposed front end therefore groups segments into speed bins and low-pass filters the record once per occupied bin with an eighth-order Butterworth filter whose cutoff is $0.45 \times 128 \times f_{\text{shaft,min}}$ (57.6 orders) before angular interpolation.

The two hyper-parameters of this filter, the Butterworth order and the cutoff coefficient, are set from the response requirement rather than tuned on diagnostic accuracy, and Table 2 reports both the design quantities and the resulting performance. The requirement has two sides. The filter must leave the retained band up to 32 orders untouched, since that is where the diagnostic content lies, and it must attenuate the angular Nyquist frequency at 64 orders enough that what folds back is negligible against the noise floor. A cutoff coefficient of 0.45 places the -3 dB point at 57.6 orders, which is 1.8 times the top of the retained band and 0.9 times the angular Nyquist frequency; with an eighth-order filter applied in the zero-phase forward-backward form this gives 0.00 dB at 32 orders and -18.5 dB at 64 orders, averaged over 500–3000 rpm. Measured on real recordings, the vibration energy above the angular Nyquist frequency falls from 95.8% of the total without filtering to 0.15% with this setting. Varying the coefficient between 0.35 and 0.50 and the order between 4 and 12 moves the residual fold-back over three orders of magnitude, from below 0.001% to 2.2%, while exact pair recovery changes by at most 1.9 percentage points and no setting is significantly better than another. The front end is therefore insensitive to these choices over the useful range; what matters is that the filter is present at all, since removing it entirely admits 95.8% fold-back and costs 3.7 percentage points (Table 3). The proposed setting sits in the middle of this plateau and was chosen because it is the widest cutoff that still keeps the measured fold-back below the 0.2% level while leaving the retained band exactly flat.

Table 2. Sensitivity of the anti-aliasing front end to the Butterworth order and to the cutoff coefficient. The two gain columns are the analytic magnitude response of the zero-phase forward-backward filter at the top of the retained order band and at the angular Nyquist frequency, averaged over 500–3000 rpm. Residual fold-back is the measured share of vibration energy above the angular Nyquist frequency after filtering, averaged over six recordings. Exact pair recovery is the mean over six folds and three seeds under the protocol of Table 3; every setting other than the anti-alias filter is held at the proposed configuration.

Butterworth Order	Cutoff Coefficient	Gain at 32 Orders (dB)	Gain at Angular Nyquist (dB)	Residual Fold-Back (%)	Exact Pair
none (no anti-aliasing)	-	0.00	0.00	95.76	0.6420
4	0.45	-0.06	-11.42	2.164	0.6728
8	0.35	-0.03	-56.51	0.000	0.6728
8	0.40	0.00	-36.16	0.005	0.6852
8	0.45 (proposed)	0.00	-18.49	0.152	0.6790
8	0.50	0.00	-6.02	1.938	0.6914
12	0.45	0.00	-26.47	0.019	0.6914

Table 3. Effect of the order-domain front end on compound-fault recovery under the conventional protocol, in which the decoder is told that exactly two faults are present. All rows share the model, the routing, the six leave-one-speed-load-combination-out folds and the three seeds, so each row isolates one property of the representation. Exact pair is the fraction of the 54 held-out compound recordings whose two constituents are both identified; ontology-free top-2 applies the same criterion to the two highest-scoring primitives without the admissible-pair constraint, and the last column is the difference between the two, that is the contribution of the ontology.

Front End	Exact Pair (9-Pair Ontology)	Constituent Micro-F1	Ontology-Free Top-2	Ontology Contribution
Conventional front end B0 (1.0-order, no anti-aliasing, absolute log spectrum)	0.6481 ± 0.1334	0.8148	0.5494	+9.9 pp
0.25-order resolution only	0.7346 ± 0.1328	0.8673	0.6173	+11.7 pp
Proposed (anti-aliased COT + envelope + scale-invariant)	0.6790 ± 0.1420	0.8395	0.6667	+1.2 pp
Proposed – anti-aliasing	0.6420 ± 0.1454	0.8210	0.5988	+4.3 pp
Proposed + zero-sequence retained	0.7099 ± 0.0944	0.8549	0.6975	+1.2 pp
Proposed with absolute log spectrum	0.7346 ± 0.0775	0.8673	0.6852	+4.9 pp

Throughout this paper, this is contrasted with a conventional front end, denoted B0, that reproduces widespread practice on this type of data: direct angular interpolation without anti-aliasing, order bins pooled to a resolution of one shaft order, absolute log-magnitude spectra, retention of the zero-sequence current channel, and no envelope order spectrum. B0 shares the segmentation, the model, the folds and the seeds with the proposed configuration, so every comparison between them isolates the front end.

3.5. Envelope Order Spectrum

Rolling-element defect information is carried by impact-excited structural resonances well above the retained order band. Following standard bearing-diagnostic practice [26,34,35], a band-pass filter (1.5–6 kHz), Hilbert transform and mean removal produce the vibration envelope, which is then passed through the same anti-aliased angular resampling. The envelope order spectrum is concatenated with the raw order spectrum in the vibration branch. Section 4.2 ablates this component.

3.6. Scale-Invariant Spectral Representation

For each channel and segment, the magnitude spectrum $|X|$ is evaluated on a 0.25-order grid from 0.25 to 32 orders (128 bins) and transformed according to Equation (1):

$$S = \log \left(1 + \frac{|X|}{\text{RMS}_{\text{bins}}(|X|)} \right), \quad \tilde{S} = S - \text{median}_{\text{bins}}(S) \quad (1)$$

The first step removes the channel amplitude scale, which varies strongly with speed and load; the second removes a uniform lift of the spectral floor. The Clarke branch retains α , β and the space-vector modulus; the zero-sequence component is discarded, because in a three-wire machine it is physically near zero—measured RMS is 0.55% of the phase-current RMS—while additive sensor noise survives the $(a + b + c)/3$ combination at $\sigma/\sqrt{3}$. Figure 5 shows the consequence: under 20 dB additive white Gaussian noise the standardized features of the zero-sequence channel move by 116σ , against 3.4– 3.7σ for α , β and the space-vector modulus, and 0.02– 0.09σ for the vibration channels. Removing the zero-sequence channel and adopting the representation above reduces the branch-level shift from 4.68σ (phase current) and 26.5σ (Clarke) to 0.81σ and 0.82σ .

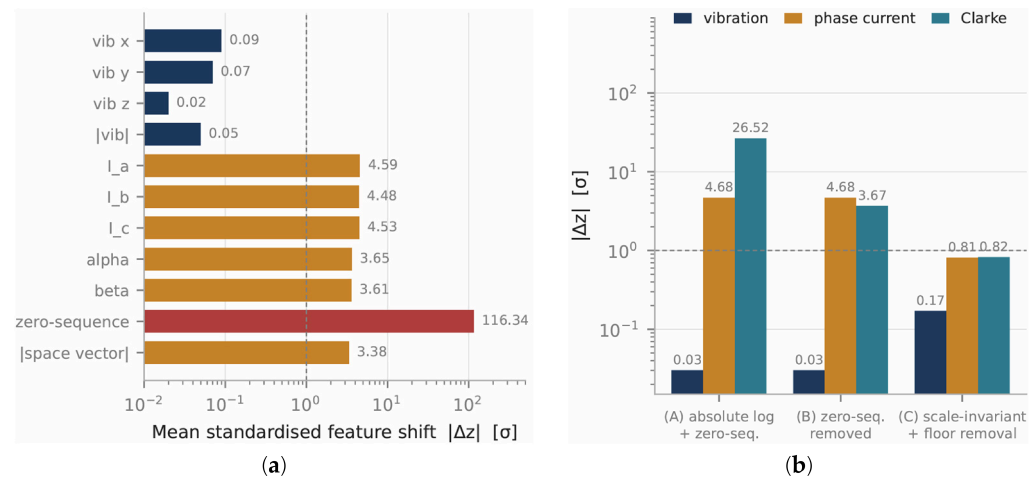


Figure 5. Sensitivity of the spectral representation to additive sensor noise at 20 dB. **(a)** Mean standardized feature shift per channel; the zero-sequence channel moves by 116σ while the vibration channels move by less than 0.1σ . **(b)** Branch-level shift for the representation of front end B0, for the same representation after removing the zero-sequence channel, and for the proposed scale-invariant, floor-removed representation.

Two groups of settings have to be distinguished before this representation is moved to another machine or another dataset. The first group is dimensionless and transfers unchanged. The order grid from 0.25 to 32 orders, the 16-revolution segment with an 8-revolution hop, the 128 samples per revolution, the anti-alias cutoff coefficient and the two operations of the representation itself, division by the root-mean-square over the order bins and subtraction of the median over the order bins, are all defined on the shaft-angle axis or on the spectrum shape, so none of them depends on the motor rating, on the sampling rate or on where the accelerometer is mounted. The quality-control thresholds on segment speed are of the same kind once expressed relative to the machine's speed range. The second group is dimensional and has to be re-derived. The envelope band of 1.5–6 kHz is chosen to contain the structural resonance excited by rolling-element impacts, which depends on the mass and stiffness of the specific bearing housing and on the accelerometer mounting; on a different machine it should be re-selected—for example, by spectral kurtosis or by a hammer test—and the analysis bandwidth must in any case satisfy the sampling rate of the new acquisition. The upper order limit of 32 is adequate here because the highest defect frequency of interest is 5.4 orders and its harmonics decay before that limit; a gearbox or a machine with a higher pole count needs a correspondingly wider grid, and the samples per revolution must be raised with it so that the angular Nyquist frequency stays above the retained band. Finally, the routing prior and the feasible-state set are properties of the machine and of its maintenance history, not of the signal processing, and must be restated for a new installation.

3.7. Primitive Experts and Physics-Regularized Routing

Following the mixture-of-experts principle [28,43], three matched-capacity encoders process the vibration branch (raw + envelope order spectra), the phase-current branch and the Clarke branch. Each encoder is a two-layer perceptron (128 and 96 units, LayerNorm, GELU, dropout 0.12/0.08) followed by a nine-dimensional head, so every modality produces evidence for every primitive. The final logit for primitive j is given by Equation (2):

$$z_j = \sum_m g_{m,j} z_{m,j}, \quad \sum_m g_{m,j} = 1, \quad m \in \{\text{vibration}, \text{abc}, \text{Clarke}\} \quad (2)$$

The prior q_j places 0.90 on the modality indicated by the fault mechanism (vibration for the four mechanical primitives, phase current for broken rotor bar, winding fault and voltage unbalance, Clarke current for the two eccentricity-labeled primitives) and 0.05 on each of the others. The routing weights are initialized at q_j and regularized toward it with a weight of 0.04 while remaining trainable. Training minimizes a positive-weighted binary cross-entropy on the fused logits plus 0.12 times the mean of the per-branch losses, with AdamW [49] (lr 9×10^{-4} , weight decay 2×10^{-4}), batch size 256, gradient clipping at 5.0, for 45 epochs.

3.8. Calibration and MAP Decoding over Feasible Machine States

Segment probabilities are averaged within a recording. Per-primitive temperature parameters T_j are fitted by negative-log-likelihood minimization [47] on source recordings only; neither target-condition recordings nor compound recordings are used. Calibrated probabilities are $\tilde{p}_j = \sigma(\text{logit}(p_j)/T_j)$. The feasible state set is defined by Equation (3):

$$\mathcal{S}_{19} = \{\mathbf{0}\} \cup \{\mathbf{e}_j\}_{j=1}^9 \cup \{\mathbf{e}_a + \mathbf{e}_b : (a, b) \in \mathcal{A}\} \quad (3)$$

where $\mathcal{A}$ is the set of nine compound compositions present in the machine's maintenance history. The diagnosis is then obtained by Equation (4):

$$\hat{s} = \underset{s \in \mathcal{S}}{\text{argmax}} \sum_{j=1}^9 [s_j \log \tilde{p}_j + (1 - s_j) \log (1 - \tilde{p}_j)] \quad (4)$$

Because the all-zero state and the single-fault states are candidates, the fault count is an output rather than an input. Two weaker configurations are reported as controls: $\mathcal{S}_{46}$, which admits all 36 unordered pairs, and an ontology-free rule $\hat{y}_j = 1[\tilde{p}_j \geq \tau]$ with τ selected to maximize macro-F1 on source recordings.

3.9. Evaluation Protocol

Six folds are used; in each, all recordings of one speed-load combination are held out. Training uses the healthy and single-fault recordings of the remaining five combinations. Evaluation covers all recordings of the held-out combination—1 healthy, 13 or 14 single-fault and 9 compound, 143 recordings in total across folds—and reports exact match over the nine primitives, micro-F1, Jaccard index, and fault-count accuracy, following standard multi-label evaluation practice [41,42]. Three random seeds are used, giving 18 independent training runs per configuration. Statistical comparisons average seeds within a condition and then apply a paired Wilcoxon signed-rank test and a condition-level bootstrap over the six folds.

For the analysis of operating regime, the criterion is defined as follows. Each 16-revolution segment carries the 16 instantaneous shaft speeds measured between its consecutive key-phase edges, and two statistics are computed from them: the slope of a least-squares straight line fitted to those speeds against time, expressed in rpm per second, and their coefficient of variation, defined as the population standard deviation divided by the mean. A segment is classified as steady when the absolute slope is below 25 rpm/s and the coefficient of variation is below 0.025; otherwise it is classified as accelerating when the slope is at least +25 rpm/s, and as decelerating in every remaining case. Applied to the 13,720 segments this gives 9349 steady (68.1%), 2065 accelerating (15.1%) and 2306 decelerating (16.8%), and the steady share ranges from 63.0% to 71.8% across the six operating combinations. In 257 segments, 1.9% of the total and 11.1% of the decelerating class, the coefficient of variation exceeds its threshold while the absolute slope does not: the speed varies inside the segment but not monotonically, typically because the segment straddles a turning point or a step of the speed profile, so a straight-line fit returns a small slope.

For raw-signal robustness, additive white Gaussian noise is injected into the held-out sensor waveforms before order tracking at 20, 10, 5, 0, -5 and -10 dB SNR. Three independent noise realizations are drawn at each level, and each realization is evaluated with all three independently trained checkpoints over the six folds, so the 54 measurements available at each level span both training variability and noise-realization variability and the confidence intervals reported in Section 4.5 are computed over them. Key-phase robustness is evaluated by randomly dropping 1% or 5% of detected edges, inserting 1% spurious edges, or adding zero-mean Gaussian timing jitter with a standard deviation of 0.1%, 0.5%, 1% or 2% of the mean revolution period. Because edge perturbations alter the valid-segment list, each perturbed recording is evaluated by averaging over all of its valid perturbed segments rather than reusing the clean positional segment indices.

All results reported in this paper were obtained from a single end-to-end re-execution of the pipeline on Linux using Python 3.12.3, NumPy 2.4.4, pandas 3.0.2, SciPy 1.17.1, scikit-learn 1.8.0, and PyTorch 2.13.0. Computation was performed on a CPU with one PyTorch intra-op thread and one inter-op thread. Random seeds 0, 1, and 2 were run independently. Generative-AI tools were used for exploratory code review, literature screening, and language editing, but not to generate experimental data or replace the scripted analyses.

4. Results

4.1. Front-End and Representation Ablation

Table 3 reports the conventional protocol in which the decoder is told that two faults are present, so that the effect of the front end can be isolated from the effect of removing the fault-count prior.

Three observations follow from Table 3. First, restoring the native 0.25-order resolution raises exact recovery by 8.7 percentage points relative to B0. Second, removing the speed-tracking anti-alias filter from the proposed front end lowers exact recovery by 3.7 percentage points. Third, retaining the zero-sequence channel or using the absolute spectrum does not reduce clean-condition accuracy; the absolute-spectrum control is in fact higher on exact pair recovery. The scale-invariant and zero-sequence-removal choices are therefore not justified as clean-accuracy optimizers, but by their much smaller noise-induced feature shift (Figure 5) and the robustness analysis of Section 4.5. The admissible-pair ontology contributes only 1.2 points with the proposed front end, so the two highest unconstrained primitives are already an admissible composition in almost every recording. Because Table 3 supplies the fault count, it is reported as a controlled front-end ablation rather than as the primary endpoint; the primary evaluation is the cardinality-free protocol in Table 4, where the proposed pipeline achieves the highest overall exact-match accuracy, overall micro-F1, fault-count accuracy, and compound exact-match accuracy.

Table 4. Diagnosis without a fault-count prior. Every recording of the held-out operating condition is evaluated, healthy, single-fault and compound alike, and the decoder chooses among the 19 physically feasible machine states so that the number of coexisting faults is an output rather than an input. Fault-count accuracy is the fraction of recordings for which the predicted number of primitives is correct.

Method	All Recordings, Exact	All, Micro-F1	Fault-Count Accuracy	Compound, Exact	Healthy + Single, Exact
Flat multi-class + top-2 decomposition	0.5920	0.8202	0.5943	0.0247	0.9362
Semantic/attribute embedding	0.6105	0.8079	0.6105	0.0000	0.9810
Cross-talk multi-output	0.6199	0.8135	0.6223	0.0062	0.9923
Proposed routing + front end B0	0.6246	0.7745	0.6315	0.1111	0.9362
Proposed, without envelope order spectrum	0.7314	0.8498	0.7661	0.3951	0.9370
Proposed (full)	0.7665	0.8810	0.8083	0.4568	0.9556

4.2. Cardinality-Free Diagnosis

Table 4 is the main result. The fault count is not supplied; the decoder chooses among the 19 feasible machine states, and every recording of the held-out condition is evaluated.

The three comparison methods reproduce the architectural principle of established approaches within this protocol. They are re-implementations rather than the original implementations, and are reported as such. All three separate healthy and single-fault recordings almost perfectly (0.936–0.992), yet essentially never emit a two-fault decision, so their compound exact-match ranges from 0.000 to 0.025. Their fault-count accuracies are close to their overall exact-match accuracies, while exact recovery of compound recordings remains near zero. These results indicate that failure to infer fault multiplicity is a major limitation of the comparison methods under the cardinality-free protocol.

Against front end B0, the proposed pipeline improves overall exact match by +0.142 (95% condition-level bootstrap interval +0.111 to +0.176, Wilcoxon $p = 0.0312$) and compound exact match by +0.346 (+0.259 to +0.414, $p = 0.0312$); fault-count accuracy rises from 0.6315 to 0.8083. The improvement is positive in all six held-out speed-load combinations (Table 5).

Table 5. Per-condition results of the paired comparison between the proposed front end and front end B0 under cardinality-free decoding, averaged over the three seeds. Each row is one held-out speed-load combination, so the six rows are the six folds of the leave-one-combination-out protocol.

Held-Out Combination	Proposed, All Exact	B0, All Exact	Δ	Proposed, Compound	B0, Compound
20 N·m–1000 rpm	0.819	0.653	+0.167	0.519	0.074
20 N·m–2000 rpm	0.806	0.639	+0.167	0.481	0.148
20 N·m–3000 rpm	0.710	0.609	+0.101	0.259	0.111
40 N·m–1000 rpm	0.792	0.681	+0.111	0.481	0.148
40 N·m–2000 rpm	0.722	0.625	+0.097	0.593	0.185
40 N·m–3000 rpm	0.750	0.542	+0.208	0.407	0.000

Comparing the last two rows of Table 4 isolates the envelope order spectrum, which contributes +3.5 percentage points overall and +6.2 percentage points in compound recordings. Figure 6 summarizes Table 4 graphically.

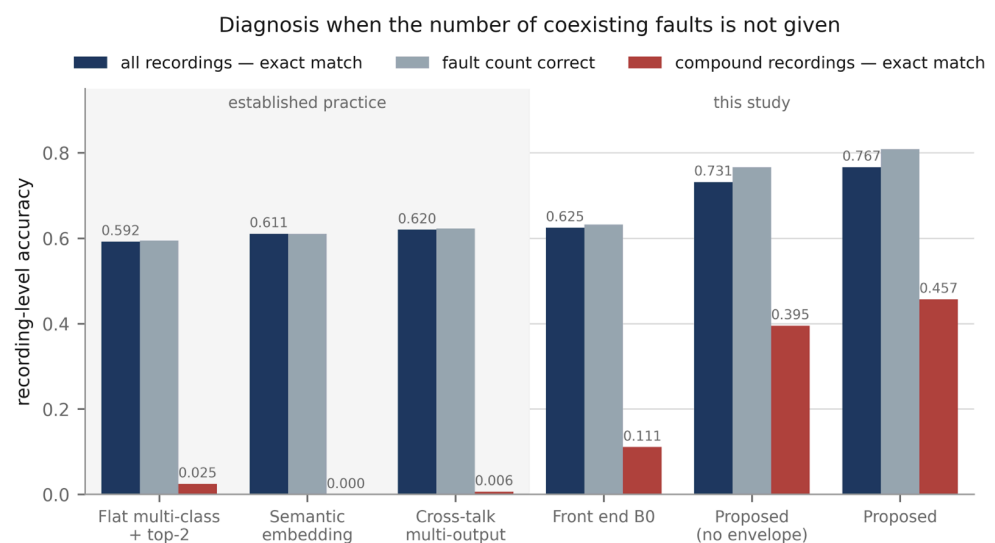


Figure 6. Diagnosis without a fault-count prior, evaluated on every recording of the held-out operating condition. The three comparison methods separate healthy and single-fault recordings well but essentially never emit a two-fault decision.

4.3. Decoder Ablation

To separate the contribution of the state ontology from that of the representation and routing, Table 6 compares source-selected threshold decoding, unrestricted 46-state MAP decoding, and the proposed 19-state feasible-state MAP decoder.

Table 6. Effect of the state ontology on cardinality-free diagnosis with the proposed front end. The three decoders receive identical calibrated primitive probabilities and differ only in how a machine state is selected from them: an independent threshold fitted on source recordings, maximum a posteriori inference over all 46 states formed by the healthy state, the nine single faults and all pairs, and maximum a posteriori inference restricted to the 19 physically feasible states.

Decoder	All Exact	Compound Exact	Fault-Count Accuracy
Threshold τ selected on source data	0.7114	0.4383	0.7510
46-state MAP (healthy + singles + all pairs)	0.7526	0.4568	0.7990
19-state MAP (feasible states)	0.7665	0.4568	0.8083

Narrowing the ontology from 46 to 19 states leaves compound recovery unchanged at 0.4568 and improves overall exact match by 1.4 percentage points. The performance therefore originates primarily in the representation and routing rather than in prior knowledge of the admissible compositions.

The decoder searches a finite set of machine states, so its behavior on a combination that the set does not contain has to be characterized rather than assumed. This was measured directly: for every held-out compound recording, the true composition was deleted from the 19-state set and the decoder was re-run, which reproduces the situation of a genuinely novel combination, over 162 recording decisions. The degradation is graceful in one respect and misleading in another. In 95.7% of cases, the returned state still contains at least one true constituent, and in 84.6%, the decoder retreats to a single fault that is a subset of the truth rather than asserting a different pair; a wrong pair is returned in 13.0% of cases and the healthy state in 2.5%. Mean constituent recall is 0.478, so on average, one of the two faults survives. What does not degrade gracefully is the decoder's own confidence. The median log-likelihood margin between the best and the second-best state rises from 1.04 with the true state present to 1.80 with it removed, because deleting the correct explanation also removes its closest competitor. The margin is therefore not a novelty detector, and a large margin cannot be read as evidence that the returned state is right. This is an intrinsic limitation of structured decoding over a closed state set: the decoder can be wrong confidently, and it has no way to express that none of its states fits. Detecting a genuinely unlisted combination requires a signal outside the decoder, such as an explicit open-set score or an absolute likelihood threshold, and is not solved here.

4.4. Where the Modality Assignment Comes From

To test whether the primitive-to-modality assignment can be recovered from source-domain performance alone, Table 7 compares the proposed physics prior with source-selected and dynamic-eccentricity reassignment controls.

The source-only control holds out one source condition as an inner validation domain, trains a single-modality expert per branch, and assigns each primitive to the modality with the highest validation average precision. It uses no target condition and no compound data. It agrees with the physics prior for only four of the nine primitives, namely the three bearing primitives and the bent shaft. Broken rotor bar, dynamic eccentricity, static eccentricity and winding fault are instead all assigned to vibration, whose source-domain

average precision is high for the isolated-fault task, and voltage unbalance is assigned to the Clarke branch. Compound recovery nevertheless falls from 0.5926 to 0.2778.

Table 7. Routing controls evaluated on front end B0 over six folds with one seed. All four configurations share the same features, the same folds and identical one-hot weighting, and differ only in where the primitive-to-modality map comes from: a source-domain average-precision criterion, two reassignments of the dynamic-eccentricity-labeled primitive, and the proposed physics prior.

Assignment	Exact Pair	Ontology-Free Top-2
Source-domain validation AP selection	0.2778	0.1852
Physics prior, dynamic eccentricity → vibration	0.5370	0.4630
Physics prior, dynamic eccentricity → phase current	0.5370	0.4815
Physics prior as proposed	0.5926	0.5000

The reason is visible in the source-domain average precisions themselves: vibration is sufficient to separate every isolated fault, so a criterion based on isolated-fault performance cannot distinguish sufficiency from compositional transferability. The physics prior resolves an underdetermination that the source task leaves open.

The dynamic-eccentricity controls also settle the nomenclature caveat of Section 3.1 empirically: reassigning that primitive to vibration or to the phase current lowers exact recovery from 0.593 to 0.537 in both cases, so the Clarke-current assignment is retained on measured grounds rather than on the strength of its label. These four assignments are compared in Figure 7.

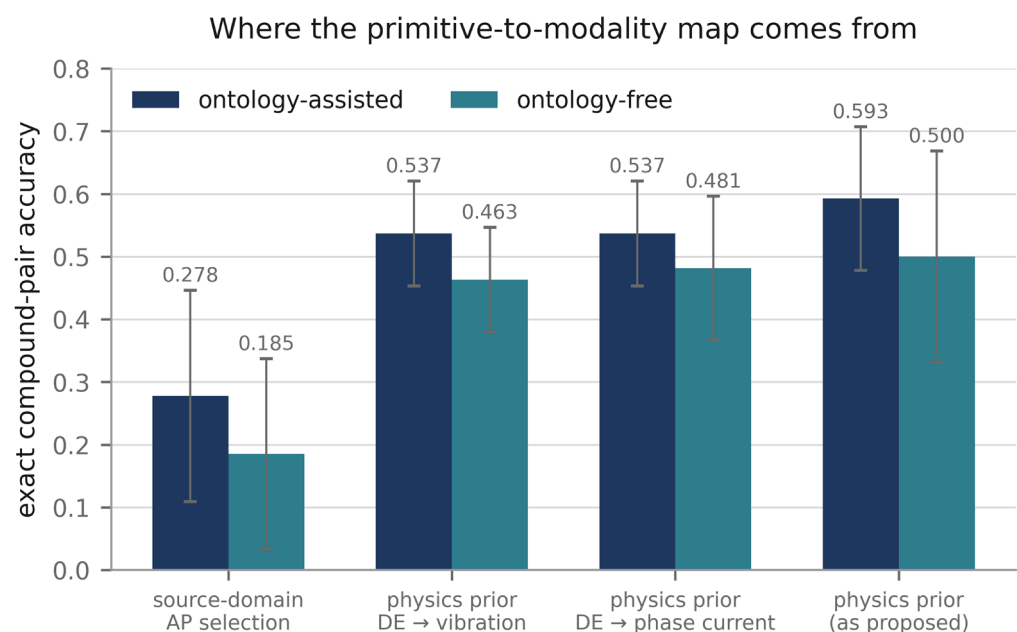


Figure 7. Effect of the source of the modality assignment on compound recovery. All four configurations share the same one-hot weighting, the same features and the same folds; only the primitive-to-modality map differs.

The routing weights are initialized at the physics prior and then updated under an L2 penalty toward that prior, so two contributions are conflated: the prior itself and the subsequent data-driven adjustment. Table 8 separates these contributions using two additional controls. In the frozen control, the weights are held at the smoothed physics prior for the whole of training and are never updated; in the free control, they are initialized uniformly and trained without any penalty, so the assignment is learned entirely from

source data. Freezing the weights costs 0.5 percentage points of overall exact accuracy and gains 1.9 percentage points of compound recovery, neither of which is significant (Wilcoxon signed-rank over 18 paired runs, $p = 0.72$ and $p = 0.20$). The learned weights explain why: they move from 0.90 to 0.92 on the modality named by the prior and reproduce the prior assignment for all nine primitives, so the trainable gate is doing almost nothing. Removing the prior instead is decisive. The freely learned gate converges to a nearly uniform mixture, 0.42 on vibration against 0.28 and 0.29 on the two current branches, agrees with the physics assignment for only five of nine primitives, and loses 38.9 percentage points of compound recovery ($p = 0.0002$); fixed uniform weighting loses 42.0 points. Essentially all of the benefit of the routing mechanism is therefore attributable to the physical prior and essentially none of it to the data-driven refinement.

Table 8. Separate contributions of the physics prior and of the trainable gate to the routing mechanism, on the proposed front end over six folds and three seeds. The first three columns use the cardinality-free protocol of Table 4; the last column uses the conventional protocol of Table 3, in which the decoder is told that two faults are present. The p -values are Wilcoxon signed-rank tests against the proposed configuration over the 18 paired runs.

Routing Weights	All Recordings, Exact	Compound, Exact	Fault-Count Accuracy	Exact Pair (Count Supplied)
Physics prior, initialized and trained (proposed)	0.7665	0.4568	0.8083	0.6790
Physics prior, frozen throughout training	0.7619	0.4753 ($p = 0.20$)	0.8013	0.6667
Uniform, frozen	0.6061	0.0370 ($p = 0.0002$)	0.6061	0.7222
Freely learned, no physics prior	0.6385	0.0679 ($p = 0.0002$)	0.6431	0.7593

One observation from the same experiment is worth recording because it is easy to misread. Under the conventional protocol of Section 4.1, in which the decoder is told that two faults are present, the freely learned gate reaches 0.7593 exact pair recovery and uniform weighting reaches 0.7222, both above the proposed routing. The ordering reverses completely once the fault count is withheld. An unspecialized mixture produces evidence that ranks the two true primitives highly but is not calibrated well enough for the decoder to prefer a two-fault state over a single-fault state, which is invisible when the cardinality is supplied and fatal when it is not. This is the same effect that separates the comparison methods in Table 4 from the proposed pipeline, and it reinforces the central argument of this paper: a protocol that supplies the fault count can rank methods in the opposite order to a protocol that does not.

4.5. Robustness and Deployment Behavior

Additive sensor noise: Table 9 applies additive white Gaussian noise to the raw time-domain sensor signals of the held-out compound recordings before order tracking, which is the physically meaningful injection point because the corruption then passes through key-phase detection, angular resampling and envelope demodulation exactly as a real sensor fault would. Three independent noise realizations are evaluated at each of six SNR levels (20, 10, 5, 0, -5 , and -10 dB), using three trained checkpoints over six folds, giving 324 evaluations in total. Table 9 reports 95% confidence intervals over folds, model seeds, and noise realizations. The spread between realizations is not negligible: the mean of a single realization differs from the mean of another by up to 8.6 percentage points at the same signal-to-noise ratio, which is of the same order as several of the ablation effects

reported elsewhere in this paper. Single-realization robustness figures should be treated as point estimates with that uncertainty attached.

Table 9. Exact compound-pair recovery of the proposed pipeline under additive white Gaussian noise injected into the raw time-domain sensor signals before order tracking. Three independent noise realizations per level, each evaluated with three trained checkpoints over six folds, give 54 measurements per level; the bracketed interval is a 95% confidence interval over those measurements. Retention is the mean exact rate divided by the clean value in the first row. SNR, signal-to-noise ratio; CI, confidence interval.

SNR	Exact Pair (95% CI)	Micro-F1	Retention vs. Clean
clean	0.6790	0.8395	100.0%
20 dB	0.4136 [0.3752, 0.4519]	0.7068	60.9%
10 dB	0.2407 [0.2078, 0.2737]	0.6204	35.5%
5 dB	0.2531 [0.2271, 0.2791]	0.6255	37.3%
0 dB	0.2202 [0.1854, 0.2549]	0.5998	32.4%
−5 dB	0.1790 [0.1448, 0.2133]	0.5401	26.4%
−10 dB	0.1811 [0.1515, 0.2106]	0.5000	26.7%

Three features of the curve matter. The pipeline is substantially sensitive to broadband raw-signal corruption: exact pair recovery falls to 0.41 at 20 dB, which is 61% of the clean value, and the informative part of the degradation is concentrated between 20 dB and 10 dB rather than below it. Performance then saturates rather than collapsing. Between 10 dB and −10 dB, the exact rate stays between 0.18 and 0.25 and constituent micro-F1 stays between 0.50 and 0.63, and even at −10 dB, where the noise power is ten times the signal power, the exact rate of 0.18 remains above the 0.11 expected from a uniformly random choice among the nine admissible pairs. The residual skill comes from the current branches, which are band-limited and therefore lose a smaller fraction of their in-band energy to broadband noise than the vibration branch does. The practical reading is that the operating margin of this pipeline against broadband sensor corruption lies above roughly 20 dB; below it the diagnosis retains some information but is no longer trustworthy at the recording level. This is also the point at which the acquisition conditions of the benchmark matter, since the recordings were made in a laboratory with low electromagnetic interference [48] and an industrial installation would start from a poorer signal-to-noise ratio.

Key-phase perturbation: Table 10 tests missed pulses, false pulses and timing jitter in the shaft reference. Drop and false-edge rates are percentages of detected key-phase edges; jitter is the Gaussian timing standard deviation expressed as a fraction of the mean revolution period. Because edge perturbation changes the valid-segment list, each perturbed recording is scored over all of its valid perturbed segments, whereas the clean pipeline uses the balanced subsample of the frozen segment index. The rows of Table 10 are therefore comparable with one another but not with the clean value of 0.6790 in Table 3. Read within Table 10, exact pair recovery is essentially flat from 0.1% to 1% timing jitter and across 1–5% edge drops and 1% false edges, and falls sharply to 0.5123 at 2% jitter. The edge-drop rows should not be read as robustness in the ordinary sense: quality control removes the segments whose intra-segment speed ratio is corrupted by a merged revolution, so a 5% drop rate leaves only 130.6 valid segments per recording against 298.2 for the jitter cases. Quality control converts degradation into discarded data rather than into wrong answers, which is the desirable behavior but is not the same as tolerating the perturbation.

Table 10. Key-phase perturbation robustness of the proposed pipeline. Rows are mutually comparable but are not comparable with the clean values in Table 3, because perturbed recordings are scored over all valid perturbed segments rather than over the balanced subsample of the frozen segment index. The last column reports the mean and range of the number of segments that survive quality control.

Perturbation	Exact Pair	Micro-F1	Mean Valid Segments/Record
1% edge drop	0.6790 $\pm$ 0.1420	0.8395	248.5 (68–388)
5% edge drop	0.6852 $\pm$ 0.1488	0.8426	130.6 (36–209)
1% false edges	0.6914 $\pm$ 0.1403	0.8457	254.5 (76–402)
0.1% timing jitter	0.6914 $\pm$ 0.1403	0.8457	298.2 (88–469)
0.5% timing jitter	0.7469 $\pm$ 0.0638	0.8735	298.2 (88–469)
1% timing jitter	0.6914 $\pm$ 0.0898	0.8457	298.2 (88–469)
2% timing jitter	0.5123 $\pm$ 0.1214	0.7531	298.2 (88–469)

Sensor and modality loss: Table 11 removes branches at inference time and renormalizes the routing weights over the remaining modalities. Removing vibration is the most damaging: ontology-free top-2 recovery falls from 0.6667 to 0.0556. Removing the Clarke branch lowers exact recovery from 0.6790 to 0.5494, a loss of 13.0 percentage points. Removing the phase-current branch, however, increases exact recovery to 0.7469.

Table 11. Inference with missing sensors. Branches are removed at inference time from the proposed trained pipeline, and the routing weights are renormalized over the modalities that remain, so no retraining is involved, and each row measures how much of the diagnosis survives the loss of a sensing channel.

Available Branches	Exact Pair	Micro-F1	Ontology-Free Top-2
all three	0.6790 $\pm$ 0.1420	0.8395	0.6667
vibration + Clarke (no phase current)	0.7469 $\pm$ 0.1653	0.8735	0.7469
vibration + phase current (no Clarke)	0.5494 $\pm$ 0.1042	0.7747	0.5370
vibration only	0.6728 $\pm$ 0.1399	0.8364	0.6543
current only (no vibration)	0.3333 $\pm$ 0.1476	0.6296	0.0556

This asymmetry is consistent with the branch ablation pattern across folds and seeds: with the proposed representation, the phase-current branch is largely redundant with the Clarke branch, and the two-branch configuration (vibration + Clarke) is the stronger deployment choice in this benchmark. The three-branch model is retained as the main configuration because the routing prior is defined over all three representations and because the phase-current branch carries direct electrical evidence for several primitives, but the claim that all three sensing representations are individually necessary is not supported by these data.

The scope of this conclusion needs to be stated precisely because it is easy to over-generalize. What the experiment shows is that in this benchmark, with these nine primitives and this sensor layout, the phase-current branch is redundant given the Clarke-transformed current: the Clarke branch is a linear function of the same three phase currents and retains the components that separate the electrical primitives present here, so the extra branch adds parameters and correlated inputs without adding information. It does not follow, and is not claimed, that phase-current measurement is unnecessary for motor compound-fault diagnosis in general. Phase quantities carry information that the Clarke transform discards by construction, specifically the per-phase asymmetry that identifies which phase is affected. Faults whose signature is precisely that asymmetry, such as inter-turn short circuits localized to one phase, supply-voltage unbalance, an open phase, connection or contactor faults, and unbalanced cable impedance, are naturally diagnosed from the

individual phase currents, and severity estimation for those faults generally requires them. The benchmark includes slight and severe winding-fault grades, whereas voltage unbalance is represented only at the slight grade. The finding should be read as an argument for auditing modality redundancy on the fault set at hand, not as a recommendation to remove phase-current sensing.

Operating-envelope extrapolation: Table 12 removes an entire speed or an entire load level from training, which is strictly harder than the leave-one-combination protocol.

Table 12. Entire-level holdout, in which one complete speed level or one complete load level is removed from training, which is strictly harder than the leave-one-combination-out protocol used elsewhere. Values are exact pair recovery over three seeds for both front ends.

Level Withheld Entirely	Front End B0	Proposed
1000 rpm	0.4444 ± 0.0556	0.7778 ± 0.0962
2000 rpm	0.7407 ± 0.1156	0.7222 ± 0.0000
3000 rpm	0.4444 ± 0.1470	0.7222 ± 0.0000
20 N·m	0.2469 ± 0.0428	0.5556 ± 0.0370
40 N·m	0.3827 ± 0.0214	0.4815 ± 0.0370

Front end B0 shows a pronounced asymmetry between speed levels: withholding the intermediate level gives 0.741, while withholding either boundary level gives 0.444. Such a pattern is ordinarily read as the difference between interpolating within the observed speed range and extrapolating beyond it. With the anti-aliased front end the three speed-level results fall within a much narrower range, 0.722 to 0.778, consistent with the interpretation that much of the earlier asymmetry arose from speed-dependent spectral fold-back rather than an intrinsic limit on extrapolation: the decimation ratio, and therefore the amount of aliased energy, is largest at the lowest speed and smallest at the highest. Load extrapolation remains the harder case, because the benchmark provides only two load levels so that withholding either is always an extrapolation, but it improves from 0.247 and 0.383 with B0 to 0.556 and 0.482 with the proposed front end.

The asymmetry between speed and load extrapolation has a physical explanation and is not merely a numerical observation. Order normalization is exact for the effect of speed: a change in shaft speed rescales the time axis, and resampling at uniform shaft-angle increments removes that rescaling, so a speed level that was not trained on differs from the training levels mainly in the position of the structural resonances relative to the order axis. Load is not removed by any such transformation. Raising the torque increases the stator current and the slip, shifts the operating point on the magnetization curve, and raises the radial bearing reaction force. These changes affect the amplitudes and, for some fault signatures, the order-domain locations of the components: the rolling-element impact force scales with the contact load, so the envelope order amplitudes of the raceway defects grow with torque; the slip-dependent sideband spacing of the broken-bar signature moves with load; and the eccentricity-related components scale with the radial magnetic pull, which itself depends on the flux level. The scale-invariant representation of Section 3.6 removes the channel-wide gain but not these relative changes between components, so a held-out load level presents the classifier with a feature distribution that is shifted along directions the training set never spanned. With only two load levels in the benchmark, withholding one leaves a single level for training and no basis on which the load dependence could be interpolated, which is why load extrapolation is the harder of the two settings and why its result should be read as a bound on the operating envelope rather than as a property of the method.

4.6. Operating Regime and Error Structure

To determine whether the remaining errors are associated with transient operation, Table 13 stratifies compound recovery by steady, accelerating, and decelerating regimes.

Table 13. Compound recovery of the proposed pipeline stratified by operating regime, with the fault count not supplied. Segments are assigned to the steady, accelerating or decelerating regime by the speed criterion of Section 3.9, and each recording is scored using only the segments of that regime.

Regime	Proposed, Compound Exact	Proposed, Micro-F1
Steady	0.5085	0.8057
Accelerating	0.3580	0.7682
Decelerating	0.3272	0.7462

Within the final proposed pipeline, steady-speed plateaus are the most informative (0.5085 compound exact, micro-F1 0.8057), while accelerating and decelerating intervals are harder and similar (0.3580 and 0.3272 exact, respectively). This indicates that the cardinality-free compound decision remains sensitive to transient operating regimes even after order normalization. To localize the residual compositional failures, Table 14 further breaks exact recovery down by the nine unseen compound compositions.

Table 14. Exact recovery of the proposed pipeline for each of the nine unseen compound compositions, with the fault count not supplied, ordered from the hardest composition to the easiest. Each value is the mean over six folds and three seeds of the fraction of that composition's recordings whose two constituents are both identified and nothing else is added.

Composition	Exact
bearing inner + bearing outer	0.000
bearing outer + dynamic eccentricity	0.111
bearing inner + dynamic eccentricity	0.167
bearing inner + broken rotor bar	0.500
bearing inner + winding fault	0.500
bearing outer + static eccentricity	0.556
bearing outer + winding fault	0.667
bearing inner + static eccentricity	0.722
bearing outer + broken rotor bar	0.889

Figure 8 visualizes the same composition-wise results, highlighting the strong recovery of several cross-subsystem pairs and the unresolved simultaneous inner-race and outer-race bearing defects.

The composition that is never recovered is the one in which both primitives belong to the same bearing, and the mechanism can be identified from the measurements rather than left to discussion. Two effects act together. The first is spectral. With the bearing geometry of this rig, the outer-race and inner-race defect frequencies are 3.585 and 5.415 orders, and the inner-race component is amplitude-modulated at shaft rate so that it appears together with sidebands at plus and minus one and two orders. On the 0.25-order grid used here, four of the eight outer-race harmonics inside the retained band fall in a bin that is also occupied by an inner-race harmonic or one of its sidebands; the third outer-race harmonic at 10.755 orders and the second inner-race harmonic at 10.830 orders are separated by 0.075 order, less than half a bin. Half of the outer-race evidence is therefore not separable from inner-race evidence at this resolution. The second effect is a consequence and is visible in the model output. For each primitive, the calibrated evidence it receives when it occurs alone can be compared with the evidence it receives inside a compound. Averaged over folds and seeds, the outer-race primitive attains probability 0.999 on its own single-fault

recordings and 0.025 when an inner-race defect is also present, a suppression of 0.975, whereas the inner-race primitive falls only from 1.000 to 0.687 in the same recordings. The stronger and denser inner-race sideband family masks the outer-race comb, and not the reverse. Across all nine compositions, the mean suppression is 0.644 within one subsystem against 0.182 across subsystems, and suppression is negatively correlated with exact recovery (Pearson $r = -0.64$). Compound faults that span two subsystems are recovered well because their evidence is carried by different modalities and different order families; compound faults inside one subsystem compete for the same evidence, and the weaker signature is extinguished.

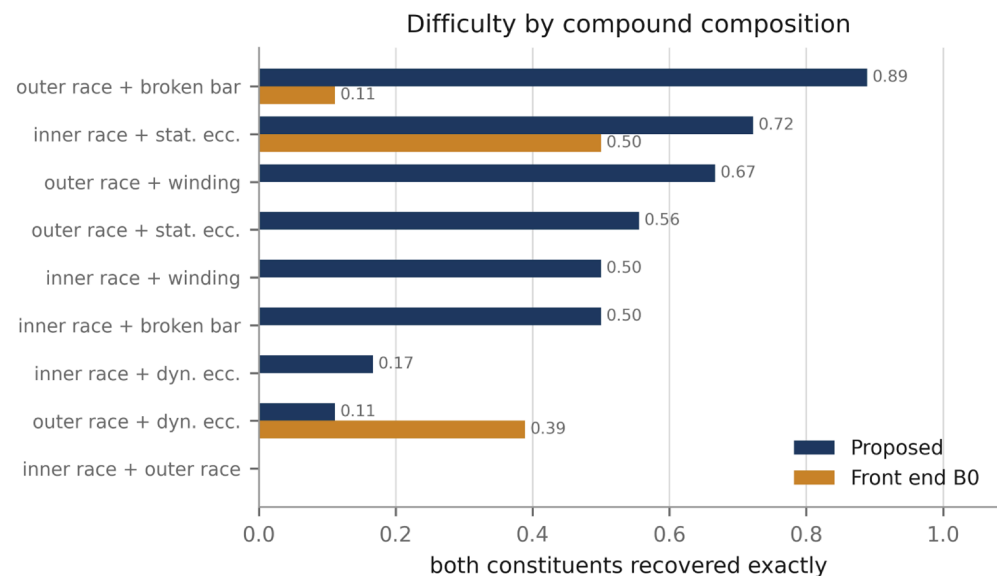


Figure 8. Exact recovery per compound composition for the proposed pipeline and the conventional front end B0. Simultaneous inner- and outer-race defects remain unrecovered. Abbreviations: stat. ecc., static eccentricity; dyn. ecc., dynamic eccentricity.

4.7. Computational Cost

Online deployment is limited by the cost of the signal path, not by the cost of inference. Table 15 reports the measured time of every stage on a single CPU core with no GPU, for a complete 90 s six-channel recording that yields 469 analysis segments. The whole path from raw samples to a decoded machine state takes 1.93 s, a real-time factor of 46.5, and the mean cost per segment is 4.1 ms. Inference accounts for 0.2% of that total: the routed model has 270,902 parameters, occupies 1.08 MB in single precision, and takes 7.9 microseconds per segment, against 5.2 for a vibration-only expert of matched capacity and 8.0 for early fusion, so the routing structure costs essentially nothing relative to a single fused encoder. The signal path dominates and is itself dominated by two operations: the speed-binned anti-aliased resampling at 51% of the total and the band-pass and Hilbert envelope path at 42%. Both are the computational costs of the two front-end components whose contributions are quantified in Tables 3 and 4, and both are ordinary filtering operations that parallelize over channels and over speed bins. A single core therefore processes about 46 machines in real time in this configuration; on a continuous monitoring installation, the practical constraint is the 16-revolution segment itself, which at 1000 rpm corresponds to a decision latency of about 1 s of acquisition, an order of magnitude above the processing time.

Table 15. Measured cost of each stage of the deployed path on a single CPU core with no GPU, for one complete 90 s six-channel recording that yields 469 analysis segments. Shares are of the total excluding the file read. The real-time factor is the duration of the signal divided by the time taken to process it.

Stage	Time (s)	Per Segment (ms)	Share (%)
Key-phase edge detection	0.025	0.054	1.3
Segment quality control	<0.001	0.001	0.0
Anti-aliased angular resampling	0.991	2.114	51.2
Band-pass, Hilbert envelope and resampling	0.821	1.750	42.4
Order spectra of all branches	0.094	0.200	4.9
Expert inference, whole recording	0.003	0.006	0.2
MAP decoding	<0.001	<0.001	0.0
Total (real-time factor 46.5)	1.935	4.125	100.0

5. Discussion

Isolated-fault performance does not predict compositional generalization, and the conventional protocol conceals the size of the gap. Every configuration examined here separates isolated faults well, yet compound recovery under cardinality-free decoding ranges from 0.000 to 0.457 across the methods in Table 4. Reporting isolated-fault accuracy, or compound accuracy with a fault-count prior, therefore overstates deployability.

The fault-count prior, not the composition ontology, carries much of the conventional protocol. With front end B0, compound recovery falls from 0.6481 under the supplied-two-fault protocol to 0.1111 when the decoder must also decide how many faults are present; with the proposed front end, it falls from 0.6790 to 0.4568. In contrast, weakening the composition ontology from 19 to 46 states leaves compound recovery unchanged at 0.4568. Evaluations of compound diagnosis should therefore state explicitly whether fault multiplicity is supplied.

Modality specialization must come from physics because the source task cannot supply it. Vibration attains source-domain average precisions of 0.795–0.968 for every primitive, so a data-driven selection criterion routes almost everything to vibration and compound recovery collapses. The point is not that current measurements are intrinsically superior for those primitives, but that the evidence they provide is what survives superposition with a bearing defect. Section 4.4 sharpens this: freezing the physics prior costs nothing measurable, whereas removing it costs 38.9 percentage points of compound recovery, so the benefit is attributable to the prior and not to the trainable gate.

Sensor necessity is not symmetric, and the phase-current branch is redundant in this benchmark. Removing vibration reduces ontology-free top-2 recovery from 0.6667 to 0.0556, whereas removing the phase-current branch increases exact recovery from 0.6790 to 0.7469. Removing the Clarke current instead lowers exact recovery to 0.5494. Thus, a two-branch deployment, vibration plus Clarke, is both smaller and slightly more accurate here. This does not weaken the routing argument, which concerns which representation supplies the evidence for each primitive, but it does mean that a claim of joint necessity for all three representations would be unsupported.

The residual failures are concentrated within one physical subsystem. Simultaneous inner- and outer-race defects are never recovered exactly, and the two compositions pairing a bearing defect with the dynamic-eccentricity-labeled condition are recovered in at most 17% of recordings. Section 4.6 identifies the mechanism: half of the outer-race harmonics share a 0.25-order bin with an inner-race component, and the outer-race evidence is suppressed from 0.999 to 0.025 when both defects are present. Combining evidence across physically distinct subsystems is comparatively easy; separating two overlapping defect

signatures inside the same bearing is not. This within-subsystem compositional ambiguity is a natural target for future work, for example through per-primitive envelope bands selected by spectral kurtosis [34,35] or cyclostationary indicators that separate inner- and outer-race modulation structures [36].

Limitations

All nine compound compositions in the benchmark contain at least one rolling-element bearing defect, and three primitives—bearing ball defect, bent shaft and voltage unbalance—never appear in a compound recording. Compositional generalization is therefore evaluated on only a subset of the primitive vocabulary. The evaluation uses 54 compound recordings across six operating conditions, which limits statistical resolution. The comparison methods were reimplemented within the present protocol rather than evaluated using the original authors' implementations.

Entire-level holdout remains difficult for load extrapolation, with exact pair recovery ranging from 0.482 to 0.556. Raw-signal additive white Gaussian noise causes substantial degradation even at 20 dB. Three independent noise realizations per level are evaluated, and Table 9 reports 95% confidence intervals across folds, model seeds, and noise realizations. The difference between realization-specific mean recovery rates reaches 8.6 percentage points, showing that single-realization estimates can be sensitive to the particular noise draw.

Synthetic key-phase perturbations are tolerated reasonably well up to 1% timing jitter, whereas 2% jitter reduces exact pair recovery to 0.5123. Because perturbed recordings are scored over a different segment population from the clean balanced subsample, the perturbation rows support relative comparisons among themselves rather than direct comparison with the clean result in Table 3. Complete tachometer loss and tachless phase estimation [31,32] were not evaluated. Source-side noise augmentation is also absent from the core training protocol.

The finite-state MAP decoder is closed by construction. As shown in Section 4.3, removing the true composition from its candidate set causes the decoder to return a subset of the true faults in 84.6% of cases, while its confidence margin increases rather than decreases. The margin therefore cannot be used as a detector of unlisted compositions. The observed redundancy of the phase-current branch is specific to the benchmark, fault set, and sensor layout and should not be generalized beyond the conditions evaluated here.

Finally, every compound recording in the benchmark contains exactly two primitives. The compositional claim is therefore established only for two-fault combinations, although the formulation can accommodate higher-order combinations without changing the expert architecture.

6. Conclusions

This study examined whether compound motor faults can be diagnosed when no compound recording is available for training and the number of coexisting faults is unknown at inference time. An anti-aliased key-phase-guided order-domain front end, a band-pass envelope order spectrum, a scale-invariant spectral representation, physics-regularized primitive-specific modality routing, and MAP decoding over the physically feasible machine states were combined into a single pipeline trained exclusively on healthy and single-fault data.

Evaluated on every recording of six held-out speed-load combinations, the pipeline identifies the exact machine condition of 76.7% of recordings and both constituents of 45.7% of previously unseen compound recordings, compared with 62.5% and 11.1% for the same model with a conventional order-domain front end and 59.2–62.0% and 0.0–2.5%

for three re-implemented baselines. Removing the fault-count prior reduces compound recovery from 0.6481 to 0.1111 for B0 but from 0.6790 to 0.4568 for the proposed front end, while narrowing the composition ontology from 46 to 19 states leaves compound recovery unchanged. The result therefore depends more strongly on the representation and primitive-specific routing than on prior knowledge of admissible compositions.

Anti-aliasing also changes the interpretation of the operating-envelope experiments: once speed-dependent fold-back is removed, entire-speed-level holdout becomes much more uniform, while load extrapolation remains difficult because the load-dependent changes in the fault-related components are not removed by angular resampling. Robustness tests delimit the present deployment envelope: key-phase edge drops and up to 1% timing jitter are tolerated reasonably well, whereas 2% timing jitter and broadband raw-signal noise cause substantial degradation, with the operating margin against broadband corruption lying above roughly 20 dB. On one CPU core, the complete path from raw samples to a decoded machine state runs 46.5 times faster than real time, so the method is not limited by computation.

Four directions follow from the limitations above, and the first three address weaknesses for which this paper offers no remedy. Load extrapolation is the weakest point of the operating envelope, and the reason is that neither order tracking nor amplitude normalization removes the load dependence of the feature amplitudes. A representation that is load-invariant by construction is therefore more promising than further normalization of the present features; magnetic flux monitoring, whether from a search coil or from an external stray-flux probe, is attractive because the eccentricity and rotor-asymmetry harmonics it carries are referred to the magnetomotive force distribution rather than to the mechanical load path, and because it adds a sensing channel that is inexpensive to retrofit and has recently been shown to separate stator faults without access to the motor terminals [50]. Second, performance during acceleration and deceleration is consistently below the steady-speed plateaus, which is expected from a front end whose segment is a fixed number of revolutions and whose spectrum assumes stationarity within that segment. Time-frequency representations designed for strong non-stationarity, such as adaptive chirplet or reassigned representations and order-domain formulations that model the speed variation within the segment instead of averaging over it, are the natural next step. Third, the compositional vocabulary of the benchmark is incomplete: every compound recording contains at least one rolling-element defect, and three primitives, the bearing ball defect, the bent shaft and the voltage unbalance, never occur inside a compound recording, so compositional generalization is demonstrated over a subset of the primitive vocabulary rather than over all of it. Validation on broader benchmarks, including compositions of three or more primitives and combinations that exclude bearing defects entirely, is required before the compositional claim can be made in general; the large pretrained models now being applied to machinery diagnosis [51] are one route to that breadth, although they do not by themselves supply the cardinality-free decoding studied here. Fourth, the dependence on a dedicated key-phase sensor should be removed. Table 10 shows that the pipeline tolerates timing jitter of up to 1% but degrades sharply at 2%. Tacholeless order tracking estimates the instantaneous phase from the current or vibration signal itself [32], for example through multi-order Viterbi tracking of instantaneous angular speed [52] or automatic ridge estimation on the time–frequency plane [53]. Such an approach would remove the requirement for a dedicated shaft-reference sensor and its associated pulse-detection errors. It is therefore an important direction for deployment on existing installations.

Author Contributions: Conceptualization, T.M.; methodology, T.M.; software, T.M.; validation, T.M. and J.L.; formal analysis, T.M.; investigation, T.M.; data curation, T.M. and J.L.; writing—original draft preparation, T.M.; writing—review and editing, T.M. and J.L.; visualization, T.M.; supervision,

T.M.; project administration, T.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Technology Innovation Program (Materials and Parts Technology Development Program) (RS-2026-00434151, Development and demonstration of integrated operating system technology for steam turbine power generation facilities) funded by the Ministry of Trade, Industry and Resources (MOTIR, Republic of Korea).

Data Availability Statement: The MCC5-THU motor benchmark analyzed in this study is publicly available from the original data descriptor at <https://doi.org/10.17632/6s3dggj9mw.1>; the speed-circulation subset was used. The authors plan to make a curated reproducibility package containing the principal analysis scripts and supporting materials required to reproduce the main results reported in this study available through a public repository following acceptance, with release targeted no later than publication of the accepted article. The repository information and persistent identifier will be provided once the archive has been finalized.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nandi, S.; Toliyat, H.A.; Li, X. Condition monitoring and fault diagnosis of electrical motors—A review. *IEEE Trans. Energy Convers.* **2005**, *20*, 719–729. [[CrossRef](#)]
2. Benbouzid, M.E.H. A review of induction motors signature analysis as a medium for faults detection. *IEEE Trans. Ind. Electron.* **2000**, *47*, 984–993. [[CrossRef](#)]
3. Thomson, W.T.; Fenger, M. Current signature analysis to detect induction motor faults. *IEEE Ind. Appl. Mag.* **2001**, *7*, 26–34. [[CrossRef](#)]
4. Riera-Guasp, M.; Antonino-Daviu, J.A.; Capolino, G.-A. Advances in electrical machine, power electronic, and drive condition monitoring and fault detection: State of the art. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1746–1759. [[CrossRef](#)]
5. Henao, H.; Capolino, G.-A.; Fernandez-Cabanas, M.; Filippetti, F.; Bruzzese, C.; Strangas, E.; Pusca, R.; Estima, J.; Riera-Guasp, M.; Hedayati-Kia, S. Trends in fault diagnosis for electrical machines: A review of diagnostic techniques. *IEEE Ind. Electron. Mag.* **2014**, *8*, 31–42. [[CrossRef](#)]
6. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [[CrossRef](#)]
7. Zhang, S.; Zhang, S.; Wang, B.; Habetler, T.G. Deep learning algorithms for bearing fault diagnostics—A comprehensive review. *IEEE Access* **2020**, *8*, 29857–29881. [[CrossRef](#)]
8. Zhao, M.; Zhong, S.; Fu, X.; Tang, B.; Pecht, M. Deep residual shrinkage networks for fault diagnosis. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4681–4690. [[CrossRef](#)]
9. Fyfe, K.R.; Munck, E.D.S. Analysis of computed order tracking. *Mech. Syst. Signal Process.* **1997**, *11*, 187–205. [[CrossRef](#)]
10. Bonnardot, F.; El Badaoui, M.; Randall, R.B.; Danière, J.; Guillet, F. Use of the acceleration signal of a gearbox in order to perform angular resampling (with limited speed fluctuation). *Mech. Syst. Signal Process.* **2005**, *19*, 766–785. [[CrossRef](#)]
11. Lampert, C.H.; Nickisch, H.; Harmeling, S. Attribute-based classification for zero-shot visual object categorization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *36*, 453–465. [[CrossRef](#)] [[PubMed](#)]
12. Xian, Y.; Lampert, C.H.; Schiele, B.; Akata, Z. Zero-shot learning—A comprehensive evaluation of the good, the bad and the ugly. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 2251–2265. [[CrossRef](#)] [[PubMed](#)]
13. Feng, L.; Zhao, C. Fault description based attribute transfer for zero-sample industrial fault diagnosis. *IEEE Trans. Ind. Inform.* **2021**, *17*, 1852–1862. [[CrossRef](#)]
14. Gao, Y.; Gao, L.; Li, X.; Zheng, Y. A zero-shot learning method for fault diagnosis under unknown working loads. *J. Intell. Manuf.* **2020**, *31*, 899–909. [[CrossRef](#)]
15. Xing, S.; Lei, Y.; Wang, S.; Lu, N.; Li, N. A label description space embedded model for zero-shot intelligent diagnosis of mechanical compound faults. *Mech. Syst. Signal Process.* **2022**, *162*, 108036. [[CrossRef](#)]
16. Xu, J.; Liang, S.; Ding, X.; Yan, R. A zero-shot fault semantics learning model for compound fault diagnosis. *Expert Syst. Appl.* **2023**, *221*, 119642. [[CrossRef](#)]
17. Xu, J.; Zhang, H.; Zhou, L.; Fan, Y. Zero-shot compound fault diagnosis method based on semantic learning and discriminative features. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3518313. [[CrossRef](#)]
18. Chen, H.; Hong, J.; Cheng, C.; Kong, Y.; Zheng, X. Zero-shot compound fault diagnosis method using semantic construction and embedding. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–13. [[CrossRef](#)]

19. Ma, C.; Wang, X.; Li, Y.; Cai, Z. Broad zero-shot diagnosis for rotating machinery with untrained compound faults. *Reliab. Eng. Syst. Saf.* **2024**, *242*, 109704. [\[CrossRef\]](#)
20. Wang, J.; Xu, Z.; Niu, F.; Liu, J.; Zhu, Z. Adaptively weighted semantic autoencoder for zero-shot compound fault diagnosis. *IEEE Sens. J.* **2024**, *24*, 37472–37481. [\[CrossRef\]](#)
21. Song, E.; Zhu, R.; Yao, C.; Ke, Y.; Ning, Y.; Lu, L. Generalized zero-shot bearing compound fault diagnosis based on reserved gate pre-classifier and physical information semantics. *Struct. Health Monit.* **2025**. *OnlineFirst*. [\[CrossRef\]](#)
22. Xia, H.; Meng, T.; Zuo, Z.; Ma, W. Fault semantic knowledge transfer learning: Cross-domain compound fault diagnosis method under limited single fault samples. *Reliab. Eng. Syst. Saf.* **2025**, *260*, 111050. [\[CrossRef\]](#)
23. Yi, W.; Jung, W.; Nam, H.; Jang, K.; Park, Y.-H. Multi-output classification using a cross-talk architecture for compound fault diagnosis of motors in partially labeled condition. *Mech. Syst. Signal Process.* **2026**, *244*, 113786. [\[CrossRef\]](#)
24. Zhu, R.; Song, E.; Ke, Y. Dynamic weighted semantic based two-stage training strategy for rotating machinery generalized zero-shot compound fault diagnosis. *Measurement* **2026**, *260*, 119802. [\[CrossRef\]](#)
25. Shen, F.; Lu, X.; Wen, H.; Zhou, L.; Tang, R. Zero-shot compound fault diagnosis with semantic graph embedding and multi-stage fusion. *Neurocomputing* **2026**, *668*, 132389. [\[CrossRef\]](#)
26. Randall, R.B.; Antoni, J. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, *25*, 485–520. [\[CrossRef\]](#)
27. Xia, P.; Huang, Y.; Qin, C.; Liu, C. Multi-modal cross-domain mixed fusion model with dual disentanglement for fault diagnosis under unseen working conditions. *Mech. Syst. Signal Process.* **2026**, *258*, 114693. [\[CrossRef\]](#)
28. Zhou, Q.; Chai, B.; Tang, C.; Guo, Y.; Wang, K.; Wu, W.; Cao, B.; Ye, Y. Enhancing multimodal fault diagnosis in mechanical systems via mixture of experts. *Complex Intell. Syst.* **2025**, *11*, 425. [\[CrossRef\]](#)
29. Faiz, J.; Ojaghi, M. Different indexes for eccentricity faults diagnosis in three-phase squirrel-cage induction motors: A review. *Mechatronics* **2009**, *19*, 2–13. [\[CrossRef\]](#)
30. Randall, R.B. *Vibration-Based Condition Monitoring: Industrial, Automotive and Aerospace Applications*, 2nd ed.; Wiley: Chichester, UK, 2021.
31. Coats, M.D.; Randall, R.B. Single and multi-stage phase demodulation based order-tracking. *Mech. Syst. Signal Process.* **2014**, *44*, 86–117. [\[CrossRef\]](#)
32. Zhao, M.; Lin, J.; Wang, X.; Lei, Y.; Cao, J. A tacho-less order tracking technique for large speed variations. *Mech. Syst. Signal Process.* **2013**, *40*, 76–90. [\[CrossRef\]](#)
33. Groover, C.L.; Trethewey, M.W.; Maynard, K.P.; Lebold, M.S. Removal of order domain content in rotating equipment signals by double resampling. *Mech. Syst. Signal Process.* **2005**, *19*, 483–500. [\[CrossRef\]](#)
34. Antoni, J. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mech. Syst. Signal Process.* **2006**, *20*, 282–307. [\[CrossRef\]](#)
35. Antoni, J.; Randall, R.B. The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mech. Syst. Signal Process.* **2006**, *20*, 308–331. [\[CrossRef\]](#)
36. Antoni, J. Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions. *J. Sound Vib.* **2007**, *304*, 497–529. [\[CrossRef\]](#)
37. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64–65*, 100–131. [\[CrossRef\]](#)
38. Sui, Z.; Wang, P.; Malik, H.; Akter, K. Zero-shot-motivated intelligent fault diagnosis: A survey of methods, applications, and future directions. *IET Gener. Transm. Distrib.* **2026**, *20*, e70361. [\[CrossRef\]](#)
39. Qin, N.; Yin, Y.; Huang, D.; You, Y.; Hou, R. Generalized zero-shot learning for fault diagnosis in high-speed train bogies based on enhanced diffusion generative models. *IEEE Trans. Reliab.* **2025**, *74*, 2867–2879. [\[CrossRef\]](#)
40. Yang, Q.; Lin, Y.; Ruan, J. A zero-shot fault attribute transfer learning method for compound fault diagnosis of power circuit breakers. *Meas. Sci. Technol.* **2024**, *35*, 056111. [\[CrossRef\]](#)
41. Tsoumakas, G.; Katakis, I. Multi-label classification: An overview. *Int. J. Data Warehous. Min.* **2007**, *3*, 1–13. [\[CrossRef\]](#)
42. Zhang, M.-L.; Zhou, Z.-H. A review on multi-label learning algorithms. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1819–1837. [\[CrossRef\]](#)
43. Jacobs, R.A.; Jordan, M.I.; Nowlan, S.J.; Hinton, G.E. Adaptive mixtures of local experts. *Neural Comput.* **1991**, *3*, 79–87. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Guo, L.; Lei, Y.; Xing, S.; Yan, T.; Li, N. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Trans. Ind. Electron.* **2019**, *66*, 7316–7325. [\[CrossRef\]](#)
45. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; Lempitsky, V. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* **2016**, *17*, 1–35.
46. Vlachou, V.I.; Kudelina, K.; Efsthathiou, D.E.; Vologiannidis, S.D.; Baraskova, T.; Shirokova, V.; Karakatsanis, T.S. Correlation-driven multisensory fusion for intelligent fault analysis in induction motors. *Machines* **2026**, *14*, 606. [\[CrossRef\]](#)
47. Guo, C.; Pleiss, G.; Sun, Y.; Weinberger, K.Q. On calibration of modern neural networks. In Proceedings of the 34th International Conference on Machine Learning (ICML), Sydney, Australia, 6–11 August 2017; PMLR 70, pp. 1321–1330.

48. Chen, S.; Liu, Z.; Li, C.; Zou, D.; He, X.; Zhou, D. Multi-mode fault diagnosis datasets of three-phase asynchronous motor under variable working conditions. *Data Brief* **2026**, *65*, 112583. [[CrossRef](#)] [[PubMed](#)]
49. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. In Proceedings of the 7th International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019.
50. Louzada, A.O.; Souza, W.A.; Vitor, A.L.O.; Castoldi, M.F.; Goedel, A. Detection of stator faults in three-phase induction motors using stray flux and machine learning. *Energies* **2025**, *18*, 1516. [[CrossRef](#)]
51. Yan, R.; Ren, J.; Wen, J.; Guo, C.; Zhao, Z.; Chen, X. Large models for machinery fault diagnosis: Current advances and future directions. *Chin. J. Mech. Eng.* **2026**, *39*, 100277. [[CrossRef](#)]
52. Yoo, J.; Park, J.; Kim, T.; Ha, J.M.; Youn, B.D. Weighted multi-order Viterbi algorithm (WMOVA): Instantaneous angular speed estimation under harsh conditions. *Mech. Syst. Signal Process.* **2024**, *211*, 111187. [[CrossRef](#)]
53. Zhang, D.; Idrus, Z.; Hamzah, R. An automatic ridge estimation method for tachless order tracking of wind turbine gearbox under varying-speed operation. *Meas. Sci. Technol.* **2025**, *36*, 086112. [[CrossRef](#)]

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