

Physics-informed Neural Networks for Acoustic Diagnostics and Prognostics of Complex Systems

Christian Fernandez Lopez,¹ Bryan T. Adey,² and Manuel Arias Chao³

^{1,2} *ETH, Zurich, 8093, Switzerland*
cferna@ethz.ch
adeyb@ethz.ch

^{1,3} *ZHAW, Winterthur, 8041, Switzerland*

³ *Delft University of Technology, 292F+VC Delft, Netherlands*
m.a.c.ariaschao@tudelft.nl

ABSTRACT

Acoustic diagnostics have emerged as a low-cost and non-invasive alternative to vibration-based condition monitoring for rotating machinery such as pumps and bearings. However, acoustic signals are highly sensitive to environmental noise, operating variability, and sensor placement, limiting the robustness and generalization of existing data-driven approaches. In addition, most acoustic diagnostic systems remain purely data-driven and rarely incorporate physical knowledge such as characteristic fault frequencies, resonances, and degradation dynamics.

This PhD research proposes a Physics-Informed framework for unsupervised acoustic fault diagnosis and prognosis of pumps and bearings. The framework integrates detection, diagnosis, and prognosis for continuous health monitoring and remaining useful life estimation. Physics-informed constraints are incorporated into the learning process to improve robustness, interpretability, and transferability across varying industrial operating conditions.

1. PROBLEM STATEMENT

Fault diagnosis and prognosis of rotating machinery, such as pumps and bearings, are critical in industrial environments to prevent unplanned downtime, reduce maintenance costs, and extend equipment lifespan.

Traditionally, condition monitoring has relied on high-precision vibration sensors to detect faults through the analysis of characteristic frequencies.

Christian Fernandez Lopez et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

More recently, acoustic diagnostics have gained attention as a low-cost and non-invasive alternative, showing promising results for early fault detection (Tang, Tian, Huang, Shi, & Ji, 2023).

Acoustic monitoring offers several advantages over contact-based sensing, particularly in harsh or inaccessible environments involving high temperatures, radiation, toxic substances, or oily surfaces, where vibration sensors may be impractical (Tang et al., 2023). However, acoustic signals are highly sensitive to environmental noise, microphone placement, and operating variability, making robust fault detection significantly more challenging.

Most existing acoustic diagnostic systems rely on supervised learning approaches that require large labeled datasets covering diverse fault conditions. In practice, obtaining representative fault data is difficult because machinery is rarely allowed to degrade to failure under normal industrial operation. Consequently, many datasets are generated using accelerated aging tests or induced defects, which often fail to capture realistic degradation dynamics and complex fault interactions. This limitation motivates the use of unsupervised and residual-based approaches that can learn primarily from healthy operating conditions.

Although unsupervised deep learning methods have demonstrated promising results, their performance often degrades when transferred from controlled laboratory environments to real industrial settings. To address this domain shift, previous works have explored domain adaptation techniques (Gu et al., 2021) and noise reduction methods (“Enhancing Sound-Based Anomaly Detection Using Deep Denoising Autoencoder”, n.d.); however, these approaches remain largely limited to simulated or controlled scenarios.

Another major limitation is that most acoustic diagnostic models remain purely data-driven and do not incorporate physical knowledge. In contrast, vibration-based diagnostics increasingly integrate physics-informed approaches by embedding characteristic fault frequencies, resonances, and degradation mechanisms into the learning process. Existing physics-informed acoustic studies remain limited. Although recent works such as (Keshun, Puzhou, Peng, & Yingkui, 2025) and (Jiang, Luo, Zhang, Shu, & Sun, 2024) incorporate physical knowledge into synthetic data generation, the integration of physical constraints within acoustic-based deep learning frameworks remains largely underexplored.

To the best of our knowledge, no prior work combines: (i) unsupervised acoustic monitoring, (ii) physics-informed constraints, (iii) simulation-driven augmentation, (iv) continual adaptation, and (v) explainable prognostics within a unified framework for rotating machinery.

2. EXPECTED CONTRIBUTIONS

This Ph.D. aims to develop a novel Physics-Informed based framework for unsupervised fault diagnosis and prognosis of pumps and bearings using acoustic data, addressing the following research questions through the defined work packages.

Research Questions:

- **RQ1:** How can anomalies be detected and characterized across evolving failure modes using acoustic signals, under varying operating conditions and noisy environments, without relying on labelled fault data?
- **RQ2:** How can simulation-based augmentation be used to generate realistic failure scenarios, and how can these be integrated to support online health monitoring and anomaly detection?
- **RQ3:** How can physics-informed modeling (e.g., fault harmonics, resonances, degradation dynamics) be effectively integrated into the learning process to improve fault mode differentiation, health index, and remaining useful life (RUL) estimation using acoustics.

2.1. WP1: Robust Unsupervised Anomaly Detection

Current methods for acoustic condition monitoring often lack robustness under diverse operational and environmental conditions. Many rely on supervised learning, which requires large volumes of labeled fault data and often fails to generalize across different operating scenarios and data acquisition setups. In this work package, we aim to explore unsupervised and self-supervised learning approaches to detect anomalies and monitor system health based solely on acoustic signals.

Preliminary analyses indicate that frequency patterns commonly associated with bearing faults are also observable in the acoustic domain, reinforcing the viability of audio-based

condition monitoring. However, further investigation is required to understand the impact of lacking directional vibration information (e.g., X, Y, Z axes) and to assess the diagnostic benefits of capturing high-frequency content. Acoustic sensors typically capture signals up to 20 kHz, and recent observations suggest that high-frequency excitations often emerge as systems begin to degrade.

To improve robustness and interpretability, we propose leveraging dimensionality reduction techniques and graph neural networks (GNNs).

Beyond offline analysis, this work also addresses how models can be maintained in production. As new failure modes or operating conditions emerge, system dynamics may change, leading to persistent shifts in reconstruction error. We therefore investigate criteria to detect such changes and to decide whether the model should be retrained, incrementally updated, or extended to capture new dynamics in an interpretable manner. This follows a continual learning perspective, aiming to adapt the model over time while preserving previously learned behavior.

2.2. WP2: Physics-Informed Data Augmentation

Recent studies have shown that synthetic datasets simulating known failure modes can significantly enhance a model's ability to distinguish between healthy and degraded operating states (Wu & Peng, 2024). These synthetic scenarios allow the learning of critical failure signatures and improve generalization to fault types that may not be explicitly represented in the training set.

This work package aims to develop a robust, physics-informed data augmentation framework for acoustic health monitoring. The generation of realistic fault data is a major challenge in many domains due to safety concerns, high cost, and operational constraints. In contrast to vibration diagnostics—where failure modes are often better characterized—the diversity and realism of acoustic fault data remain limited. We therefore propose to simulate a range of failure scenarios using validated bearing simulators and other physics-based tools.

A key objective is to reduce the domain gap between simulated and real-world acoustic data. This will involve testing techniques such as domain randomization and parameter sweeps in the simulation space, with the goal of improving downstream generalization performance of machine learning models in the presence of previously unseen faults.

2.3. WP3: Physics-Informed Prognostics

Once the current health state of the system is assessed, the next step is to estimate its RUL. In this work package, we propose to explore semi-supervised approaches for prognos-

tics based on acoustic signals, building upon anomaly detection techniques developed in WP1.

Defining RUL from acoustic signals is inherently challenging. Unlike in vibration analysis—where thresholds on signal energy (e.g., RMS) are often used to define end-of-life—there is no standardized metric in the acoustic domain. Moreover, in real-world scenarios, faulty components may continue operating for extended periods or be replaced preemptively, resulting in ambiguous end-of-life labels. To address this, we aim to approximate RUL by identifying and modeling distinct degradation trajectories and failure modes. By characterizing these patterns, we aim to model the system’s degradation trajectory over time.

This research also investigates how integrating physics-informed knowledge into acoustic analysis can enhance the identification and differentiation of failure modes. Understanding the physical principles behind the observed signals helps guide the learning process toward physically plausible and meaningful representations, thereby improving model generalization and failure prediction accuracy. For instance, specific frequency components linked to particular failure modes often manifest in both acoustic and vibration signals. Given the maturity of vibration-based diagnostics, insights from this domain can inform acoustic feature extraction and support more robust identification of fault signatures.

2.4. WP4: Orchestration and Explainability in Health Monitoring

To support long-term autonomous operation and effective decision-making, this work package focuses on building a comprehensive orchestration framework that coordinates multiple intelligent agents across anomaly detection (WP1), simulation and data augmentation (WP2), and prognosis (WP3). The system will be designed to operate in real-time, dynamically adapting to changes in operating conditions, emerging failure patterns, or sensor noise, while ensuring robust health assessment over extended time horizons.

Given that degradation processes in components such as bearings or pumps often unfold over several months, the system must be capable of recognizing both abrupt and gradual changes that may invalidate earlier assumptions or models. When such deviations are detected—e.g., a shift in the dominant fault mode or a change in background noise conditions—the orchestration module should automatically trigger retraining of models or initiate new data generation pipelines based on WP2.

A second key objective of this work package is to address the need for interpretability and human-centric decision support. To this end, we will develop a natural language explanation module that translates diagnostic insights into clear, actionable maintenance information. This module will leverage

pre-trained language models and prompt-based techniques to generate reports explaining the detected anomalies, likely causes, and recommended next steps in an intuitive format suitable for operators and engineers. The goal is to move toward a proactive acoustic PHM assistant that can continuously monitor equipment, autonomously adapt to new scenarios, and communicate system health in a transparent, trustworthy, and operationally relevant manner.

3. WORK IN PROGRESS

3.1. System Description

The investigated setup consists of a radial pump skid used for cooling applications, operated in a controlled laboratory environment where the pump is the only active noise source. Several test runs are being performed. The system is monitored using multiple sensors, but this work focuses exclusively on acoustic measurements, vibration sensors are used for baselining purposes and labeling the failure modes. Acoustic data are collected using a microphone placed about one meter from the pump, enabling non-contact measurement.

This test run provides an illustrative example of realistic run-to-failure data. A typical maintenance error was introduced before the start of the test, including slight misalignment between the motor and pump shafts. As a result, the progressive degradation of the pump skid can be investigated.

Later in the experiment, additional mechanical degradation was intentionally introduced to accelerate the degradation process and observe the system response under more severe fault conditions. These changes are visible in the spectrogram computed from the in-place microphones, as shown in Fig. 1. By analyzing the Fast Fourier transform (FFT) of the vibration sensors installed in the system, it was possible to label the different failure modes throughout the experiment.

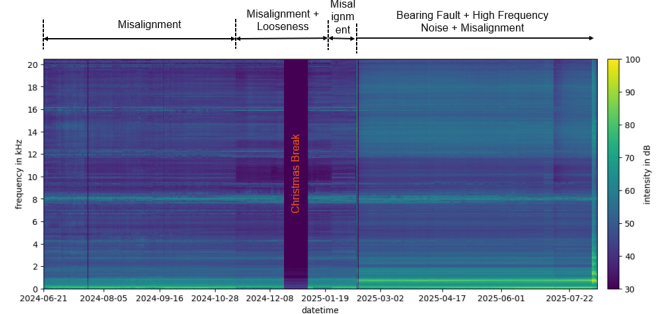


Figure 1. Spectrogram of the entire test.

3.2. Methodology

A central goal in condition monitoring is the derivation of a meaningful health index that reflects the degradation state of a system over time. In this work, we investigate two comple-

mentary autoencoder-based approaches to learn such a health representation.

The first approach is a standard residual-based autoencoder, trained exclusively on healthy operating data, assumed to correspond to the first three days of operation, in order to learn a compact representation of normal spectral patterns. Given an input signal $x \in R^d$, the encoder and decoder are defined as

$$z = f_\theta(x), \quad \hat{x} = g_\phi(z), \quad (1)$$

where $f_\theta(\cdot)$ and $g_\phi(\cdot)$ denote the encoder and decoder networks, respectively. The model is trained by minimizing the reconstruction loss on healthy data:

$$\mathcal{L}_{\text{rec}} = \frac{1}{N_h} \sum_{i \in \mathcal{H}} \|x_i - \hat{x}_i\|^2, \quad (2)$$

where N_h denotes the number of samples in healthy operating conditions. After training, the model is applied to the full dataset. The anomaly score is defined as the reconstruction error:

$$e_i = \|x_i - \hat{x}_i\|^2. \quad (3)$$

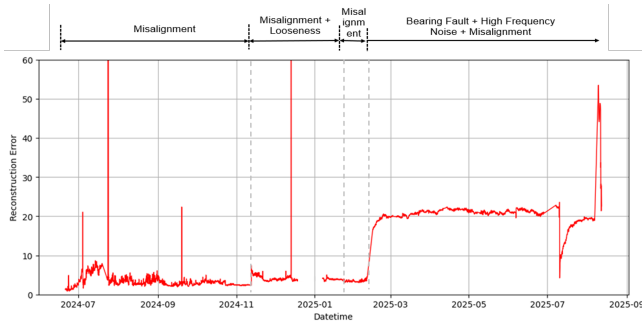


Figure 2. Reconstruction error of the healthy-trained autoencoder on the complete test set.

The second approach is a latent-constrained autoencoder designed to learn an interpretable representation of system health by explicitly structuring the latent space as a one-dimensional health indicator. Unlike the previous model, it is trained using both healthy and degraded data and maps each input to a scalar latent variable:

$$z_i = f_\theta(x_i), \quad \hat{x}_i = g_\phi(z_i), \quad z_i \in R. \quad (4)$$

In this formulation, the latent variable is directly interpreted as a health index, where higher or lower values correspond to different stages of system degradation along a one-dimensional manifold. The model is trained using a composite loss function:

$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_{\text{rec}} + \beta \mathcal{L}_{\text{anchor}} + \gamma \mathcal{L}_{\text{reg}} + \delta \mathcal{L}_{\text{dis}}. \quad (5)$$

The individual loss components are defined as follows. The reconstruction loss is computed over the complete test, including both healthy and degraded operating conditions:

$$\mathcal{L}_{\text{rec}} = \frac{1}{N} \sum_{i=1}^N \|x_i - \hat{x}_i\|^2, \quad (6)$$

where N denotes the total number of available samples. Anchor loss (healthy constraint): enforces latent representation of healthy samples $\mathcal{H}$ to be close to 1, defining a health reference point.

$$\mathcal{L}_{\text{anchor}} = \frac{1}{N_h} \sum_{i \in \mathcal{H}} (z_i - 1)^2, \quad (7)$$

Latent regularization: encourages a compact and stable latent space.

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^N z_i^2. \quad (8)$$

Latent variance maximization: encourages dispersion of latent representations around their mean. This term prevents latent collapse and enhances the separation between different operating conditions and degradation states.

$$\mathcal{L}_{\text{dis}} = -\frac{1}{N} \sum_{i=1}^N \left| z_i - \frac{1}{N} \sum_{j=1}^N z_j \right|^2. \quad (9)$$

The two approaches are evaluated in terms of anomaly detection performance and interpretability of the resulting health indicators. Both methods successfully capture system changes, while the latent-constrained approach provides improved interpretability and better separation of failure modes, partly due to training on both healthy and degraded data. Ongoing work explores a continual learning framework to incrementally identify, segment, and distinguish different failure modes and their combinations over time.

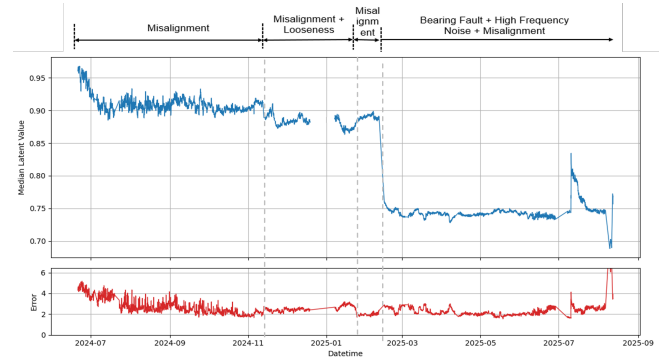


Figure 3. Latent space evolution of the constrained autoencoder and corresponding reconstruction error over time.

REFERENCES

- Enhancing sound-based anomaly detection using deep denoising autoencoder. (n.d.). , 12.
- Gu, X., Li, R., Kang, M., Lu, F., Tang, D., & Peng, J. (2021, October). Unsupervised adversarial domain adaptation abnormal sound detection for machine condition monitoring under domain shift conditions. In *2021 IEEE 20th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)* (pp. 139–146). doi: 10.1109/ICCICC53683.2021.9811305
- Jiang, M., Luo, M., Zhang, C., Shu, M., & Sun, G. (2024, November). Rolling bearing fault diagnosis based on acoustic-vibration data fusion and mode decomposition combined with the crested porcupine optimization algorithm. *Heliyon*, 10(22), e40351. doi: 10.1016/j.heliyon.2024.e40351
- Keshun, Y., Puzhou, W., Peng, H., & Yingkui, G. (2025). A sound-vibration physical-information fusion constraint-guided deep learning method for rolling bearing fault diagnosis. , 253, 110556. doi: 10.1016/j.res.2024.110556
- Tang, L., Tian, H., Huang, H., Shi, S., & Ji, Q. (2023). A survey of mechanical fault diagnosis based on audio signal analysis. , 220, 113294. doi: 10.1016/j.measurement.2023.113294
- Wu, X., & Peng, D. (2024). Acoustic monitoring of rolling element bearings using simulation-based partial domain adaptation.