*mathematics*

MDPI

*Article*

# An Uncertainty-Adaptive Robust Graph-Based Belief Rule Base for Interpretable Four-State Fault Diagnosis of Industrial Robots

**Minyang Wang [1,2], Bangcheng Zhang [1,3,*] and Zhi Gao [1]**

[1] School of Mechatronic Engineering, Changchun University of Technology, Changchun 130012, China; 1202201007@stu.ccut.edu.cn (M.W.); gaozhi@ccut.edu.cn (Z.G.)
[2] Aviation Sub-College, Jilin Communications Polytechnic, Changchun 130015, China
[3] School of Mechanical and Electrical Engineering, Changchun Institute of Technology, Changchun 130012, China
* Correspondence: zhangbangcheng@ccut.edu.cn

**Abstract**

Industrial robot severity diagnosis must combine evidence from coupled subsystems while remaining stable under fluctuating measurements. We develop an uncertainty-adaptive robust graph-based belief rule base (RG-BRB) that maps monitored attributes to transmission, servo-control, and motion-execution nodes. A row-stochastic graph propagates node evidence, while local fluctuation adjusts antecedent widths and attribute reliability before rule activation. Normalized evidential reasoning yields a belief distribution over normal, slight-degradation, severe-degradation, and failure-related states. An interpretability-constrained adaptive covariance search estimates graph and rule parameters under ordering and normalization constraints. The evaluation uses one record of 2000 observations, with a fixed 600-observation test partition for the primary comparison. Across 10 random-seed runs on that partition, RG-BRB achieved 91.55% mean Accuracy and 0.9115 mean Macro-F1. Condition-shift, noise, calibration, and ablation analyses on the same record show how subsystem-constrained propagation and reliability-aware matching contribute to robust four-state diagnosis.

**Keywords:** industrial robot; fault diagnosis; belief rule base; evidential reasoning; graph reasoning; uncertainty quantification; interpretable artificial intelligence

**MSC:** 62P30; 68T37; 90B25; 94D05

Check for updates

Academic Editor: Yiu Yin Raymond Lee

Received: 4 October 2026
Revised: 4 November 2026
Accepted: 6 November 2026
Published: 8 November 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) license.

## 1. Introduction

Industrial robots are coupled electromechanical systems in which transmission wear, servo instability, and motion anomalies can propagate across subsystem boundaries. A diagnostic system should distinguish normal operation from successive degradation stages and retain the evidence behind each decision. These requirements are central to condition-based maintenance and model-based fault diagnosis [1–3]. They are difficult under variable loads and non-stationary noise because an abrupt observation may indicate either a transient disturbance or a genuine state transition [4,5].

Machine-learning diagnosis has progressed from engineered descriptors to convolutional, recurrent, capsule, attention-based, and graph-based representations [4,6–10]. These models capture nonlinear patterns in vibration, current, attitude, and process measurements. Industrial deployment also calls for stable decisions across operating conditions and an intelligible link between measurements and subsystem states [11–15]. A belief

distribution adds a useful measure of decision confidence, especially when its calibration can be examined alongside classification accuracy [16,17].

Belief rule bases (BRBs) offer a knowledge-guided route. RIMER represents each rule consequent as a belief distribution and combines activated rules by evidential reasoning (ER) [18]. Evidence aggregation supplies the mathematical foundation [19,20]; constrained training learns rule and attribute weights while preserving their semantics [21]. Multilayer, hierarchical, and deep BRBs extend this framework to complex systems [22–24]. Electromechanical diagnosis has also incorporated textual maintenance knowledge and multisource uncertainty [25,26]. RG-BRB builds on these advances by coupling fluctuation-sensitive rule matching with explicit subsystem evidence flow.

Graph models address structural dependence by representing components or variables as nodes connected by functional or causal relations. They have improved diagnosis under coupled industrial processes and changing conditions [27,28]. For maintenance use, the graph should expose how subsystem evidence enters rule activation and the final state belief [29,30].

This paper develops an uncertainty-adaptive robust graph-based belief rule base (RG-BRB) for direct diagnosis of four ordered states: normal, slight-degradation, severe-degradation, and failure-related. The contributions are as follows:

1. A graph-constrained BRB architecture is derived in which a row-stochastic functional propagation operator maps monitoring evidence to subsystem evidence before rule activation. The construction keeps propagated node evidence within the range of its inputs and prevents amplification in the infinity norm.
2. A reliability-aware matching mechanism couples local fluctuation, adaptive antecedent widths, rule matching, and ER fusion. The ER normalization is derived from rule-level probability masses, and the resulting diagnostic output remains on the probability simplex under the stated conditions.
3. A diagnosis-oriented adaptive covariance search combines class-balanced, ordinal, and knowledge-preservation terms under explicit feasibility constraints. The resulting belief outputs support direct four-state decisions, class-sensitive evaluation, calibration, and selective review.

Section 2 positions the method among data-driven, BRB, and graph-based diagnosis. Section 3 defines the four-state task, Section 4 derives RG-BRB, and Section 5 presents the evaluation. Section 6 interprets the findings, followed by the conclusions in Section 7.

## 2. Related Work

### *2.1. Data-Driven Industrial Fault Diagnosis*

Deep learning has shifted fault diagnosis toward learned multiscale representations. Convolutional and capsule networks capture impacts and local structures, while recurrent or convolutional-recurrent models encode temporal dependence [4,7,31]. Harmonic-drive and industrial robot studies demonstrate the value of task-specific signal transforms and multiscale feature learning [5,8]; sparse self-attention further emphasizes informative components of long monitoring sequences [9]. Robot-specific studies in Mathematics have evaluated transfer learning under varying loads, speeds, and motion patterns [32], while a focused review identifies data scarcity, computational burden, and the need for transparent prognostics and health management in industrial robots [33]. The same journal has also reported graph synthesis for class-imbalanced mechanical diagnosis [10] and explainable Bayesian learning for uncertainty-aware industrial abnormality detection [17].

Mechanism–data fusion, transfer learning, domain-feature decoupling, and class-incremental learning broaden diagnosis across small samples and changing operating

conditions [12–15]. Reviews also emphasize robustness and trustworthy interpretation in machinery diagnosis [3,6,11]. RG-BRB addresses these needs through a visible subsystem-to-rule evidence path and reliability-aware matching.

### *2.2. Belief Rule Bases Under Uncertain Evidence*

A BRB assigns a belief distribution to each rule consequent and combines the activated rules without forcing an early crisp label. RIMER provides the canonical inference structure [18], and ER supplies an analytical aggregation mechanism for multiple uncertain attributes [20]. Constrained optimization then calibrates a BRB while preserving non-negativity and normalization [21].

Recent work has improved representational capacity through multilayer, hierarchical, and deep structures [22–24]. The parallel–serial BRB for industrial robots uses feature selection and grouping to reduce rule complexity in health-state assessment [34]. RG-BRB instead organizes monitoring evidence through an explicit subsystem graph and uses fluctuation-dependent matching for direct four-state diagnosis. Other BRB models address trustworthy vehicle diagnosis and the conversion of qualitative maintenance reports into electromechanical rules [25,26]. The same numerical distance can represent strong evidence under a stable baseline but weaker evidence during a volatile interval. RG-BRB therefore lets reliability alter each attribute's contribution to rule activation.

### *2.3. Graph Structure and Explainable Diagnosis*

Graph representations encode dependencies among components or process variables, and graph neural networks provide a general mechanism for message passing [27]. Causal-intervention graph networks use structural information in complex industrial diagnosis [28]. For rule-based diagnosis, the graph can also make the pathway from measurement evidence to subsystem context, activated rules, and final state belief visible [29,30].

### *2.4. Research Gap*

The present task combines three requirements: measurement reliability changes with local fluctuation; subsystem interactions shape the meaning of each attribute; and four diagnostic stages retain an ordinal relationship. RG-BRB connects these requirements through functional graph propagation, reliability-aware rule activation, normalized ER fusion, and an objective tailored to direct four-state diagnosis.

## 3. Problem Formulation

### *3.1. Uncertain Multistate Observations*

At sampling instant $t$, the monitored feature vector is modeled as

$$\mathbf{x}_t = \Phi(y_t, \mathbf{o}_t) + \boldsymbol{\xi}_t + \boldsymbol{\varepsilon}_t, \qquad y_t \in \mathcal{Y} = \{D_1, D_2, D_3, D_4\}. \tag{1}$$

Here, $\mathbf{x}_t \in \mathbb{R}^M$ contains $M$ monitoring features; $y_t$ is the true state; $\mathbf{o}_t \in \mathbb{R}^C$ contains $C$ operating-condition variables; $\Phi(\cdot)$ is the ideal state–condition response; $\boldsymbol{\xi}_t$ denotes structured operating uncertainty; $\boldsymbol{\varepsilon}_t$ is random measurement noise; and $\mathcal{Y}$ is the diagnostic label set. States $D_1$–$D_4$ denote normal, slight-degradation, severe-degradation, and failure-related operation. The diagnostic target is the current severity stage.

### *3.2. Functional Evidence Graph*

Subsystem coupling is represented by a weighted directed graph

$$\mathcal{G} = (\mathcal{U}, \mathcal{E}, \mathbf{A}), \qquad \mathbf{A} = [a_{vu}] \in \mathbb{R}_+^{K \times K}. \tag{2}$$

Here, $\mathcal{V} = \{V_1, \ldots, V_K\}$ is the set of $K$ functional nodes; $\mathcal{E}$ is the admissible directed-edge set; and $a_{vu}$ measures how evidence at source node $V_u$ contributes to receiving node $V_v$. An entry is fixed to zero when $(V_u, V_v) \notin \mathcal{E}$, preventing unsupported evidence transfer.

### *3.3. Diagnostic Output*

The required model maps the observation and graph to a state-belief vector:

$$\hat{\boldsymbol{\beta}}_t = \mathcal{D}_{\boldsymbol{\Theta}}(\mathbf{x}_t, \mathcal{G}), \qquad \hat{\boldsymbol{\beta}}_t \in \Delta_4 = \left\{\mathbf{b} \in \mathbb{R}_+^4 : \mathbf{1}_4^{\mathsf{T}}\mathbf{b} = 1\right\}. \tag{3}$$

Here, $\mathcal{D}_{\boldsymbol{\Theta}}$ is the RG-BRB diagnostic mapping with trainable parameter vector $\boldsymbol{\Theta}$; $\hat{\boldsymbol{\beta}}_t = [\hat{\beta}_{1,t}, \ldots, \hat{\beta}_{4,t}]^{\mathsf{T}}$; $\Delta_4$ is the probability simplex over the four diagnostic states; $\mathbf{1}_4$ is a vector of four ones; and $\hat{\beta}_{n,t}$ is the support assigned to $D_n$.

The direct decision and residual uncertainty are

$$\hat{y}_t = D_{\arg\max_{1\le n\le 4}\hat{\beta}_{n,t}}, \qquad q_t = 1 - \max_{1\le n\le 4}\hat{\beta}_{n,t}. \tag{4}$$

Here, $\hat{y}_t$ is the diagnosed state and $q_t \in [0, 3/4]$ is residual decision uncertainty. Ordered severity is retained through

$$\hat{s}_t = \boldsymbol{\nu}^{\mathsf{T}}\hat{\boldsymbol{\beta}}_t, \qquad \boldsymbol{\nu} = [0, 1/3, 2/3, 1]^{\mathsf{T}}, \tag{5}$$

where $\hat{s}_t$ is expected severity and $\boldsymbol{\nu}$ assigns increasing utility to the four ordered states. This scalar regularizes ordinal distance but does not replace the class decision in Equation (4).

## 4. Proposed RG-BRB Method

### *4.1. Feature-to-Node Mapping and Graph Propagation*

Let $x_m^{\min,\mathrm{tr}}$ and $x_m^{\max,\mathrm{tr}}$ be the minimum and maximum of feature $m$ in the training set. The input to RG-BRB is

$$\tilde{x}_{m,t} = \mathrm{clip}_{[0,1]}\left(\frac{x_{m,t} - x_m^{\min,\mathrm{tr}}}{\max\{x_m^{\max,\mathrm{tr}} - x_m^{\min,\mathrm{tr}}, \epsilon_x\}}\right), \qquad m = 1, \ldots, M, \tag{6}$$

where $\mathrm{clip}_{[0,1]}(\cdot)$ truncates its argument to the unit interval and $\epsilon_x > 0$ supplies a positive denominator only when a training feature has zero range. Thus, $\tilde{\mathbf{x}}_t = [\tilde{x}_{1,t}, \ldots, \tilde{x}_{M,t}]^{\mathsf{T}} \in [0,1]^M$, and validation and test observations cannot alter the normalization constants. Initial node evidence is

$$\mathbf{e}_t = \mathbf{H}\tilde{\mathbf{x}}_t, \qquad h_{km} \ge 0, \qquad \sum_{m=1}^{M} h_{km} = 1. \tag{7}$$

Here, $\mathbf{e}_t \in [0,1]^K$ is local node evidence; $\mathbf{H} = [h_{km}] \in \mathbb{R}_+^{K\times M}$ is the physically specified feature-to-node matrix; and $h_{km}$ is the contribution of feature $m$ to node $k$. Row normalization makes each node an interpretable convex aggregation of its assigned features.

Define the receiving-node degree and normalized adjacency entries as

$$d_v^{\mathrm{in}} = \sum_{u=1}^{K} a_{vu}, \qquad \bar{a}_{vu} = \begin{cases} a_{vu}/d_v^{\mathrm{in}}, & d_v^{\mathrm{in}} > 0, \\ 1, & d_v^{\mathrm{in}} = 0 \text{ and } u = v, \\ 0, & d_v^{\mathrm{in}} = 0 \text{ and } u \ne v, \end{cases} \qquad \overline{\mathbf{A}} = [\bar{a}_{vu}]. \tag{8}$$

Here, $d_v^{\mathrm{in}}$ is the total incoming evidence weight of node $v$; $\bar{a}_{vu}$ is its normalized counterpart; and $\overline{\mathbf{A}}$ is row-stochastic. The second and third cases give an isolated node a unit self-loop without requiring an inverse of a zero degree. A one-step propagation is then

$$\mathbf{z}_t = \mathbf{P}\mathbf{e}_t, \qquad \mathbf{P} = \rho\mathbf{I}_K + (1-\rho)\overline{\mathbf{A}}, \qquad 0 \le \rho \le 1. \tag{9}$$

Here, $\mathbf{z}_t$ is propagated node evidence; $\mathbf{P}$ is the propagation operator; $\mathbf{I}_K$ is the $K \times K$ identity matrix; and $\rho$ retains local evidence. The remaining weight incorporates only graph-admissible neighbors.

The rule layer receives an explicit convex assignment of propagated node evidence and retained normalized features:

$$\mathbf{v}_t = \mathbf{B}\begin{bmatrix}\mathbf{z}_t \\ \tilde{\mathbf{x}}_t\end{bmatrix}, \qquad b_{mj} \geq 0, \qquad \sum_{j=1}^{K+M} b_{mj} = 1. \tag{10}$$

Here, $\mathbf{v}_t \in [0,1]^M$ is the BRB antecedent vector; $\mathbf{B} = [b_{mj}] \in \mathbb{R}_+^{M\times(K+M)}$ is a fixed, documented engineering assignment; and $b_{mj}$ links rule attribute $m$ to propagated or directly retained evidence component $j$. A one-hot row selects one source, whereas a nontrivial row represents an interpretable convex aggregation. The matrices $\mathbf{H}$ and $\mathbf{B}$ are fixed before model fitting; the normalized graph weights are learned only on edges admitted by $\mathcal{E}$.

The graph normalization uses nonnegative weights in every non-isolated row, assigns a unit self-loop to each isolated node, and assumes finite normalized monitoring features. Under these conditions, every row of $\mathbf{P}$ sums to one, so $\min_k e_{k,t} \leq z_{v,t} \leq \max_k e_{k,t}$ for every receiving node $v$, and $\|\mathbf{P}\|_\infty = 1$. This boundedness follows because both $\mathbf{I}_K$ and $\overline{\mathbf{A}}$ are row-stochastic and their convex combination is also row-stochastic. Each $z_{v,t}$ is therefore a convex combination of the entries of $\mathbf{e}_t$, while the maximum absolute row sum of the nonnegative propagation matrix remains one.

Figure 1 summarizes the evidence pathway. Graph propagation is performed before rule matching, so the functional structure cannot be bypassed by the decision layer.

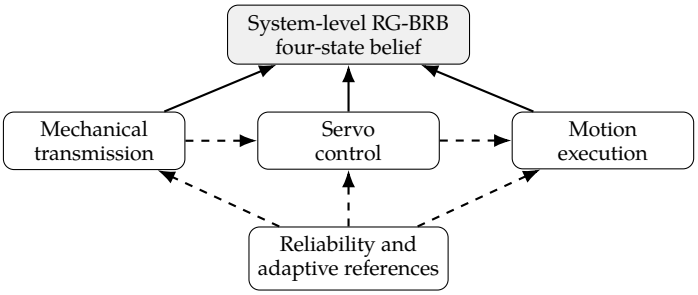



**Figure 1.** Functional organization of monitoring evidence before four-state RG-BRB inference. Solid arrows indicate aggregation into the system-level belief distribution; dashed arrows indicate reliability control and admissible cross-subsystem evidence flow.

### *4.2. Local Fluctuation and Adaptive Reference Intervals*

For a maximum causal window length $W \geq 2$, define the available history and its size by $\mathcal{W}_t = \{\max(1, t-W+1), \ldots, t\}$ and $W_t = |\mathcal{W}_t|$. When $W_t \geq 2$, the local mean and standard deviation of antecedent attribute $m$ are

$$\overline{v}_{m,t} = \frac{1}{W_t}\sum_{i\in\mathcal{W}_t} v_{m,i}, \qquad \sigma_{m,t} = \left[\frac{1}{W_t - 1}\sum_{i\in\mathcal{W}_t}(v_{m,i} - \overline{v}_{m,t})^2\right]^{1/2}. \tag{11}$$

Here, $i$ is a temporal index within the current-and-past window; $W$ is fixed on validation data; $\overline{v}_{m,t}$ is the local mean; and $\sigma_{m,t}$ is local fluctuation. At the start of a monitored sequence, the causal buffer is initialized with the final development observations; if fewer than two valid observations are available, diagnosis is withheld rather than assigning a zero variance. Let $\sigma_{m,0}$ be the corresponding normal-state baseline estimated from training samples. Excess fluctuation and reliability are

$$\gamma_{m,t} = \max\left\{0, \frac{\sigma_{m,t}}{\sigma_{m,0} + \epsilon_\sigma} - 1\right\}, \qquad r_{m,t} = \exp\left(-\frac{\gamma_{m,t}^2}{2\tau_m^2}\right). \tag{12}$$

Here, $\gamma_{m,t}$ is excess fluctuation relative to baseline; $\epsilon_\sigma > 0$ is the variance-ratio stabilizer; $\tau_m > 0$ controls reliability decay; and $r_{m,t} \in (0,1]$ is feature reliability. Evidence at or below the normal fluctuation level retains unit reliability.

For antecedent reference level $\ell$, the adaptive half-width is bounded as

$$w_{m\ell t} = \min\{w_{m\ell}^{\max},\ \kappa_{m\ell}(1 + v_m\gamma_{m,t})\}, \qquad \mathcal{J}_{m\ell t} = [c_{m\ell} - w_{m\ell t},\ c_{m\ell} + w_{m\ell t}]. \tag{13}$$

Here, $c_{m\ell}$ is the ordered reference center; $\kappa_{m\ell} > 0$ is its nominal half-width; $v_m \geq 0$ controls expansion; $w_{m\ell}^{\max} \geq \kappa_{m\ell}$ is a fixed upper bound; $w_{m\ell t}$ is the current half-width; and $\mathcal{J}_{m\ell t}$ is the resulting interval. Expansion reduces abrupt rule switching during short volatile periods, while the upper bound prevents all antecedents from becoming indistinguishable under extreme noise.

### *4.3. Reliability-Aware Rule Activation*

If rule $k$ selects reference level $\ell_{mk}$ for attribute $m$, its triangular matching degree is

$$\alpha_{mkt} = \max\left\{0, 1 - \frac{|v_{m,t} - c_{m,\ell_{mk}}|}{w_{m,\ell_{mk},t}}\right\}. \tag{14}$$

Here, $\ell_{mk}$ is the antecedent-level index and $\alpha_{mkt} \in [0,1]$ is the resulting match. Let $\theta_k \in [0,1]$ be rule importance, $\delta_m \in [0,1]$ be attribute importance, and $\bar{\delta}_m = \delta_m / \max_j \delta_j$ with $\max_j \delta_j > 0$. Reliability modifies the attribute exponent through $\delta_{m,t}^{\text{eff}} = r_{m,t}\bar{\delta}_m$. To prevent an exactly zero match from permanently vetoing a rule when the corresponding attribute is unreliable, introduce a fixed numerical floor $\epsilon_\alpha \in (0,1)$ and define the unnormalized support and activation by

$$\begin{aligned} \widetilde{\omega}_{kt} &= \theta_k \prod_{m=1}^{M} \varphi(\alpha_{mkt}, \delta_{m,t}^{\text{eff}}), \\ \varphi(a,e) &= \exp(e\log[\epsilon_\alpha + (1-\epsilon_\alpha)a]), \qquad S_t = \sum_{k'=1}^{L} \widetilde{\omega}_{k't}, \\ \omega_{kt} &= \begin{cases} \widetilde{\omega}_{kt}/S_t, & S_t > \epsilon_S, \\ 0, & S_t \leq \epsilon_S, \end{cases} \end{aligned} \tag{15}$$

Here, $L$ is the number of rules and $k'$ indexes the normalization sum. The generic arguments of $\varphi$ are $a, e \in [0,1]$. The activation-coverage threshold $\epsilon_S > 0$ is fixed using development data. When $S_t > \epsilon_S$, the activations are nonnegative and satisfy $\sum_k \omega_{kt} = 1$. The logarithm receives a value in $[\epsilon_\alpha, 1]$, so $\varphi$ remains defined at $a = 0$ and $e = 0$.

As $r_{m,t}$ decreases, the $m$th exponent approaches zero and that attribute contributes a factor approaching one to each rule. Applying one common reliability multiplier to all rules would cancel during activation normalization. When $S_t \leq \epsilon_S$, activations are set to zero and the prior fallback below is used.

Rule $k$ has the form

$$R_k: \quad \text{IF} \bigwedge_{m=1}^{M} (v_{m,t} \text{ matches } \mathcal{J}_{m,\ell_{mk},t}), \quad \text{THEN } \{(D_n, \beta_{nk})\}_{n=1}^{4}, \quad \beta_{nk} \geq 0, \quad \sum_{n=1}^{4} \beta_{nk} = 1. \tag{16}$$

Here, $R_k$ is the belief rule; $\bigwedge$ denotes conjunction across its $M$ antecedents; and $\beta_{nk}$ is its consequent belief for state $D_n$. Complete consequent distributions are used so every activated rule has a direct probabilistic interpretation.

### *4.4. Derivation of Evidential-Reasoning Fusion*

For $S_t > \epsilon_S$, activation discounts each complete rule consequent into state-specific and unassigned masses:

$$m_{nkt} = \omega_{kt}\beta_{nk}, \qquad m_{Ukt} = 1 - \omega_{kt}. \tag{17}$$

Here, $m_{nkt}$ is committed support for $D_n$; $U$ denotes the undifferentiated full state frame; and $m_{Ukt}$ is the mass left unassigned because rule $k$ is not fully activated. Equation (16) gives $\sum_n m_{nkt} + m_{Ukt} = 1$. Under ER combination, a product contributes to $D_n$ only if every selected factor either supports $D_n$ or remains unassigned. This gives

$$G_{nt} = \prod_{k=1}^{L} (m_{nkt} + m_{Ukt}), \qquad G_{Ut} = \prod_{k=1}^{L} m_{Ukt}. \tag{18}$$

Here, $G_{nt}$ collects combinations compatible with $D_n$, and $G_{Ut}$ is the wholly unassigned product included in every $G_{nt}$. Subtracting this duplicated product leaves the non-conflicting state masses $G_{nt} - G_{Ut}$. Define the standard ER normalizer and their total by

$$\begin{aligned} Z_t^{\mathrm{ER}} &= \sum_{n=1}^{4} (G_{nt} - G_{Ut}), \qquad \mu_t = (Z_t^{\mathrm{ER}} + G_{Ut})^{-1}, \\ \hat{\beta}_{n,t} &= \frac{\mu_t (G_{nt} - G_{Ut})}{1 - \mu_t G_{Ut}} = \frac{G_{nt} - G_{Ut}}{Z_t^{\mathrm{ER}}}. \end{aligned} \tag{19}$$

Here, $Z_t^{\mathrm{ER}}$ is the total committed mass and $\hat{\beta}_{n,t}$ is the fused diagnostic belief. For four states, $Z_t^{\mathrm{ER}} + G_{Ut} = \sum_n G_{nt} - 3G_{Ut}$; its reciprocal $\mu_t$ is the standard ER conflict normalizer. Equation (19) specializes analytical ER to four states with complete rule consequents [18,20]. Its final form gives the same belief with a simpler numerical implementation.

The committed-mass sum $Z_t^{\mathrm{ER}}$ is computed first. Let $\epsilon_{\mathrm{ER}} > 0$ be a fixed conflict threshold. If $S_t \le \epsilon_S$ or $Z_t^{\mathrm{ER}} \le \epsilon_{\mathrm{ER}}$, the divisions in Equation (19) are bypassed; instead,

$$\hat{\boldsymbol{\beta}}_t = \boldsymbol{\pi}^{(0)}, \qquad \boldsymbol{\pi}^{(0)} \in \Delta_4, \tag{20}$$

where $\boldsymbol{\pi}^{(0)}$ is a fixed diagnostic prior estimated from the development data or elicited before test evaluation. The thresholds $\epsilon_S$ and $\epsilon_{\mathrm{ER}}$ are fixed before test evaluation. This branch makes insufficient rule coverage or complete evidence conflict explicit and prevents an arbitrary average of unrelated rule consequents.

The two-branch inference rule also preserves a valid belief distribution. If the rule consequents satisfy Equation (16), the activations satisfy Equation (15), and the prior satisfies Equation (20), then $\hat{\boldsymbol{\beta}}_t \in \Delta_4$. When $S_t > \epsilon_S$ and $Z_t^{\mathrm{ER}} > \epsilon_{\mathrm{ER}}$, the discounted masses in Equation (17) are nonnegative and sum to one for every rule. Hence $G_{nt} \ge G_{Ut}$, making every numerator in the stable form of Equation (19) nonnegative. Their sum is $Z_t^{\mathrm{ER}}$, so division by this positive quantity produces a nonnegative vector with unit sum. In the fallback branch, Equation (20) belongs to $\Delta_4$ by construction. The diagnostic vector therefore lies in $\Delta_4$ in either branch.

### *4.5. Constrained Diagnosis-Oriented Learning*

Let the training set contain $N_{\mathrm{tr}}$ observations, of which $N_n$ belong to state $D_n$. To reduce majority-state dominance, define the class weight $\varrho_n = N_{\mathrm{tr}}/(4N_n)$. The class-balanced cross-entropy is

$$\mathcal{J}_{\mathrm{cls}}(\boldsymbol{\Theta}) = -\frac{1}{N_{\mathrm{tr}}} \sum_{t=1}^{N_{\mathrm{tr}}} \sum_{n=1}^{4} \varrho_n \mathbb{I}(y_t = D_n) \log\big(\max\{\hat{\beta}_{n,t}, \epsilon_p\}\big). \tag{21}$$

Here, $t$ retains its sampling-index meaning over the training observations; $\mathbb{I}(\cdot)$ is the indicator function; $\epsilon_p \in (0,1)$ is a fixed probability-log floor; and $\mathcal{J}_{\mathrm{cls}}$ directly rewards support for the correct state. Define the severity utility function by $\nu(D_n) = (n-1)/3$. The ordered-state penalty is

$$\mathcal{J}_{\mathrm{ord}}(\boldsymbol{\Theta}) = \frac{1}{N_{\mathrm{tr}}} \sum_{t=1}^{N_{\mathrm{tr}}} (\hat{s}_t - \nu(y_t))^2, \tag{22}$$

where $\nu(y_t)$ is the utility assigned to the true class. This term penalizes a normal-to-failure error more heavily than confusion between adjacent states. Expert structure is retained by

$$\mathcal{J}_{\text{know}}(\boldsymbol{\Theta}) = (\boldsymbol{\Theta} - \boldsymbol{\Theta}^0)^{\top}\mathbf{R}_{\Theta}(\boldsymbol{\Theta} - \boldsymbol{\Theta}^0), \tag{23}$$

where $\boldsymbol{\Theta}^0$ is expert initialization and $\mathbf{R}_{\Theta} \succeq 0$ assigns different preservation strengths to graph, reference, rule, and consequent parameters. The complete problem is

$$\boldsymbol{\Theta}^{\star} = \arg\min_{\boldsymbol{\Theta}\in\mathcal{F}}\left(\mathcal{J}_{\text{cls}} + \lambda_{\text{o}}\mathcal{J}_{\text{ord}} + \lambda_{\text{k}}\mathcal{J}_{\text{know}}\right), \tag{24}$$

where $\lambda_{\text{o}}, \lambda_{\text{k}} \geq 0$ control the ordinal and knowledge terms. For receiving node $\mathsf{V}_v$, let $\mathcal{N}_v^- = \{u : (\mathsf{V}_u, \mathsf{V}_v) \in \mathcal{E}\}$ denote its admitted source-node indices. The feasible set is

$$\begin{aligned}\mathcal{F} = \{\boldsymbol{\Theta} : {} & \bar{a}_{vu} \geq 0,\ \ \overline{\mathbf{A}}\mathbf{1}_K = \mathbf{1}_K,\ \ \bar{a}_{vu} = 0 \text{ if } \mathcal{N}_v^- \neq \emptyset \text{ and } u \notin \mathcal{N}_v^-, \\ & \bar{a}_{vu} = \mathbb{I}(u = v) \text{ if } \mathcal{N}_v^- = \emptyset,\ \ 0 \leq \rho \leq 1, \\ & 0 \leq \theta_k, \delta_m \leq 1,\ \theta_{k_*} = 1,\ \delta_{m_*} = 1,\ \ \tau_{\min} \leq \tau_m \leq \tau_{\max},\ 0 \leq \upsilon_m \leq \upsilon_{\max}, \\ & \kappa_{\min} \leq \kappa_{m\ell} \leq w_{m\ell}^{\max},\ \ c_m^{\min} \leq c_{m1} \leq c_{mQ_m} \leq c_m^{\max}, \\ & c_{m,\ell+1} - c_{m\ell} \geq \Delta_{m,\min},\ \ \boldsymbol{\beta}_k \in \Delta_4\},\end{aligned} \tag{25}$$

Here, $\boldsymbol{\Theta}$ stacks the trainable blocks $\overline{\mathbf{A}}$, $\rho$, $\{\theta_k\}$, $\{\delta_m\}$, $\{\tau_m\}$, $\{\upsilon_m\}$, $\{\kappa_{m\ell}\}$, $\{c_{m\ell}\}$, and $\{\boldsymbol{\beta}_k\}$. The indicator $\mathbb{I}(u = v)$ fixes an isolated graph row to its unit self-loop; for a non-isolated row, mass can be assigned only to edges in $\mathcal{E}$. The center-spacing constraint applies for $\ell = 1, \ldots, Q_m - 1$; $Q_m$ is the number of reference levels for attribute $m$; $[c_m^{\min}, c_m^{\max}]$ is its admissible normalized range; $\Delta_{m,\min} > 0$ is the minimum center separation; $0 < \kappa_{\min} \leq w_{m\ell}^{\max}$ prevents degenerate or unbounded intervals; $0 < \tau_{\min} \leq \tau_{\max}$ bounds reliability decay; and $\upsilon_{\max}$ bounds interval expansion. The indices $k_*$ and $m_*$ identify expert-selected anchor rule and attribute weights; fixing them to one removes the common-scale ambiguity in normalized activation and makes $\max_m \delta_m$ well defined. The vector $\boldsymbol{\beta}_k$ is the consequent of rule $k$. The normalized adjacency is optimized directly on its permitted row simplices, which removes the scale non-identifiability of raw edge weights and avoids duplicating the self-retention already controlled by $\rho$. The fixed support $\mathcal{E}$ prevents the optimizer from inventing physically unsupported edges. All parameter blocks are closed and convex, so their Euclidean projections are well defined.

### *4.6. Interpretability-Constrained Adaptive Covariance Search*

The objective is nonlinear and contains simplex, order, and graph constraints. RG-BRB therefore uses an interpretability-constrained adaptive covariance search (ACSO-I) based on covariance matrix adaptation [35]. Each structured block is encoded in an unconstrained vector containing $n_{\Theta}$ parameters, after which the decoded candidate is projected blockwise onto Equation (25). At generation $g$, candidate $p$ is sampled and projected as

$$\mathbf{u}_p^{(g)} = \boldsymbol{\chi}^{(g)} + \eta^{(g)}(\boldsymbol{\Sigma}^{(g)})^{1/2}\boldsymbol{\zeta}_p^{(g)}, \quad \boldsymbol{\zeta}_p^{(g)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_{\Theta}}), \quad \boldsymbol{\Omega}_p^{(g)} = \Pi_{\mathcal{F}}(\mathbf{u}_p^{(g)}). \tag{26}$$

Here, $\boldsymbol{\chi}^{(g)}$ is the search center; $\eta^{(g)} > 0$ is global step size; $\boldsymbol{\Sigma}^{(g)} \succ 0$ is the search covariance; $\mathbf{I}_{n_{\Theta}}$ is the identity of order $n_{\Theta}$; $\boldsymbol{\zeta}_p^{(g)}$ is a standard Gaussian vector; $\Pi_{\mathcal{F}}$ is the blockwise Euclidean projection onto Equation (25); and $\boldsymbol{\Omega}_p^{(g)}$ is a feasible candidate. Projection consists of bounded clipping for scalar parameters, isotonic projection with minimum spacing for each ordered center vector, and simplex projection for every graph row and rule consequent.

After ranking the population by Equation (24), let $\boldsymbol{\Omega}_{i:\Lambda}^{(g)}$ be the $i$th best of $\Lambda$ candidates, and let $\psi_i > 0$ satisfy $\sum_{i=1}^{\Lambda_{\text{s}}} \psi_i = 1$ for the best $\Lambda_{\text{s}}$ candidates. Define $\Lambda_{\text{eff}} = (\sum_{i=1}^{\Lambda_{\text{s}}} \psi_i^2)^{-1}$ and $\mathbf{y}_i^{(g)} = (\boldsymbol{\Omega}_{i:\Lambda}^{(g)} - \boldsymbol{\chi}^{(g)})/\eta^{(g)}$. Here, $\psi_i$ is the recombination weight of ranked candidate $i$. The center and covariance path update are

$$\boldsymbol{\chi}^{(g+1)} = \sum_{i=1}^{\Lambda_{\text{s}}} \psi_i \boldsymbol{\Omega}_{i:\Lambda}^{(g)}, \qquad \mathbf{p}_c^{(g+1)} = (1 - c_c)\mathbf{p}_c^{(g)} + \sqrt{c_c(2 - c_c)\Lambda_{\text{eff}}} \sum_{i=1}^{\Lambda_{\text{s}}} \psi_i \mathbf{y}_i^{(g)}. \tag{27}$$

Here, $c_c \in (0,1)$ is the path learning rate and $\mathbf{p}_c$ accumulates successful directions. Covariance is updated by

$$\mathbf{\Sigma}^{(g+1)} = (1 - c_1 - c_s)\mathbf{\Sigma}^{(g)} + c_1 \mathbf{p}_c^{(g+1)} (\mathbf{p}_c^{(g+1)})^\top + c_s \sum_{i=1}^{\Lambda_s} \psi_i \mathbf{y}_i^{(g)} (\mathbf{y}_i^{(g)})^\top, \tag{28}$$

where $c_1$ and $c_s$ are the rank-one and selected-population learning rates satisfying $c_1 + c_s < 1$. Starting from $\mathbf{\Sigma}^{(0)} \succ 0$, the positive coefficient retained on the previous covariance preserves positive definiteness. For step-size adaptation, define

$$\begin{aligned} \mathbf{p}_\sigma^{(g+1)} &= (1 - c_\sigma)\mathbf{p}_\sigma^{(g)} + \sqrt{c_\sigma(2 - c_\sigma)\Lambda_{\text{eff}}} (\mathbf{\Sigma}^{(g)})^{-1/2} \sum_{i=1}^{\Lambda_s} \psi_i \mathbf{y}_i^{(g)}, \\ \eta^{(g+1)} &= \eta^{(g)} \exp\left[ \frac{c_\sigma}{d_\sigma} \left( \frac{\|\mathbf{p}_\sigma^{(g+1)}\|_2}{\mathbb{E}\|\mathcal{N}(\mathbf{0}, \mathbf{I}_{n_\Theta})\|_2} - 1 \right) \right]. \end{aligned} \tag{29}$$

Here, $\mathbf{p}_\sigma$ is the isotropic evolution path; $c_\sigma$ is its learning rate; $d_\sigma$ is damping; and $\mathbb{E}\|\mathcal{N}(\mathbf{0}, \mathbf{I}_{n_\Theta})\|_2$ is the expected Euclidean norm of an $n_\Theta$-dimensional standard Gaussian vector. Projection after sampling is the interpretability constraint: every evaluated candidate has ordered references, admissible graph weights, bounded rule and attribute weights, and normalized consequents. To monitor the effect of projection, the implementation records both $\|\mathbf{u}_p^{(g)} - \boldsymbol{\chi}^{(g)}\|_2$ and $\|\mathbf{\Omega}_p^{(g)} - \boldsymbol{\chi}^{(g)}\|_2$; persistent discrepancies prompt rescaling of the step size or parameter encoding.

### *4.7. Computational Complexity*

Let $S_H = \text{nnz}(\mathbf{H})$ and $S_B = \text{nnz}(\mathbf{B})$ be the numbers of stored nonzero assignment coefficients. For one observation, the two mappings cost $O(S_H + S_B)$, graph propagation costs $O(K + |\mathcal{E}|)$ including self-retention, matching and activation cost $O(LM)$, and four-state ER fusion costs $O(4L)$. Local means and variances are maintained by causal rolling sums in $O(M)$ time rather than recomputed over all $W$ samples. With four fixed states, define

$$\begin{aligned} C_{\text{inf}} &= S_H + S_B + K + |\mathcal{E}| + LM, \\ T_{\text{infer}} &= O(C_{\text{inf}}), \\ T_{\text{train}} &= O\Big( N_g [\Lambda \big( N_{\text{tr}} C_{\text{inf}} + n_\Theta \log n_\Theta \big) \\ &\qquad + n_\Theta^3 + (\Lambda + \Lambda_s) n_\Theta^2 + \Lambda \log \Lambda] \Big). \end{aligned} \tag{30}$$

Here, $C_{\text{inf}}$ collects the leading inference operations; $N_g$ is the number of optimizer generations; $\Lambda$ is population size; $n_\Theta^3$ accounts for a dense covariance factorization; $(\Lambda + \Lambda_s)n_\Theta^2$ covers dense candidate generation and rank updates; $n_\Theta \log n_\Theta$ is a worst-case bound for blockwise simplex and ordered-center projection; and $\Lambda \log \Lambda$ covers population ranking. The factorization can be reused within a generation. A diagonal covariance reduces dense covariance generation and adaptation to $O(\Lambda n_\Theta)$ time and $O(n_\Theta)$ memory, while objective evaluation, projection, and ranking costs remain.

## 5. Case Analysis and Experimental Protocol

### *5.1. Data, Labels, and Task Boundary*

The case study comprises 2000 ordered industrial robot observations, 10 candidate attributes, four retained normalized attributes, and four state labels. Class counts are 600 normal, 600 slight-degradation, 500 severe-degradation, and 300 failure-related observations. The retained attributes characterize load, servo response, and motion behavior and are assigned to the functional nodes in Figure 1. Their ordered trajectories are shown in Figure 2. All reported evaluations draw on this case record.

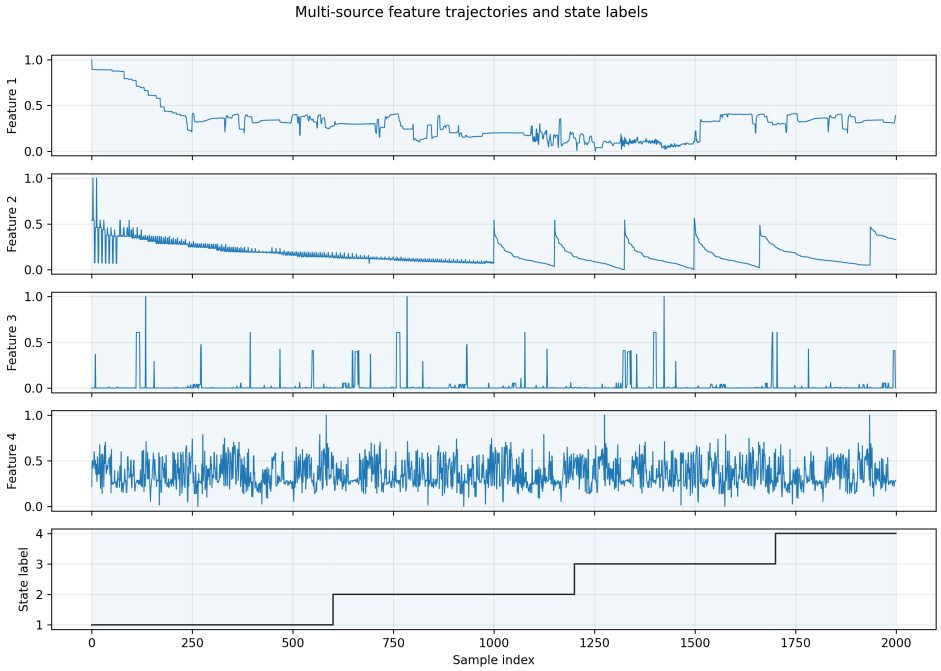



**Figure 2.** Ordered monitoring trajectories and four diagnostic-state stages in the 2000-observation robot record. Normalization and feature selection used in model fitting were estimated from the development data only.

The labels identify normal operation, two degradation stages, and failure-related operation. The analysis therefore centers on the severity stage associated with each observation. Table 1 summarizes the case data and diagnostic role.

**Table 1.** Case data and diagnostic role.

| Quantity | Diagnostic Role |
|---|---|
| 2000 ordered observations | Training, validation, and order-preserving contiguous testing |
| 10 candidate attributes | Training-only redundancy and relevance analysis |
| 4 retained attributes | Common input to RG-BRB and every comparison method |
| Four ordered labels | Normal, slight-degradation, severe-degradation, and failure-related |
| Class counts | 600, 600, 500, and 300 observations, respectively |

### *5.2. Order-Preserving Split and Baselines*

The evaluation preserves the ordering of observations within each severity stage. For every state, the first 70% forms the development block and the remaining 30% forms the test block. The fixed test set contains 180, 180, 150, and 90 observations from $D_1$–$D_4$, respectively, totaling 600. Each development block is further divided in order into 80% training and 20% validation observations. Normalization constants, feature screening, window length, baseline fluctuations, antecedent references, optimizer settings, and uncertainty thresholds are determined from the development data.

RG-BRB was compared with a conventional BRB, a long short-term memory network (LSTM), a one-dimensional convolutional neural network (CNN), and a radial-basis-function support vector machine (RBF-SVM). The graph-free formulation was evaluated as an RG-BRB ablation. All six configurations used the same four retained attributes and fixed partition. Ten random-seed runs used the same 600 test observations to measure variation in model fitting. Subsequent analyses evaluate different conditions and fitted configurations using the same 2000-observation record.

Table 2 summarizes the recorded candidate spaces and budgets; the selected configurations were evaluated on the held-out test blocks. The candidate spaces and search budgets differ by model family.

**Table 2.** Recorded tuning spaces and budgets for the principal methods.

| Method | Candidate Space | Recorded Budget |
|---|---|---|
| RG-BRB | Graph, matching and regularization parameters | 500 generations, population 20 |
| LSTM | Hidden units 32 or 64; learning rate 0.0005–0.01; dropout 0.1 or 0.2 | 2 × 4 × 2 configurations, 5 seeds |
| CNN | Filters 16 or 32; kernel 3 or 5; learning rate 0.0005–0.01 | 2 × 2 × 4 configurations, 5 seeds |
| Conventional BRB | Reference-width candidates | 10 configurations, 5 seeds |
| RBF-SVM | $C \in \{0.1, 1, 10, 100\}$; kernel scale or automatic | 8 configurations, 5 seeds |

### *5.3. Diagnostic and Calibration Metrics*

Let $\mathbf{\Gamma} = [\Gamma_{ij}] \in \mathbb{N}_0^{4\times4}$ be the confusion matrix, where $\Gamma_{ij}$ counts observations with true state $D_i$ and predicted state $D_j$. Accuracy is

$$\text{Accuracy} = \frac{\sum_{n=1}^{4} \Gamma_{nn}}{\sum_{i=1}^{4}\sum_{j=1}^{4} \Gamma_{ij}}. \tag{31}$$

Here, $i$ and $j$ index true and predicted classes. Class-wise precision, recall, and F1 are

$$\text{Prec}_n = \frac{\Gamma_{nn}}{\sum_{i=1}^{4} \Gamma_{in}}, \quad \text{Rec}_n = \frac{\Gamma_{nn}}{\sum_{j=1}^{4} \Gamma_{nj}}, \quad F1_n = \frac{2\text{Prec}_n\text{Rec}_n}{\text{Prec}_n + \text{Rec}_n}, \quad \text{Macro-F1} = \frac{1}{4}\sum_{n=1}^{4} F1_n. \tag{32}$$

Here, $\text{Prec}_n$ measures the reliability of predictions assigned to $D_n$; $\text{Rec}_n$ measures detection of $D_n$; and Macro-F1 weights every class equally. Together with Accuracy, these measures describe performance across the four ordered states [10,36].

If a precision or recall denominator is zero, the corresponding quantity and its $F1_n$ are defined as zero; the same convention applies to every method. Belief quality is assessed first using negative log-likelihood (NLL),

$$\text{NLL} = -\frac{1}{N_{\text{te}}}\sum_{t=1}^{N_{\text{te}}}\sum_{n=1}^{4} \mathbb{I}(y_t = D_n)\log\Big(\max\{\hat{\beta}_{n,t}, \epsilon_p\}\Big), \tag{33}$$

and using the multiclass Brier score

$$\text{Brier} = \frac{1}{N_{\text{te}}}\sum_{t=1}^{N_{\text{te}}}\sum_{n=1}^{4}\Big(\hat{\beta}_{n,t} - \mathbb{I}(y_t = D_n)\Big)^2, \tag{34}$$

where $N_{\text{te}}$ is the test-set size. Expected calibration error (ECE) is

$$\text{ECE} = \sum_{\iota=1}^{N_{\text{bin}}} \frac{|\mathcal{B}_\iota|}{N_{\text{te}}}|\text{acc}(\mathcal{B}_\iota) - \text{conf}(\mathcal{B}_\iota)|, \tag{35}$$

where $N_{\text{bin}}$ is the pre-specified number of confidence bins; $\iota$ is the bin index; $\mathcal{B}_\iota$ contains test observations whose maximum belief falls in that bin; and acc and conf are empirical correctness and mean maximum belief in the bin. Empty bins contribute zero. Calibration matters because an interpretable belief distribution should also correspond to observed correctness frequencies [16,17].

### *5.4. Experimental Records and Reproducibility*

The 10-run primary comparison uses a fixed 600-observation test partition. Confusion matrices and paired intervals provide a closer view of a selected fit. Condition-shift, Gaussian- and impulse-noise, calibration, selective-risk, and ablation analyses examine complementary diagnostic properties of the same source record.

Means and standard deviations describe variation across random-seed runs. The paired intervals characterize the selected fit, while the robustness curves show performance under controlled perturbations and operating conditions.

### *5.5. Overall Performance and Repeated-Run Stability*

Table 3 summarizes the 10-run results on the fixed test set. RG-BRB achieves a mean Accuracy of 91.55% and mean Macro-F1 of 0.9115, compared with 89.03% and 0.8874 for LSTM. The mean Accuracy difference is 2.52 percentage points. The Accuracy standard deviation (SD) is 0.52 percentage points for RG-BRB and 1.07 for LSTM. The graph-free variant reaches 89.58% mean Accuracy, 1.97 percentage points below RG-BRB.

**Table 3.** Performance over 10 runs on the same 600-observation test set. Values are mean ± sample SD. The graph-free variant is an ablation.

| Method | Accuracy (%, Mean ± SD) | Macro-F1 (Mean ± SD) |
|---|---|---|
| RG-BRB | 91.55 ± 0.52 | 0.9115 ± 0.0061 |
| LSTM | 89.03 ± 1.07 | 0.8874 ± 0.0112 |
| CNN | 86.62 ± 0.73 | 0.8629 ± 0.0073 |
| Conventional BRB | 86.05 ± 0.74 | 0.8563 ± 0.0072 |
| RBF-SVM | 82.82 ± 0.54 | 0.8244 ± 0.0061 |
| Without graph propagation | 89.58 ± 0.59 | 0.8921 ± 0.0058 |

Figure 3 displays the individual Accuracy values and their variation across random seeds. The fixed test partition is common to all configurations.

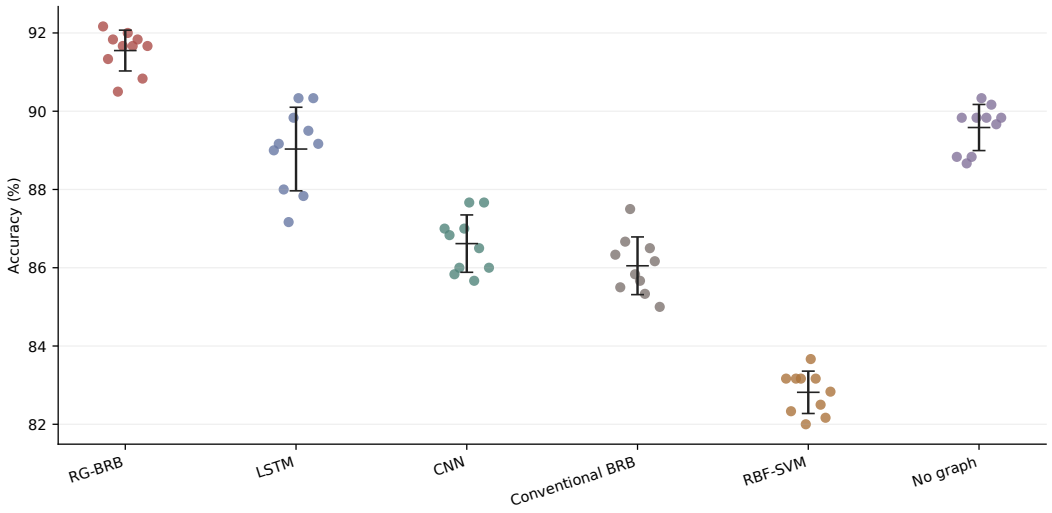



**Figure 3.** Accuracy across 10 runs on the fixed test set. Points denote individual runs and error bars denote mean ± sample SD.

Figure 4 shows the differences obtained by matching the run identifiers. RG-BRB exceeds LSTM by 0.50–3.83 percentage points and the graph-free variant by 1.17–3.33 points across these runs. These differences quantify run-matched performance on the fixed partition.

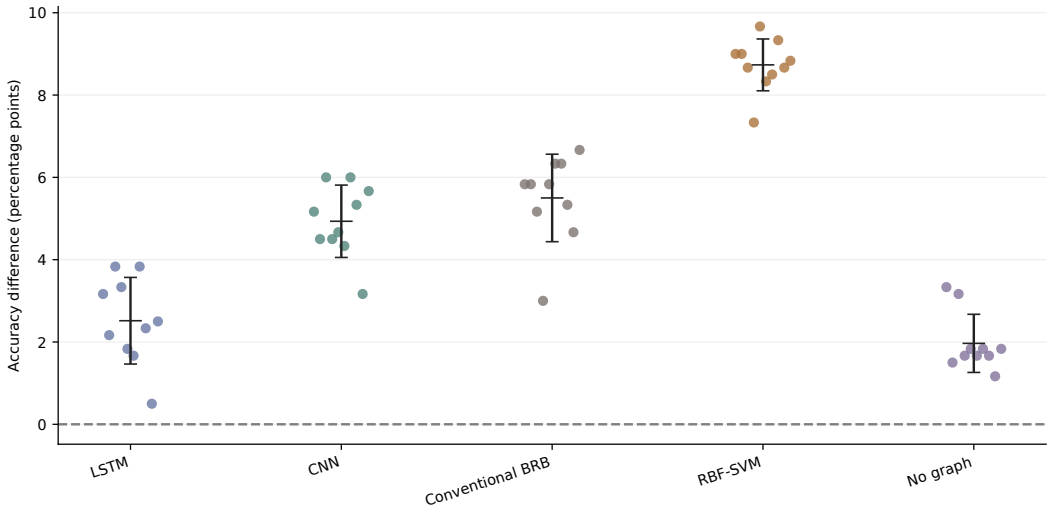



**Figure 4.** Run-matched Accuracy differences between RG-BRB and the other configurations on the fixed test set. Points denote individual differences and error bars denote mean ± sample SD.

Table 4 reports mean run-matched differences for the same 10 runs and intervals for run 01. Paired non-overlapping blocks are resampled within each state, preserving the state counts; 2000 resamples and a block length of 10 are used. The run-01 paired predictions also permit exact two-sided McNemar tests. The intervals describe the selected fit on this test partition, while the McNemar results test its paired classification outcomes.

**Table 4.** Paired comparisons on the fixed 600-observation test set. Mean differences are across the 10 runs in Table 3; block-resampling intervals and exact two-sided McNemar *p*-values apply to run 01. Accuracy differences are in percentage points.

| Comparison | 10-Run Mean | Run 01 | Run 01 95% Interval | McNemar $p$ |
|---|---|---|---|---|
| RG-BRB − LSTM | 2.52 | 3.17 | [1.17, 5.33] | 0.0026 |
| RG-BRB − no graph | 1.97 | 3.33 | [2.00, 5.00] | <0.001 |
| RG-BRB − CNN | 4.93 | 5.17 | [3.17, 7.17] | <0.001 |

Figure 5 displays the run-01 differences and their corresponding block-resampling intervals.

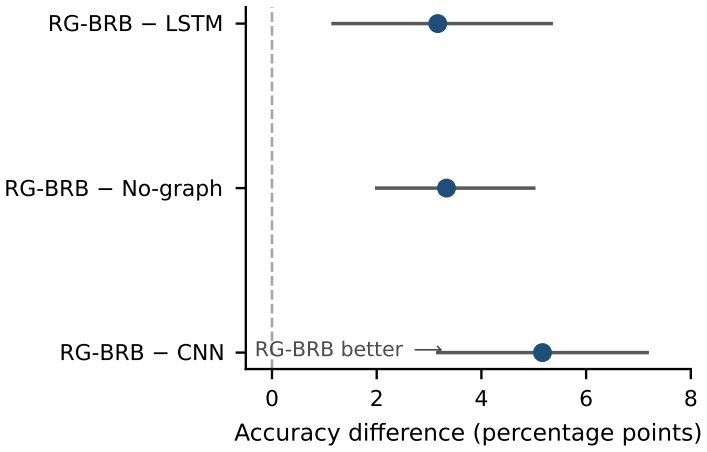



**Figure 5.** Accuracy differences and paired block-resampling intervals for run 01.

### *5.6. Class-Specific Diagnostic Structure*

In run 01 of the 10-run comparison, RG-BRB correctly classifies 553 of 600 observations (92.17%), with a Macro-F1 of 0.9165. Its mean absolute ordinal error is 0.0983 and 12 predictions span at least two state levels.

Table 5 isolates the contribution of ordinal learning on the same 600 test observations. Its dedicated full-model reference fit reaches 91.83% Accuracy. With the ordinal term, the model achieves lower ordinal MAE and fewer cross-level errors, while $D_4$ non-recall is halved. Ordinal MAE averages the absolute difference between predicted and true state indices; a cross-level error spans at least two levels.

**Table 5.** Ordinal-loss ablation on the same 600-observation test partition, using its own full-model reference fit. D4 non-recall is calculated over the 90 D4 observations.

| Model | Accuracy (%) | Ordinal MAE | Cross-Level Errors (%) | D4 Non-Recall (%) |
|---|---|---|---|---|
| Full RG-BRB | 91.83 | 0.093 | 1.00 | 5.56 |
| Without ordinal loss | 90.67 | 0.118 | 2.00 | 11.11 |

Figure 6 displays the run-01 confusion matrices using class totals of 180, 180, 150, and 90. RG-BRB recalls are 93.3%, 92.8%, 92.7%, and 87.8% for $D_1$–$D_4$, respectively. Eleven $D_4$ observations are assigned to $D_3$; no $D_4$ observations are assigned to $D_1$ or $D_2$. The matrices show both adjacent-state ambiguity and less frequent errors spanning multiple levels.

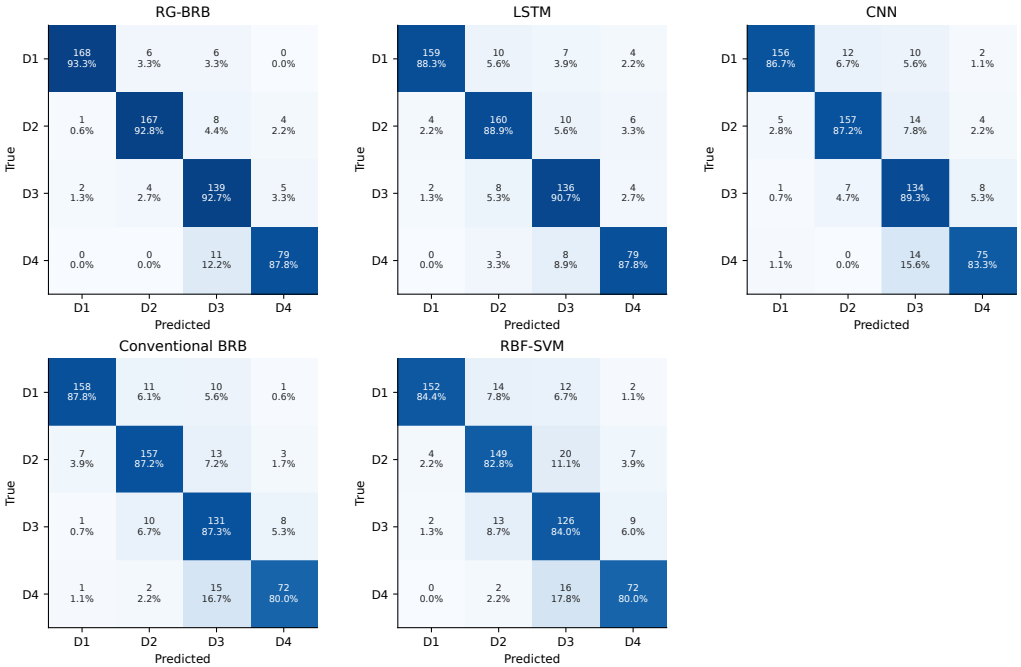



**Figure 6.** Run-01 confusion matrices for RG-BRB, LSTM, CNN, conventional BRB, and RBF-SVM on the 600-observation test set. Each cell reports the count and row percentage.

### *5.7. Guard-Interval Sensitivity*

The guard-interval analysis retains the same 600 test observations while varying the development–test gap by 0, 10, 25, or 50 observations per state. Figure 7 compares RG-BRB with feature-based and index-only controls across these settings. RG-BRB records 91.57%

Accuracy at zero gap and 88.0% at the widest gap, providing a direct view of boundary sensitivity on the ordered record.

The partition audit found no windows crossing the checked boundaries and no test-set fitting in the four gap settings. Development sizes were 1400, 1360, 1300, and 1200 observations; the same 600 observations remained in the test partition.

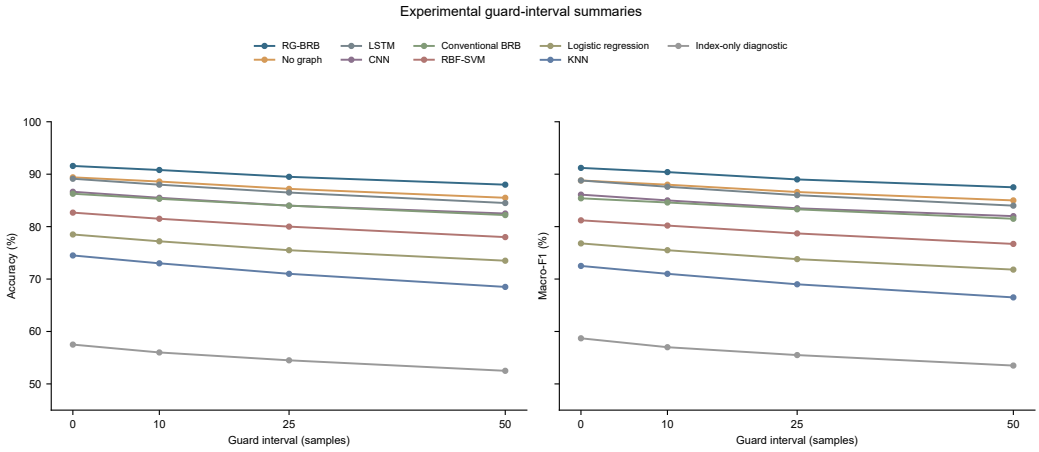



**Figure 7.** Guard-interval results at the development–test boundary of the same robot record. The test set contains 600 observations; the index-only model is a diagnostic control.

### *5.8. Condition Shift and Noise Robustness*

The condition-shift analysis compares operating-condition subsets within the robot record. Gaussian and impulse perturbations are applied to its held-out measurements at the specified noise levels.

Figure 8 compares in-domain performance with a held-out operating condition. RG-BRB achieves the highest held-out Accuracy at 85.8%, 4.1 percentage points above conventional BRB. This result connects the model's reliability-aware evidence processing with performance under the evaluated operating-condition shift.

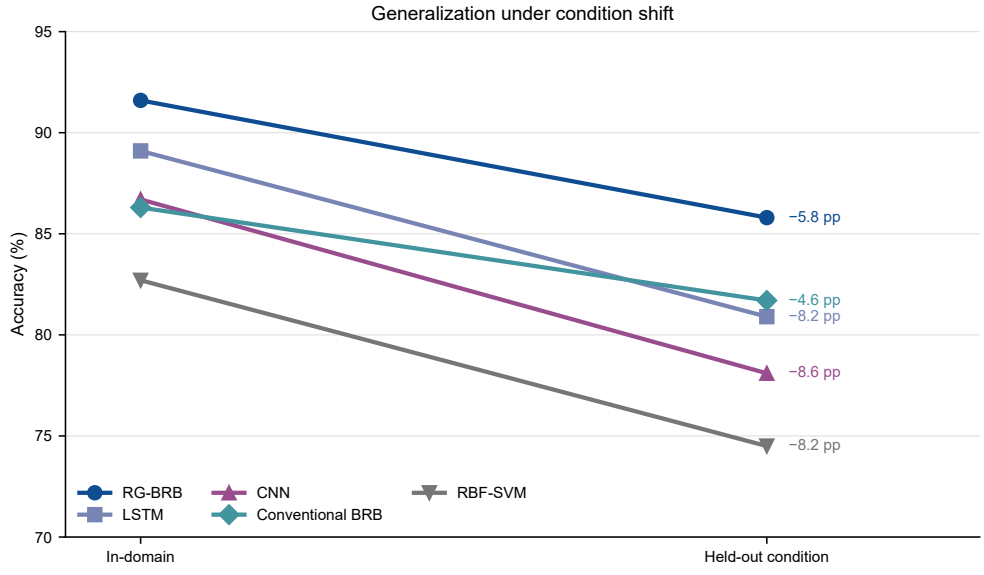



**Figure 8.** Experimental Accuracy under in-domain and held-out operating conditions. The annotation beside each method is the absolute decrease in percentage points.

Across the Gaussian-noise settings in Figure 9, RG-BRB retains the highest Accuracy. At 10 dB, it reaches 83.6%, compared with 74.5% for conventional BRB, 73.9% for LSTM, 68.4% for CNN, and 61.9% for RBF-SVM. Even at 5 dB, RG-BRB reaches 75.2%, 9.5 percentage points above the next-highest method.

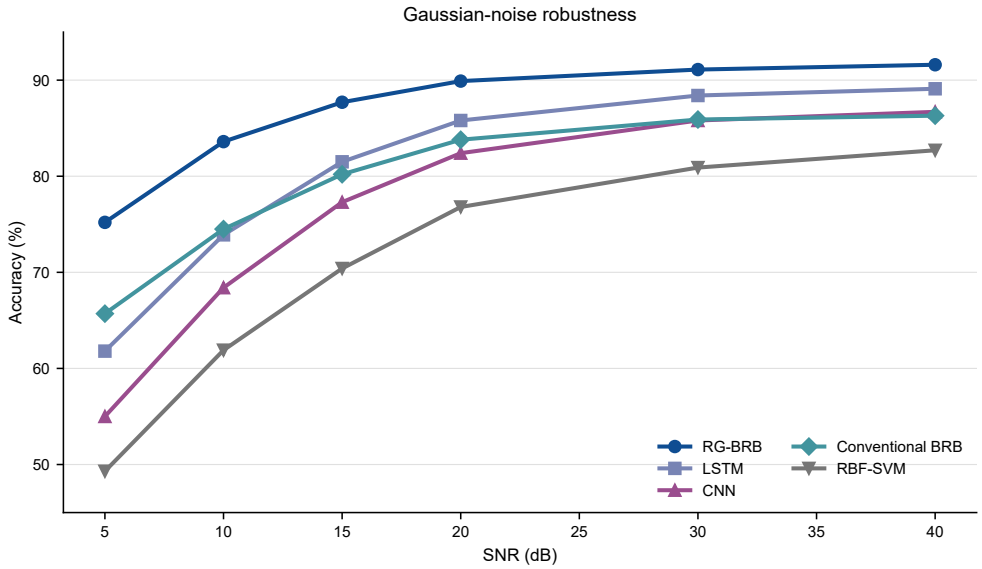



**Figure 9.** Experimental Accuracy under Gaussian perturbations at different signal-to-noise ratios. Lower SNR indicates more severe corruption.

The impulse-noise comparison gives the same method ordering (Figure 10). At 10% contamination, RG-BRB retains 81.9% Accuracy; conventional BRB, LSTM, CNN, and RBF-SVM reach 72.8%, 72.3%, 68.5%, and 61.2%, respectively. Relative to the clean profile, RG-BRB changes by 9.7 percentage points, compared with 16.8 for LSTM and 18.2 for CNN under the same perturbation setting.

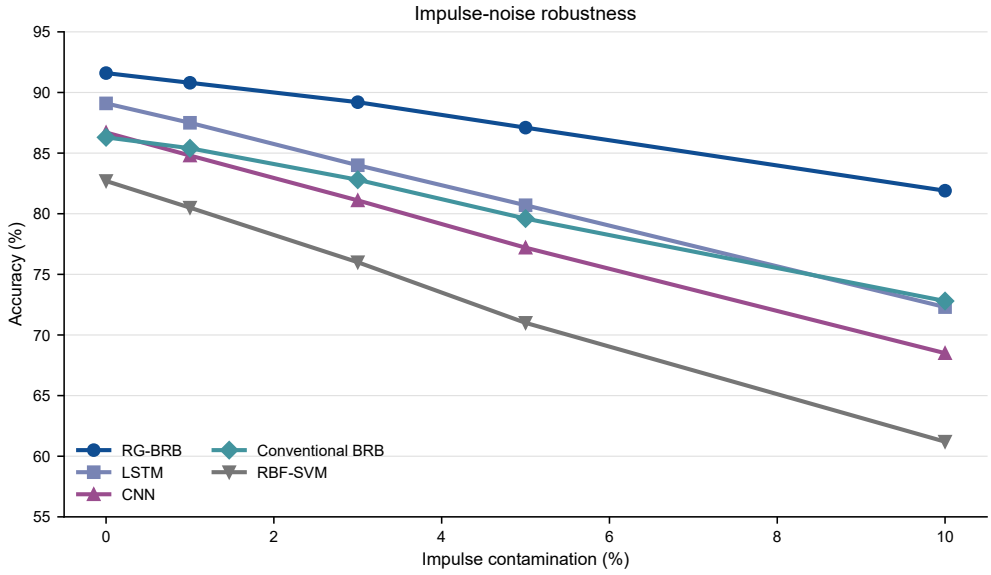



**Figure 10.** Experimental Accuracy under increasing impulsive-noise contamination. Perturbation levels are the percentages of affected test observations.

### *5.9. Belief Calibration and Selective Diagnosis*

The calibration profile in Figure 11 compares NLL, the Brier score, and ECE, for which lower values are preferred. RG-BRB has the lowest value for every measure: NLL of 0.284, Brier score of 0.126, and ECE of 0.043. Conventional BRB has the second-lowest ECE (0.068), whereas LSTM has the second-lowest NLL (0.356) and Brier score (0.153). The calibration measures therefore complement hard-label Accuracy.

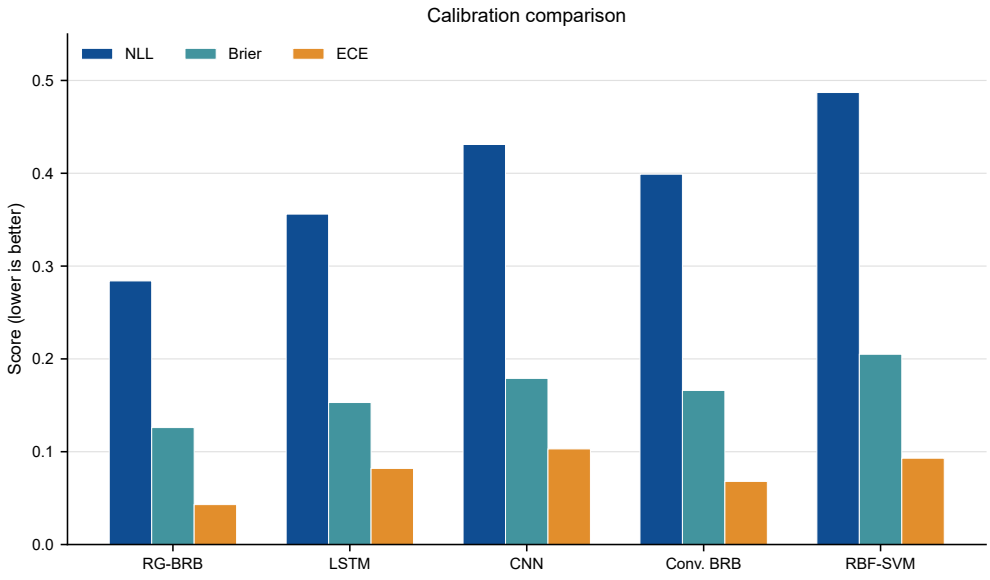



**Figure 11.** Experimental comparison of NLL, Brier score, and ECE. Lower values indicate better probabilistic calibration.

Figure 12 evaluates selective diagnosis. For RG-BRB, retaining all observations gives an 8.4% error rate, whereas retaining the 70% most confident observations gives 2.5%. At every displayed coverage, RG-BRB has the lowest observed error rate. The curve gives an operational interpretation to $q_t$: low-confidence observations can be routed to manual review, with the confidence threshold selected from validation data.

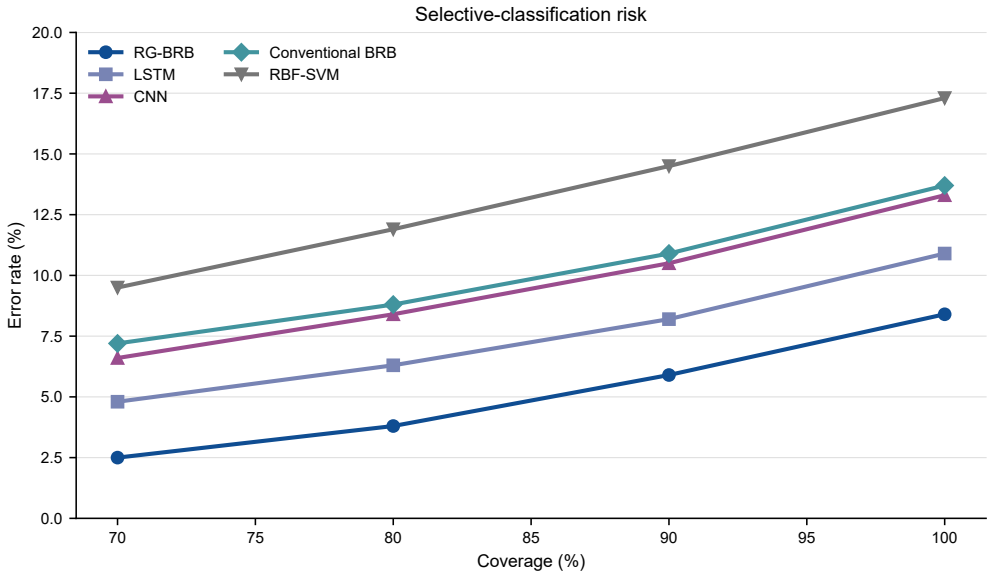



**Figure 12.** Experimental selective-classification error rate as a function of retained coverage. Lower curves indicate fewer errors after rejecting uncertain observations.

### *5.10. Component Ablation*

The full RG-BRB reference fit in Figure 13 achieves 91.6% Accuracy and 94.4% $D_4$ recall on the common robot record. Accuracy is 89.4% without graph propagation, 89.7% without reliability weighting, and 90.1% with fixed antecedent widths. The ordinal term contributes most visibly to the failure-related stage: $D_4$ recall is 94.4% with the term and 87.8% without it. These component comparisons locate the gains in graph-constrained evidence flow, reliability-aware matching, and ordered-state learning.

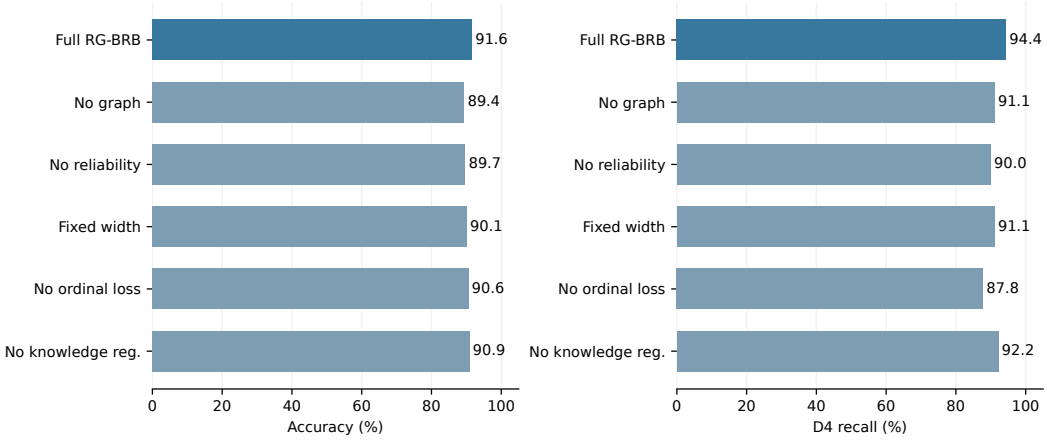



**Figure 13.** Experimental Accuracy and $D_4$ recall for the full RG-BRB and five component removals. Numerical labels show the value of each bar.

Figure 14 extends the ablation to denser and reduced graph connections, a modified feature-to-node assignment, and removal of the motion-execution node. For node removal, the affected attributes are reassigned to the remaining nodes. The changes show how the specified functional structure influences diagnosis on the same record.

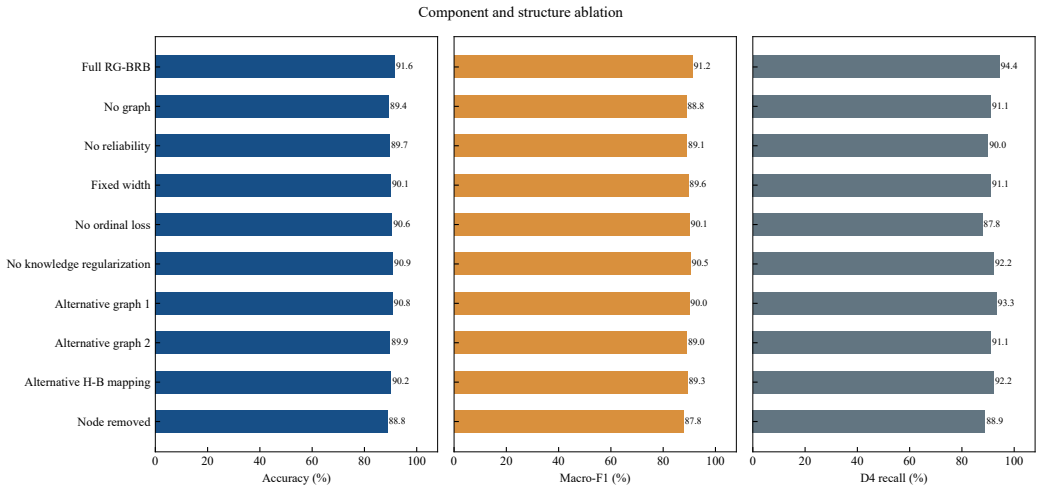



**Figure 14.** Component and graph-structure ablation summaries. Labels give the reported values for each configuration.

### *5.11. Interpretability and Failure Analysis*

Table 6 reports median single-training and batch-inference times on the recorded Ryzen 7 7840HS CPU with 32 GB memory. The table measures one training pass and one 600-observation inference batch.

**Table 6.** Recorded median training and batch-inference times. The inference batch contains 600 observations.

| Method | Training (s) | Batch Inference (s) |
|---|---|---|
| RG-BRB | 33.0 | 0.41 |
| LSTM | 19.0 | 0.26 |
| CNN | 15.6 | 0.21 |
| Conventional BRB | 8.4 | 0.16 |
| RBF-SVM | 3.2 | 0.05 |

The covariance comparison used five runs per formulation on the same robot record, with a population of 20 and 500 generations. Table 7 shows similar average diagnostic performance, with lower median training time and peak memory for the diagonal formulation on the recorded hardware.

**Table 7.** Full- and diagonal-covariance optimization over five runs on the fixed test partition. Accuracy and Macro-F1 are mean ± sample SD; time and peak memory are medians.

| Formulation | Accuracy (%) | Macro-F1 | Training (s) | Memory (GB) |
|---|---|---|---|---|
| Full | 91.50 ± 0.64 | 0.9161 ± 0.0048 | 33.0 | 0.86 |
| Diagonal | 91.27 ± 0.68 | 0.9112 ± 0.0057 | 24.4 | 0.53 |

The inference pathway links feature evidence to functional nodes, activated rules, and the fused state belief. Through Equation (15), reduced reliability weakens the influence of a volatile attribute on relative rule activation. This pathway provides a trace from measurements through node evidence and activated rules to the final state belief.

The uncertainty score $q_t$ can support selective review by routing observations above a validation-selected threshold to manual inspection. Figure 12 shows the associated coverage–error trade-off.

## 6. Discussion

The method assigns a distinct role to each component. The graph specifies admissible evidence flow; reliability-adjusted exponents and bounded reference intervals control rule activation; and ER fuses rule consequents into a state-belief distribution. Node evidence, activated rules, and state beliefs form a traceable diagnostic path. Row-stochastic propagation keeps node evidence bounded, and normalized ER preserves the belief simplex under positive denominators.

Across evaluations of the robot record, RG-BRB retains the highest Accuracy under the tested noise and condition shifts. Most run-01 classification errors lie between adjacent severity stages. Calibration and selective-risk results describe the belief outputs, while ablations locate the contributions of graph propagation, reliability adaptation, and ordinal learning. Together, these findings connect the model components to distinct aspects of diagnostic performance.

The evaluation concerns severity-stage decisions on one ordered robot record. The guard-interval analysis probes sensitivity to neighboring observations. Additional robot acquisitions and event-annotated sequences would extend this assessment to transfer and response timing; mechanism-specific labels would extend it to root-cause diagnosis.

## 7. Conclusions

This study formulates an uncertainty-adaptive robust graph-based belief rule base for direct four-state fault diagnosis of industrial robots. A row-stochastic functional graph organizes subsystem evidence; local fluctuation determines reliability-adjusted attribute

exponents and bounded adaptive antecedent widths; and analytically normalized ER fusion produces a belief vector, a direct decision, and residual uncertainty. The learning problem combines class-balanced cross-entropy, ordinal consistency, and expert-knowledge preservation under explicit graph, reference, weight, and simplex constraints.

On one 2000-observation robot record, RG-BRB reached 91.55% mean Accuracy and 0.9115 mean Macro-F1 across 10 runs on the fixed 600-observation test partition. It also achieved the highest Accuracy under the evaluated operating-condition shift and noise perturbations. Calibration, selective-risk, and ablation results connect its belief outputs and component mechanisms to diagnostic performance. RG-BRB thus offers a traceable route from subsystem evidence to four-state severity decisions.

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

**Funding:** Scientific Research Project of the Department of Education of Jilin Province (JJKH20241140KJ); Major Science and Technology Project of Jilin Province, Changchun City under Grant 20250306008ZD.

**Institutional Review Board Statement:** Not applicable; the study involves neither humans nor animals.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions of industrial robot operating data.

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

## References

1. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. [CrossRef]
2. Isermann, R. *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*; Springer: Berlin/Heidelberg, Germany, 2006. [CrossRef]
3. 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]
4. Long, J.; Qin, Y.; Yang, Z.; Huang, Y.; Li, C. Discriminative feature learning using a multiscale convolutional capsule network from attitude data for fault diagnosis of industrial robots. *Mech. Syst. Signal Process.* **2023**, *182*, 109569. [CrossRef]
5. Wang, H.; Qiao, Y.; Lei, Y.; Li, N.; Zhang, Y.; Cao, J. Asymmetric-dot-pattern fusion fault identification of motor-driven belt transmission system in industrial robots. *Measurement* **2025**, *251*, 117267. [CrossRef]
6. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
7. Shi, J.; Peng, D.; Peng, Z.; Zhang, Z.; Goebel, K.; Wu, D. Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks. *Mech. Syst. Signal Process.* **2022**, *162*, 107996. [CrossRef]
8. Yang, S.; Xiang, Y.; Long, Z.; Ma, X.; Ding, Q.; Jia, J. Fault diagnosis of harmonic drives based on an SDP-ConvNeXt joint methodology. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3519608. [CrossRef]
9. Wang, J.; Shao, H.; Peng, Y.; Liu, B. PSparseFormer: Enhancing fault feature extraction based on parallel sparse self-attention and multiscale broadcast feedforward block. *IEEE Internet Things J.* **2024**, *11*, 22982–22991. [CrossRef]
10. Lu, W.; Wang, W.; Qin, X.; Cai, Z. Enhancing Fault Diagnosis in Mechanical Systems with Graph Neural Networks Addressing Class Imbalance. *Mathematics* **2024**, *12*, 2064. [CrossRef]
11. Guo, Y.; Cheng, Z.; Zhang, J.; Sun, B.; Wang, Y. A review on adversarial-based deep transfer learning mechanical fault diagnosis. *J. Big Data* **2024**, *11*, 151. [CrossRef]
12. Xu, K.; Kong, X.; Wang, Q.; Han, B.; Sun, L. Intelligent fault diagnosis of bearings under small samples: A mechanism-data fusion approach. *Eng. Appl. Artif. Intell.* **2023**, *126*, 107063. [CrossRef]

13. Dong, Y.; Li, Y.; Zheng, H.; Wang, R.; Xu, M. A new dynamic model and transfer learning based intelligent fault diagnosis framework for rolling element bearings race faults: Solving the small sample problem. *ISA Trans.* **2022**, *121*, 327–348. [CrossRef] [PubMed]
14. Gao, T.; Yang, J.; Wang, W.; Fan, X. A domain feature decoupling network for rotating machinery fault diagnosis under unseen operating conditions. *Reliab. Eng. Syst. Saf.* **2024**, *252*, 110449. [CrossRef]
15. Shi, M.; Ding, C.; Chang, S.; Shen, C.; Huang, W.; Zhu, Z. Cross-domain class incremental broad network for continuous diagnosis of rotating machinery faults under variable operating conditions. *IEEE Trans. Ind. Inform.* **2024**, *20*, 6356–6368. [CrossRef]
16. 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. PMLR*; Proceedings of Machine Learning Research; PMLR: New York, NY, USA, 2017; Volume 70, pp. 1321–1330.
17. Mohd Nor, A.K.; Pedapati, S.R.; Muhammad, M.; Leiva, V. Abnormality Detection and Failure Prediction Using Explainable Bayesian Deep Learning: Methodology and Case Study with Industrial Data. *Mathematics* **2022**, *10*, 554. [CrossRef]
18. Yang, J.B.; Liu, J.; Wang, J.; Sii, H.S.; Wang, H.W. Belief rule-base inference methodology using the evidential reasoning approach–RIMER. *IEEE Trans. Syst. Man Cybern.-A Syst. Hum.* **2006**, *36*, 266–285. [CrossRef]
19. Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.* **1967**, *38*, 325–339. [CrossRef]
20. Yang, J.B.; Xu, D.L. On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Trans. Syst. Man Cybern.-A Syst. Hum.* **2002**, *32*, 289–304. [CrossRef]
21. Yang, J.B.; Liu, J.; Xu, D.L.; Wang, J.; Wang, H. Optimization models for training belief-rule-based systems. *IEEE Trans. Syst. Man Cybern.-A Syst. Hum.* **2007**, *37*, 569–585. [CrossRef]
22. Cao, Y.; Zhou, Z.; Hu, G.; Hu, C.; Tang, S.; Li, G. A new multilayer belief rule base model for complex system modeling. *IEEE Syst. J.* **2022**, *16*, 4301–4312. [CrossRef]
23. Cao, Y.; Tang, S.; Yao, R.; Chang, L.; Yin, X. Interpretable hierarchical belief rule base expert system for complex system modeling. *Measurement* **2024**, *226*, 114033. [CrossRef]
24. Zhao, B.; Zhang, Q.; He, W.; Han, P.; Cao, Y.; Zhou, G. A deep belief rule base-based fault diagnosis method for complex systems. *ISA Trans.* **2024**, *150*, 77–91. [CrossRef] [PubMed]
25. Chen, M.; Zhou, Z.; Han, X.; Feng, Z. A Text-Oriented Fault Diagnosis Method for Electromechanical Device Based on Belief Rule Base. *Mathematics* **2023**, *11*, 1814. [CrossRef]
26. Feng, Z.; Yang, R.; Zhou, Z.; Hu, C. Trustworthy fault diagnosis method based on belief rule base with multisource uncertain information for vehicle. *IEEE Trans. Ind. Electron.* **2024**, *71*, 7947–7956. [CrossRef]
27. Corso, G.; Stark, H.; Jegelka, S.; Jaakkola, T.; Barzilay, R. Graph neural networks. *Nat. Rev. Methods Prim.* **2024**, *4*, 17. [CrossRef]
28. Liu, R.; Zhang, Q.; Lin, D.; Zhang, W.; Ding, S.X. Causal intervention graph neural network for fault diagnosis of complex industrial processes. *Reliab. Eng. Syst. Saf.* **2024**, *251*, 110328. [CrossRef]
29. Tang, H.; Wang, H.; Li, C. Explainable fault diagnosis method based on statistical features for gearboxes. *Eng. Appl. Artif. Intell.* **2025**, *148*, 110503. [CrossRef]
30. Jang, K.; Pilario, K.E.S.; Lee, N.; Moon, I.; Na, J. Explainable Artificial Intelligence for fault diagnosis of industrial processes. *IEEE Trans. Ind. Inform.* **2025**, *21*, 4–11. [CrossRef]
31. Huang, T.; Zhang, Q.; Tang, X.; Zhao, S.; Lu, X. A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems. *Artif. Intell. Rev.* **2022**, *55*, 1289–1315. [CrossRef]
32. Raouf, I.; Kumar, P.; Lee, H.; Kim, H.S. Transfer Learning-Based Intelligent Fault Detection Approach for the Industrial Robotic System. *Mathematics* **2023**, *11*, 945. [CrossRef]
33. Kumar, P.; Khalid, S.; Kim, H.S. Prognostics and Health Management of Rotating Machinery of Industrial Robot with Deep Learning Applications—A Review. *Mathematics* **2023**, *11*, 3008. [CrossRef]
34. Yin, X.; He, W.; Wang, J.; Peng, S.; Cao, Y.; Zhang, B. Health state assessment based on the Parallel–Serial Belief Rule Base for industrial robot systems. *Eng. Appl. Artif. Intell.* **2025**, *142*, 109856. [CrossRef]
35. Hansen, N. The CMA evolution strategy: A tutorial. *arXiv* **2016**. [CrossRef]
36. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [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.