Electronics Letters

LETTER OPEN ACCESS

Compact Acoustic Signal Portraits for Resource-Efficient Mechanical Fault Detection

Oleksandr Volkov[1] | Oleksandr Bushma[2] | Artem Sieriebriakov[1] | Volodymyr Taranukha[3] | Yaroslav Linder[3]

[1]Institute of Information Technologies and Systems of National Academy of Science of Ukraine, Kyiv, Ukraine | [2]Department of Computer Sciences, Borys Grinchenko Kyiv Metropolitan University, Kyiv, Ukraine | [3]Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Correspondence: Oleksandr Bushma (o.bushma@kubg.edu.ua)

Received: 25 December 2025 | **Revised:** 2 February 2026 | **Accepted:** 11 February 2026

ABSTRACT

A resource-efficient semi-analytical method for mechanical fault detection via acoustic signal analysis is presented. A compact signal portrait is constructed by tracking dominant salient-frequency peaks over time, without a priori assumptions on peak locations. Peak feature statistics are aggregated into a 3D matrix that encodes nominal oscillatory modes. Fault detection is achieved by measuring the deviation between current and healthy portraits using a single decision threshold. Experiments on internal combustion engine recordings demonstrate improved separability and lower computational cost compared to empirical mode decomposition, confirming suitability for resource-constrained monitoring.

1 | Introduction

Recognising acoustic signal sources in mechanical systems is critical for predictive maintenance, enabling timely fault detection, minimising downtime, and preserving equipment performance [1]. Acoustic monitoring captures real-time information on friction, impacts, and resonance, providing a reliable tool for industrial diagnostics [2, 3]. While deep learning approaches achieve high diagnostic accuracy for rotating machinery, they face practical limitations: they require large labelled datasets under diverse conditions, generalise poorly to unseen environments, are difficult to interpret, and are sensitive to noise, data imbalance, or atypical events [4–6]. High computational and power demands further restrict deployment on resource-constrained platforms, limiting applicability for real-time embedded systems [7]. Traditional acoustic diagnostics using spectral, wavelet, or acoustic emission analysis rely on manually crafted features and rule-based classifiers, effectively detecting specific harmonic or impulse patterns indicative of mechanical faults [2, 8]. Their accuracy and robustness, however, decline under noisy or variable conditions, limiting industrial applicability. Nevertheless, low computational cost and predictable behaviour make them suitable for resource-constrained hardware platforms [2, 9–11].

The aim of this work is to propose and validate a computationally efficient method for mechanical fault detection in moving machinery, which represents acoustic signals as compact portraits formed from salient frequency-domain features, enabling reliable condition monitoring with minimal computational overhead.

The novelty of the proposed method lies in a semi-analytical construction of compact acoustic signal portraits that represent oscillatory behaviour of moving machinery via the temporal evolution of dominant salient-frequency peaks. Peak intervals with maximal relative energy are estimated in local signal segments without a priori assumptions on their positions or envelopes, enabling informative feature extraction at low computational cost. The obtained peak statistics are aggregated across salient frequencies to form a compact three-dimensional matrix portrait of the nominal operating state. Mechanical faults are detected

by measuring deviations from a precomputed healthy reference using a single decision threshold.

Experimental validation on acoustic recordings of internal combustion engines (ICE) demonstrates improved fault separability and lower computational cost compared to an Empirical Mode Decomposition (EMD)-based method, confirming suitability for resource-constrained condition monitoring [12].

2 | Methodology

Acoustic fault detection in mechanical systems relies on the analysis of audio signal portraits. These portraits are represented as rectangular matrices that lie conceptually between the raw signal representation (such as the Fourier transform) and feature-based descriptors (such as spectral or energy measures). Two portrait construction methods are considered below: the proposed approach and the EMD method.

Fault-related processes in complex mechanical systems (like wear, cracking, or deformation) cause distinctive acoustic events that are periodic yet brief and time-dispersed. In ICEs, such periodic variations reflect oscillatory interactions among components, which are difficult to reproduce artificially. Individual contributions cannot be discerned separately; thus, analysis must consider the combined effect on the overall acoustic signal.

The study employed audio data from ICEs of motorcycles. Original sound samples were sourced from YouTube using Google Audioset [13] and organised into two datasets [14]. The first dataset consists of 96 audio recordings, each compiled from 5-second excerpts of properly operating motorcycle engines. It includes two 2-stroke models: Simson S51 and MZ ETZ 250 (ETZ), and two 4-stroke ones: Yamaha YZF-R6 (YZF) and Honda CBR600RR (CBR).

A second semi-synthetic dataset of equal size was generated to represent malfunction conditions in the same engine models. Faulty ICE audio signals were analysed. Dataset generation accounted for untimely ignitions and misfires, as well as changes in valve opening/closing timing [15–18]. Fault-induced acoustic patterns can be derived from healthy system recordings [19, 20]. Based on this analysis, effects of different failures were applied individually to each sample. They were simulated in the form of appropriate peak alteration achieved by applying peak-dependent weighting functions to healthy ICE's audio signals.

The first dataset was split into training and testing subsets (3:1 ratio), while the second dataset was used exclusively for testing, as fault-type identification was beyond the scope of this study. The fault detection procedure relied on a quantitative measure of distance between a specific signal's portrait and the average portrait for the corresponding model from the training dataset. The computed measure is expressed in dimensionless units and reflects both the temporal behaviour and the signal magnitude, as discussed later. Owing to the acoustic characteristics of the data and the portrait extraction method, a low-complexity classifier proved sufficient for reliable fault detection.

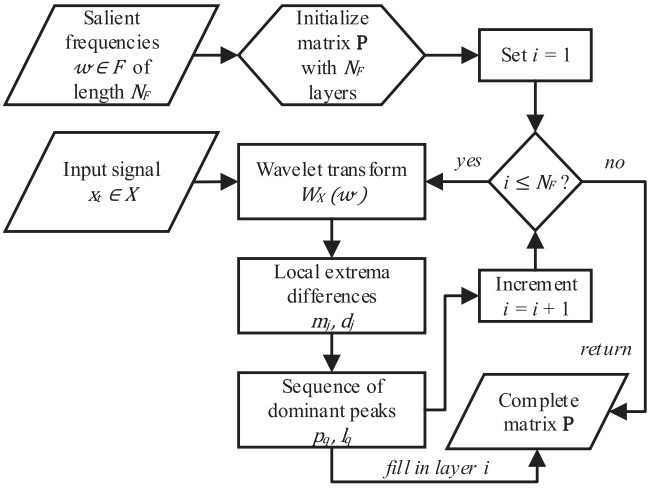

FIGURE 1 | Portrait of oscillatory modes calculation.

3 | Signal Portrait Formation

The proposed approach assumes that the considered audio signal can be viewed as oscillatory processes produced by mechanical parts of the ICE working together. When these oscillations repeat over time, they form patterns (oscillatory modes) described by quasi-periodic sound waves. The shape of such a wave can be captured by identifying its dominant peaks, and their sequence provides a clear way to describe such oscillatory modes.

The new approach, named Peak Detection at Cycle Boundaries (PDCB), further assumes that oscillations can be treated independently of the specific mechanical parts producing them. Hence, it identifies the salient frequencies of the signal and analyses each one separately. The fundamental frequency is considered quasi-constant within a fixed spectral band. Features derived from peak distributions at these frequencies describe the system's oscillatory modes and enable identification of the source. The flowchart of the respective procedure for obtaining a portrait of oscillatory modes based on salient frequencies is shown in Figure 1.

First, a discrete short-time Fourier transform is applied to X, where $x_t \in X$ is the input discrete signal, $t \in [1 : N_X]$, N_X is the number of samples with a frequency f_s [21]. As a result, the values $\chi(m, f)$ are obtained, where m and f are the time and the frequency index, respectively.

To balance the difference between large and small values, the logarithmic compression operation $\Gamma(m, f)$ is applied to χ [22]. Next, the frequencies are converted into logarithmic intervals, measured in cents. The lowest reference frequency ω_{ref} is set, below which the signal does not contain significant features. Then the discrete limits of the N_B logarithmic intervals $B(\omega)$, for the frequency $\omega \ge \omega_{ref}$ are defined as $[\lfloor(\frac{\omega}{\omega_{ref}})\rfloor, \lceil(\frac{\omega}{\omega_{ref}})\rceil]$. These limits set the bounds for indices f, in which the sum of the compressed values Γ is calculated. The result is the sum of the coefficients $\mathbf{Y}_{\log}(m, b)$ in logarithmic intervals with indices $b \in [1 : N_B]$. Next, the spectral vector of maximal energies is obtained as a row-wise sum $F_{max}(b) = \sum_m \mathbf{Y}_{\log}$. From it, based on local maxima, a set of salient frequencies $F = \{w_1, w_2, \ldots, w_{N_F}\}$ of length N_F is determined, measured in hertz.

In a loop for each salient frequency F the following is carried out. The continuous wavelet transform $Y_w(t) = W_X(\omega)$ of the signal X is calculated with values $y_t \in Y_w$ for time indices t [23]. Next, local extrema indices $e_i \in E$ are determined for local maxima $E = (t \mid y_{t-1} \langle y_t \rangle y_{t+1})$, with indices $i \in [1, N_E]$. Then, sets of differences are calculated both for the amplitudes of the extrema $m_j = Y_w(e_{j+1}) - Y_w(e_j)$ and for the distances between the extrema $d_j = e_{j+1} - e_j$, where $j \in [1, N_E - 1]$.

Among the obtained extrema, the ones with the greatest local energy are selected. Based on the reference index k_l, $k_l < j$, the growth rate is defined for the interval $[k_l, j]$ as:

$$R(k_l, j) = \sum_{k=k_l}^{j} m_k \Big/ \sum_{k=k_l}^{j} d_k$$

Then for a fixed k_l the starting points $k_>$ of intervals with sufficient growth rates are estimated as

$$A(k_l, j) = \{k_> \mid (k_l < k_> < j) \& (R(k_>, j) > R(k_l, j))\}$$

For the current index k_l iterations continue for $j > k_l$ until j^{stop}, which defines the search range. It is determined empirically based on available computational resources.

From the obtained set $A(k_l, j)$ the starting index of the interval with the highest relative growth rate p is selected, defining the local dominant peak of the signal. After that, next index k_l is searched. The passage through k_l is executed until the end of the signal. For the resulting peak sequence of length N_P, the distributions of growth rates $(p_q)_{q=1}^{N_P}$ and the inter-peak distances $(l_q)_{q=1}^{N_P}$ (in samples) are determined. For each p_q and l_q the mean and standard deviation are computed as two-component tuples s_p and s_l respectively.

The execution ends when all frequencies F have been processed. Each statistical tuple is placed in a square array $[s_p; s_l]$. These arrays are stacked into a 3D matrix $\mathbf{P}$, with the outermost layer dimension corresponding to elements of F.

The resulting output represents the portrait of the oscillatory modes $\mathbf{P}$ of signal X. The algorithm may continue to search secondary, lower-energy peaks by subtracting the identified sequence from the signal. This iterative process reveals peak sequences across decreasing energy levels, though for the current analysis only the uppermost, dominant oscillatory mode is considered.

4 | EMD-Based Signal Portrait Formation

For comparison with the PDCB algorithm, the EMD algorithm was chosen because it extracts the intrinsic mode functions (IMFs) with similar properties [12]. Within EMD, IMF denotes an oscillatory component that reflects separate levels of signal variation. It is time-dependent, representing oscillations at a specific scale, while its instantaneous frequency and amplitude vary dynamically with the signal.

The EMD algorithm is the basis for the Huang-Hilbert transform (HHT), so the results relevant for HHT are also relevant for EMD. Accordingly, any criterion based on HHT involves the use of the obtained IMFs. For signal portrait formation, results of HHT/EMD comparison with other algorithms were considered [24]. Eventually, a similar approach to PDCB was used to obtain comparable results and performance estimates.

For each IMF obtained from X, local maxima are determined. Then, local peak magnitudes and distances in-between are estimated using j^{stop} threshold. From the obtained peak sequence, the matrix $\mathbf{P}_{\text{emd}}$ is formed as described above. Layer indices of the outermost dimension correspond to the respective IMFs.

Lightweight spectral feature extraction methods were analysed and it was found that EMD-related features work better than MFCC, ZCR, spectral roll-off, energy, kurtosis, and spectral flatness [25–27]. Also, EMD-based approach outperforms WPD-based approach [28]. Noise injection did not improve performance with current setup unlike in [29, 30] so it is not used in the final experiment.

5 | Audio Signals Portraits Comparison

Mean-squared quantitative measure of difference was used to compare the portraits with interpolation reducing outlier effects. The measure was applied to compare the datasets described above and exhibited sufficient sensitivity for fault detection.

The distance between signal portraits is computed as the relative difference between the mean and standard deviation of the compared distributions for both s_p and s_l. These intermediate differences are quadratically averaged across all matrix layers, each corresponding to a salient frequency (or IMF for EMD). The resulting dimensionless measure expresses how many standard deviations separate the two portraits.

Table 1 presents differences computed by the PDCB and EMD algorithms (first and second values in each cell, respectively) relative to average healthy portraits, grouped by model. The left column ("Healthy") shows maxima of distances for healthy cases, while the right column ("Faulty") shows minima for faulty cases, enabling straightforward separation between normal and defective states.

The results for PDCB show that the values on the left are smaller than those on the right, forming a separation interval of at least 7 units for the same model. This allows reliable fault detection using a simple difference-threshold classifier. Moreover, results demonstrated correct model assignment for healthy engines. The practical threshold for a healthy ICE is a difference value below 2.

For EMD, the values on the left show stronger overlap with those on the right, with a reduced separation interval of up to 2.6 units for the same model. This demonstrates that EMD provides poorer class separability, confirming the superior performance of the proposed method. Likely the observed effect arises because EMD uses a fixed window, ignoring data-driven cues, sometimes generating suboptimal IMFs and reducing analysis quality [31].

TABLE 1 | Experimental results for the PCDB and EMD algorithm.

Model	Healthy				Faulty			
	S51	ETZ	YZF	CBR	S51	ETZ	YZF	CBR
S51	1.5 / 1.4	1.7 / 3.4	2.4 / 4.0	2.0 / 3.5	9.0 / 1.9	7.0 / 2.0	7.3 / 2.2	5.8 / 2.4
ETZ	4.0 / 2.9	1.5 / 1.6	1.9 / 1.8	1.5 / 1.7	17 / 3.5	16 / 1.9	14 / 1.7	11 / 3.0
YZF	18 / 6.1	15 / 6.8	1.4 / 2.1	2.2 / 2.9	39 / 13	18 / 5.8	17 / 4.8	14 / 9.1
CBR	11 / 5.8	8.6 / 6.6	2.8 / 2.8	1.5 / 2.0	36 / 11	15 / 1.8	20 / 1.7	14 / 3.3

To support obtained results, the Naive Bayes classifier has been applied to the features extracted by PDCB and EMD. Resulting accuracy, as the ratio of correct predictions to all predictions, is 1.0 and 0.86, respectively.

6 | Computational Complexity Comparison

A comparison of runtime for known algorithms is given as follows [32]: DFT — $O(N^2)$; FFT — $O[N\log(N)]$; STFT — $O[N\log(N)]$; CWT — $O(S\cdot N)$ if the number of scales S is constant, but it tends toward $O[S\cdot N\log(N)]$ and, in certain implementations, to $O[N\log^2(N)]$; and EMD/HHT — $O[N\log(N)]$. Additional factors arise in EMD from IMF search and window selection, yielding a complexity upper bound of $O[W\cdot S\cdot N\log(N)]$, with W denoting the window selection functions.

For the developed algorithm, instead of enumerating all oscillatory modes, salient frequencies are sorted, replacing one multiplier with another. This yields a complexity of $O[C\cdot N\log(N)]$, with C defined differently for EMD depending on analysis depth. This bound is lower than that of EMD-based approaches.

7 | Conclusion

A resource-efficient PDCB-based approach has been presented that exploits dominant salient-frequency peaks for characterising oscillatory modes in acoustic signals. Peak sequences are obtained by estimating intervals with maximal relative energy in local segments, enabling the construction of compact signal portraits suitable for low-complexity analysis. Experimental evaluation of motorcycle ICE recordings confirms the effectiveness of the proposed method for fault detection, where a simple difference-threshold classifier reliably separates healthy and faulty operating conditions. Compared with EMD-based methods, the proposed approach provides higher discriminability, reduced sensitivity to edge effects, and a computational complexity, making it well suited for condition monitoring of moving machinery on resource-constrained platforms.

Author Contribution

Oleksandr Volkov: conceptualisation, formal analysis, funding acquisition, investigation, project administration, resources, supervision. **Oleksandr Bushma:** conceptualisation, formal analysis, methodology, project administration, supervision, validation, writing – review and editing. **Artem Sieriebriakov:** conceptualisation, data curation, methodology, software, validation, visualisation, writing – original draft. **Volodymyr Taranukha:** conceptualisation, data curation, methodology, software, validation, visualisation, writing – original draft. **Yaroslav Linder:** conceptualisation, data curation, methodology, software, validation, visualisation, writing – original draft.

Acknowledgements

The research is done under the grant support of the Ukrainian National Research Fund within the framework of the "Modern hardware and software complex for unmanned aerial vehicle operators training" project (registration number 2023.04/0082) implemented within the Science to Strengthen Defence Capabilities of Ukraine competition.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Google AudioSet at https://research.google.com/audioset/.and GitHub at https://github.com/aceart1002/ICE_Malfunction_Detection, reference numbers 13 and 14.

References

1. Z. Ye, Z. Liu, and H. Wang, "Review on Sound-Based Industrial Predictive Maintenance: From Feature Engineering to Deep Learning," *Mathematics* 13, no. 11 (2025): 1724, https://doi.org/10.3390/math13111724.

2. P. Kundu, "Review of Rotating Machinery Elements Condition Monitoring Using Acoustic Emission Signal," *Expert Systems with Applications* 252 (2024): 124169, https://doi.org/10.1016/j.eswa.2024.124169.

3. H. Wang and J. Xie, "Fault Diagnosis of Rolling Bearings Based on Acoustic Signals in Strong Noise Environments," *Applied Sciences* 15, no. 3 (2025): 1389, https://doi.org/10.3390/app15031389.

4. Z. Zhu, "A Review of the Application of Deep Learning in Intelligent Fault Diagnosis of Rotating Machinery," *Measurement* 206 (2023): 112346, https://doi.org/10.1016/j.measurement.2022.112346.

5. B. A. Tama, M. Vania, S. Lee, and S. Lim, "Recent Advances in the Application of Deep Learning for Fault Diagnosis of Rotating Machinery Using Vibration Signals," *Artificial Intelligence Review* 56 (2023): 4667–4709, https://doi.org/10.1007/s10462-022-10293-3.

6. J. Zhao, W. Wang, J. Huang, and X. Ma, "A Comprehensive Review of Deep Learning-Based Fault Diagnosis Approaches for Rolling Bearings: Advancements and Challenges," *AIP Advances* 15, no. 2 (2025): 020702, https://doi.org/10.1063/5.0255451.

7. D. Neupane, M. R. Bouadjenek, R. Dazeley, and S. Aryal, "Data Driven Machinery Fault Detection: A Comprehensive Review," *Neurocomputing* 627 (2024): 129588, https://doi.org/10.1016/j.neucom.2025.129588.

8. R.-V. Sánchez, J. C. Macancela, L.-R. Ortega, D. Cabrera, F. P. García Márquez, and M. Cerrada, "Evaluation of Hand-Crafted Feature Extraction for Fault Diagnosis in Rotating Machinery: A Survey," *Sensors* 24, no. 16 (2024): 5400, https://doi.org/10.3390/s24165400.

9. D. Aiordăchioaie, "A Comparative Analysis of Fault Detection and Process Diagnosis Methods Based on a Signal Processing Paradigm," *Discover Applied Sciences* 7 (2025): 10, https://doi.org/10.1007/s42452-024-06390-3.

10. S. Zhang, J. Zhou, E. Wang, H. Zhang, M. Gu, and S. Pirttikangas, "State of the Art on Vibration Signal Processing towards Data-Driven Gear Fault Diagnosis," *IET Collaborative Intelligent Manufacturing* 4, no. 4 (2022): 249–266, https://doi.org/10.1049/cim2.12064.

11. Y. Zhou, Z. Ma, and L. Fu, "A Review of Key Signal Processing Techniques for Structural Health Monitoring: Highlighting Non-Parametric Time-Frequency Analysis, Adaptive Decomposition, and Deconvolution," *Algorithms* 18, no. 6 (2025): 318, https://doi.org/10.3390/a18060318.

12. A. Quinn, V. Lopes-dos-Santos, D. Dupret, A. Nobre, and M. Woolrich, "EMD: Empirical Mode Decomposition and Hilbert-Huang Spectral Analyses in Python," *The Journal of Open Source Software* 6, no. 59 (March 2021): 2977, 10.21105/joss.02977.

13. "Google AudioSet," Google AudioSet, https://research.google.com/audioset/.

14. "GitHub," ICE_Malfunction_Detection, https://github.com/aceart1002/ICE_Malfunction_Detection.

15. A. I. Alahmer, W. M. Adaileh, and M. A. Al Zubi, "Monitoring of a Spark Ignition Engine Malfunctions Using Acoustic Signal Technique," *International Journal of Vehicle Noise and Vibration* 10, no. 3 (2014): 201–213.

16. S. K. Mathew and Y. Zhang, "Acoustic-Based Engine Fault Diagnosis Using WPT," *PCA and Bayesian optimization Applied Sciences* 10, no. 19 (2020): 6890.

17. S. Singh, S. Potala, and A. R. Mohanty, "An Improved Method of Detecting Engine Misfire by Sound Quality Metrics of Radiated Sound," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 233, no. 12 (2019): 3112–3124.

18. B. S. Anami and V. B. Pagi, "Fault Detection and Localization in Motorcycles Based on the Chain Code of Pseudospectra and Acoustic Signals," *Journal of Mechanical Engineering and Sciences* 4 (2013): 440–451.

19. J. Salamon, D. MacConnell, M. Cartwright, P. Li, and J. P. Bello, "Scaper: A Library for Soundscape Synthesis and Augmentation," 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA 2017 (October 2017): 344–348, https://doi.org/10.1109/WASPAA.2017.8170052.

20. K. Wilkinghoff, "Self-Supervised Learning for Anomalous Sound Detection," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, April), 276–280.

21. J. B. Allen and L. R. Rabiner, "A Unified Approach to Short-Time Fourier Analysis and Synthesis," *Proceedings of the IEEE* 65, no. 11 (1977): 1558–1564, https://doi.org/10.1109/PROC.1977.10770.

22. M. Müller, *Fundamentals of Music Processing* (Springer International Publishing, 2015), https://doi.org/10.1007/978-3-319-21945-5.

23. C. Torrence and G. P. Compo, "A Practical Guide to Wavelet Analysis," *Bulletin of the American Meteorological Society* 79, no. 1 (January 1998): 61–78, https://doi.org/10.1175/1520-0477(1998)079⟨0061:APGTWA⟩2.0.CO;2.

24. A. Orhan, N. Yordanov, M. Ertarğın, M. Zhilevski, and M. Mikhov, "A Comparative Study of Time–Frequency Representations for Bearing and Rotating Fault Diagnosis Using Vision Transformer," *Machines* 13, no. 8 (August 2025): 737, https://doi.org/10.3390/machines13080737.

25. A. Kumar, S. S. Solanki, and M. Chandra, "Hilbert Spectrum Based Features for Speech/Music Classification," *Serbian Journal of Electrical Engineering* 19, no. 2 (2022): 239–259.

26. A. Ghosal, B. C. Dhara, and S. K. Saha, "Speech/Music Classification Using Empirical Mode Decomposition," in *2011 Second International Conference on Emerging Applications of Information Technology* (IEEE, February), 49–52.

27. M. Amarnath and I. P. Krishna, "Local Fault Detection in Helical Gears via Vibration and Acoustic Signals Using EMD Based Statistical Parameter Analysis," *Measurement* 58 (2014): 154–164.

28. M. Kedadouche, M. Thomas, and A. J. M. S. Tahan, "A Comparative Study Between Empirical Wavelet Transforms and Empirical Mode Decomposition Methods: Application to Bearing Defect Diagnosis," *Mechanical Systems and Signal Processing* 81 (2016): 88–107.

29. C. Yang, Z. Qiao, L. Liu, A. Kumar, and R. Zhu, "Positive-Incentive Noise in Artificial Intelligence-Enabled Machine Fault Diagnosis," *Structural Health Monitoring* (Early view).

30. Y. Changpu, Z. Qiao, R. Zhu, X. Xu, Z.-C. Lai, and S. Zhou, "An Intelligent Fault Diagnosis Method Enhanced by Noise Injection for Machinery," *IEEE Transactions on Instrumentation and Measurement* 72 (2023): 1–11.

31. H. Jeon, Y. Jung, S. Lee, and Y. Jung, "Area-Efficient Short-Time Fourier Transform Processor for Time–Frequency Analysis of Non-Stationary Signals," *Applied Sciences* 10, no. 20 (October 2020): 7208, https://doi.org/10.3390/app10207208.

32. M. Frigo and S. G. Johnson, "The Design and Implementation of FFTW3," *Proceedings of the IEEE* 93 (2005), 216–231, https://doi.org/10.1109/JPROC.2004.840301.