sensors

MDPI

Article

Robust Sparse Non-Negative Matrix Factorization for Identifying Signals of Interest in Bearing Fault Detection

Hamid Shiri [1] and Anna Michalak [2,*]

[1] Tony Davies High Voltage Laboratory, School of Electronics and Computer Science, Faculty of Engineering and Physical Sciences, University of Southampton, Southampton SO17 1BJ, UK; h.shiri@soton.ac.uk
[2] Faculty of Geoengineering, Mining and Geology, Wroclaw University of Science and Technology, Na Grobli 15, 50-421 Wrocław, Poland
* Correspondence: anna.michalak@pwr.edu.pl

Abstract

Bearings are among the most failure-prone components in rotating systems, making early fault detection crucial in industrial applications. While recent publications have focused on this issue, challenges remain, particularly in dealing with heavy-tailed or non-cyclic impulsive noise in recorded signals. Such noise poses significant challenges for classical fault selectors like kurtosis-based methods. Moreover, many deep-learning approaches struggle in these environments, as they often assume Gaussian or stationary noise and rely on large labeled datasets that are rarely available in practice. To address this, we propose a robust sparse non-negative matrix factorization (NMF) method based on the maximum-correntropy criterion, which is known for its robustness in the presence of heavy-tailed noise. This methodology is applied to identify fault frequency bands in the spectrogram of the signal. The effectiveness of the approach is validated using simulated fault signals under both Gaussian and heavy-tailed noise conditions through Monte Carlo simulations. A statistical efficiency analysis confirms robustness to random perturbations. Additionally, three real datasets are used to evaluate the performance of the proposed method. Results from both simulations and real-world data demonstrate the effectiveness of the proposed approach.

Keywords: bearing fault detection; heavy-tailed noise; non-cyclic impulsive noise; robust sparse non-negative matrix factorization; maximum-correntropy criterion; fault frequency band; Monte Carlo simulation

check for updates

Academic Editor: Hossam A. Gabbar

Received: 10 October 2025
Revised: 3 November 2025
Accepted: 11 November 2025
Published: 18 November 2025

Citation: Shiri, H.; Michalak, A. Robust Sparse Non-Negative Matrix Factorization for Identifying Signals of Interest in Bearing Fault Detection. *Sensors* **2025**, *25*, 7041. https://doi.org/10.3390/s25227041

Copyright: © 2025 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rotating machines play a crucial role in industries, and early fault detection in components such as bearings and gears is essential to prevent catastrophic failures and economic losses. In recent years, numerous studies have been published on this topic, with a particular focus on identifying impulsive behavior in signals caused by faults in rotating systems. However, this impulsive behavior is often obscured by noise, leading many researchers to develop methods for filtering noise to reveal these impulsive components better.

The well-established approach for informative band selection (IFB) is the spectral kurtosis, which offers the advantage of detecting impulsive frequency bands associated with faults while filtering out irrelevant noise [1–4]. Spectral kurtosis is effective in identifying frequency bands where the signal presents impulsive behavior, which is often associated with faults. However, it has limitations in environments with heavy-tailed impulsive noise

or when the artifacts are present in the signal. In those cases, the kurtosis may misinterpret random impulsiveness as fault-related features. In response to these limitations, several alternative approaches have been proposed. Modified sparsity measures such as L-kurtosis, the Gini index, and spectral smoothness have been introduced to improve robustness and allow for detecting the damage also under noisy or repetitive conditions [5,6]. Conditional variance (CV), originally grounded in the statistical interpretation of the 20-60-20 rule, has also proven effective in this context, particularly when implemented in a seven-subset framework tailored for non-Gaussian, impulsive environments [7,8]. Other notable contributions include the use of dependency measures, such as Pearson, Spearman, and Kendall correlations [9], as well as the application of the stability parameter from the α-stable distribution, introduced by Żak et al. [10], both offering alternative paths for identifying frequency bands carrying diagnostically relevant information in complex noise conditions.

Another powerful technique is cyclostationary analysis, which exploits the periodicity inherent in many rotating machine faults [11–15]. This method is particularly useful for detecting cyclic impulsive patterns that occur at specific rotational frequencies. Cyclostationary analysis can enhance fault detection by identifying frequency components modulated by the machine's operation, but it may struggle with non-cyclic impulsive noise or signals with weak periodicity, limiting its effectiveness in certain noisy environments [16].

In addition to these, signal decomposition methods such as empirical mode decomposition (EMD) [17–19], variational mode decomposition (VMD) [20–23], and local mean decomposition (LMD) [24,25] have gained popularity for fault detection [26–28]. These methods decompose a complex signal into simpler components, allowing for the extraction of fault-related features even in the presence of noise. EMD and its variations can adaptively separate non-linear and non-stationary signals into intrinsic mode functions (IMFs), while VMD provides better resolution by isolating specific frequency bands. LMD, on the other hand, is known for its ability to capture local signal variations effectively. Although these methods have proven useful in many scenarios, R. Randall and J. Antoni [29], have demonstrated why the previously mentioned decomposition methods offer limited benefits for bearing diagnostics by highlighting issues such as end effects and mode mixing when detecting bearing faults. In addition, these decomposition methods suffer from heavy-tailed or impulsive noise.

Deep-learning approaches have achieved notable success in fault diagnosis. Architectures tailored to rotating machinery, such as MHSNet, also reach high accuracy through multi-sensor fusion and long-range dependency modeling [30], while operator-learning frameworks like the Generalized Koopman Neural Operator demonstrate efficient data-driven modeling of railway pantograph–catenary dynamics [31]. To address data scarcity, Yan et al. [32] proposed a multimodal arc-detection network based on diffusion models that is explicitly designed for limited-data railway systems, leveraging unlabeled data and audio–visual fusion to improve robustness. Nevertheless, despite these advances, deploying supervised or semi-supervised deep models in our setting still requires modality-specific training and extensive label curation. More importantly, most deep-learning architectures rely on statistical assumptions—such as approximately Gaussian, stationary, and independently distributed noise—that are often violated by heavy-tailed and non-cyclic impulsive noise commonly observed in vibration and acoustic measurements. In many industrial systems, such as crushers or belt conveyors, these impulsive or outlier events are inherent to the mechanical process itself rather than removable anomalies, making conventional outlier suppression inappropriate and potentially destructive to informative features. Such conditions lead to unstable feature extraction, biased gradient updates, and poor generalization in deep-learning models.

Non-negative matrix factorization (NMF) and non-negative tensor factorization (NTF) have been widely applied to fault detection, particularly using spectrogram representations as inputs [33–37]. The spectrogram, which provides a time-frequency representation of the signal, is highly advantageous in fault detection because it captures both the frequency and temporal characteristics of faults, making it ideal for analyzing rotating machinery signals. By using the spectrogram as input, NMF and NTF can decompose the signal into distinct frequency bands and time components, offering a clearer understanding of the underlying fault-related structures. M. Gabor et al. [38] developed stochastic sampled orthogonal non-negative matrix factorization (SS-ONMF) to extract frequency-based features from the spectrogram of measured signals. Their proposed algorithm identifies a selective filter specifically tailored to the frequency band of the SOI. The results presented emphasize the importance of orthogonality constraints in producing highly selective filters that effectively capture only the informative frequency band. A. Michalak et al. [39] investigated the impact of heavy-tailed noise on the extraction of the SOI using NMF. They conducted an empirical analysis to evaluate how different noise levels affect the effectiveness of SOI extraction with two NMF algorithms: a multiplicative algorithm that minimizes the Euclidean objective (Standard NMF) and a generalized version of the hierarchical alternating least squares algorithm (β-HALS) for minimizing Beta-divergence. Their findings from simulations and real data analysis demonstrated that the β-HALS NMF algorithm outperforms standard NMF in terms of extraction efficiency.

However, all the mentioned algorithms, which are typically based on least-squares cost functions or divergence-based approaches, are sensitive to heavy-tailed noise and outliers, which can distort the decomposition results and lead to inaccurate fault detection.

To overcome this limitation, we propose a procedure using the robust sparse NMF method, based on the MCC, which offers enhanced robustness against outliers and heavy-tailed noise compared to traditional least-squares-based approaches. MCC leverages similarity measures rooted in higher-order statistics, making it more effective in handling heavy-tailed noise, as it downplays large deviations that are typical in such noise distributions. Additionally, the inclusion of a sparsity constraint focuses the decomposition on the most relevant components, allowing for a more precise extraction of key features. This is particularly advantageous for detecting impulsive signals, as the sparsity encourages the model to isolate sharp, localized events, thereby improving the identification of meaningful patterns hidden in noisy data. Additionally, NNDSVD [40] initialization is used in our proposed method, providing an accurate starting point compared to random initialization, which leads to faster convergence and improved decomposition accuracy. This ensures that the algorithm captures key components more reliably, further enhancing its ability to detect impulsive signals, especially in noisy environments. Additionally, the procedure utilized the Envelope Spectrum Indicator (ENVSI) to identify which filters from the NMF family effectively capture the fault. In summary, the contributions of this work can be summarized as follows:

1. **Robust cost function:** In the literature, various NMF frameworks and robust cost functions have been explored for bearing condition monitoring. In this work, we further adapt the NMF framework to this purpose by introducing the **Maximum Correntropy Criterion (MCC)**, which significantly improves robustness against heavy-tailed and impulsive noise compared with classical least-squares- and β-divergence-based NMF methods.
2. **Sparsity constraint:** An ℓ_1-norm sparsity term is incorporated into the NMF formulation to emphasize localized impulsive components that characterize bearing defects, enhancing the clarity and separability of cyclic features.

3. **Stable initialization:** The **NNDSVD initialization** is applied to both W and H matrices to ensure stable and reproducible decompositions and to accelerate convergence of the algorithm.
4. **Automatic filter selection:** The **Envelope Spectrum Indicator (ENVSI)** is employed to automatically identify the most informative NMF component corresponding to the fault-related frequency band, ensuring objective and repeatable filter selection.
5. **Comparative validation:** The proposed **Sparse NMF-MCC** is benchmarked against the **classic NMF** and **β-HALS NMF** algorithms, demonstrating that the MCC-based approach consistently achieves higher ENVSI values and clearer fault-frequency localization across both simulated and real-world datasets (belt conveyor, ore crusher, and test rig).

The paper is organized as follows: After the introduction, Section 2 presents the critical components of the proposed approach in theory. Then, in Section 3, the suggested model is simulated, and the results are presented. The results of applying the proposed approach to three benchmark data sets are presented in Section 4, along with an indication of all intermediate steps. Finally, the conclusions are formed in Section 5.

2. Methodology and Theory

The proposed methodology's framework is illustrated in Figure 1. The process starts with the input of the raw signal, followed by the configuration of key parameters: the number of filters or rank of decomposition (r), the maximum correntropy bandwidth (σ_k), the sparsity coefficient (λ), and the number of iterations. Next, a spectrogram of the raw signal is computed and decomposed using the proposed Sparse NMF-MCC method, resulting in the basis matrix **W** and the coefficient matrix **H**.

Each basis vector in **W** is then used to filter the raw signal, producing a set of candidate filtered signals. To automate the selection of the most informative filter, the ENVSI is computed for each filtered signal. The filtered signal with the highest ENVSI value is selected as the final output, as it most clearly emphasizes the fault-related spectral components. Each step of the proposed methodology is described in detail in the following subsections.

A complete list of symbols and parameters used in the formulas is provided in Appendix B (Table A1).

2.1. Short-Time Fourier Transform (STFT)

The short-time Fourier transform (STFT) of a one-dimensional discrete data **y** is defined as [41]:

$$\mathbf{STFT}_{t,f} = \sum_{n=0}^{L-1} y[n] w[n-t] e^{-\frac{j2\pi fn}{N}}, \quad (1)$$

where $w[n-t]$ is a window function centered around t used to localize the signal in time, L is the number of samples in the analyzed segment of the signal (window length), N is the number of points used for the Discrete Fourier Transform, f is the frequency, and j is the imaginary unit, with $j^2 = -1$.

The spectrogram **S** is the absolute value of **STFT**, forming a non-negative matrix that represents the time-frequency distribution of a signal. The window function $w[n-t]$ determines the resolution trade-off between time and frequency: a narrow window yields better time resolution but poorer frequency resolution, while a wide window offers better frequency resolution but poorer time resolution.

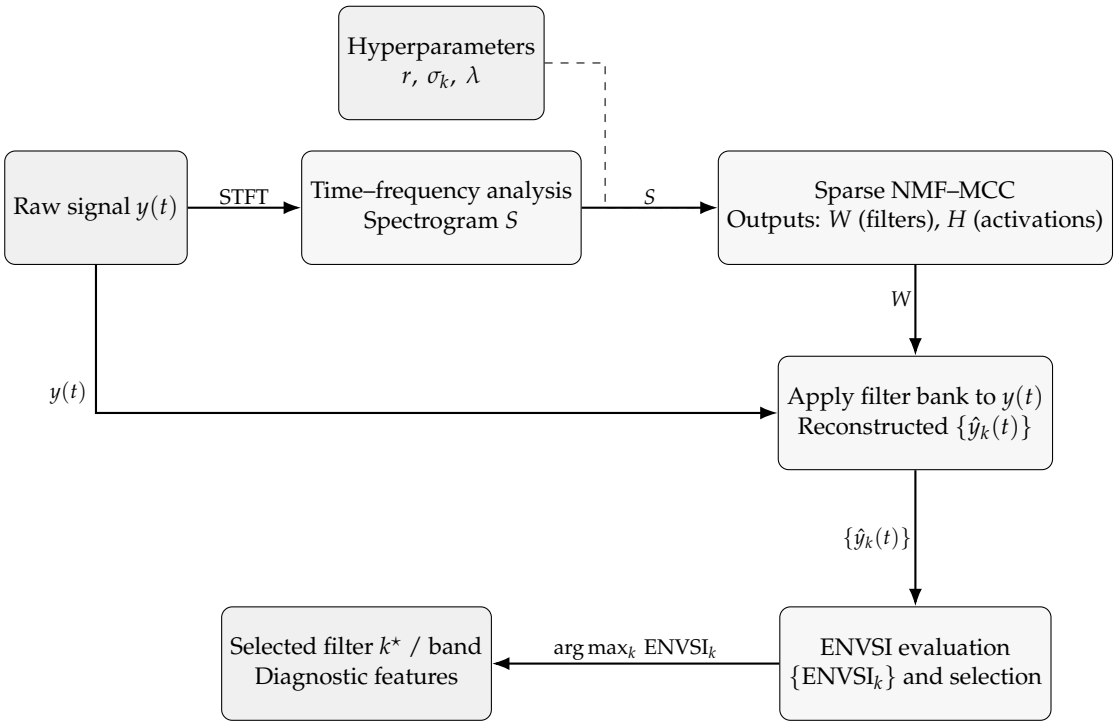

Figure 1. Framework of the proposed Sparse NMF–MCC pipeline. Inputs: raw signal $y(t)$ and hyperparameters (r, σ_k, λ). The spectrogram S is factorized into basis W and activations H. Filters derived from W are applied to $y(t)$ to obtain reconstructed signals $\{\hat{y}_k(t)\}$, which are scored by ENVSI; the component with the highest ENVSI$_k$ is selected for diagnostics.

2.2. NMF Model

Non-negative Matrix Factorization (NMF) is a technique used to decompose a non-negative data matrix **S** into a product of two lower-rank non-negative matrices: **W** (the basis matrix) and **H** (the coefficient matrix). In general, the NMF model can be expressed as follows [42–45]:

$$\mathbf{S} = \mathbf{WH} + \mathbf{E}, \tag{2}$$

where $\mathbf{S} \in \mathbb{R}_+^{m \times n}$ is the input matrix, $\mathbf{W} \in \mathbb{R}_+^{m \times r}$ (the basis matrix) and $\mathbf{H} \in \mathbb{R}_+^{r \times n}$ (the coefficient matrix), $\mathbf{E} \in \mathbb{R}^{m \times n}$ is the error matrix and r is the rank of the factorization. In the proposed approach, the input matrix with non-negative entries is the spectrogram matrix. One of the resulting matrices, the base matrix **W**, can be used as a filter characteristic. This allows the filtering of the signal and enables the identification of the faults.

2.3. Maximum Correntropy Criterion (MCC) with Gaussian Kernel

Let $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ and $\mathbf{y} = \{y_1, y_2, \ldots, y_N\}$ represent two sets of random variables. The correntropy between **x** and **y** is defined as [46,47]:

$$V(\mathbf{x}, \mathbf{y}) = \mathbb{E}\big[\kappa_{\sigma_k}(x - y)\big], \tag{3}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator and $\kappa_{\sigma_k}(\cdot)$ is a Gaussian kernel function given by:

$$\kappa_{\sigma_k}(x - y) = \exp\left(-\frac{(x - y)^2}{2\sigma_k^2}\right), \tag{4}$$

with $\sigma_k > 0$ as the kernel bandwidth parameter.

The Gaussian kernel in MCC has two important properties that make it an effective tool in analyzing data with noise:

- **Robustness**: The Gaussian kernel suppresses large deviations by assigning lower values to them. For large deviations, the value of $\exp\left(-\frac{(x-y)^2}{2\sigma_k^2}\right)$ approaches zero, making corentropy less sensitive to outlier points than traditional measures such as mean squared error, which enhances the robustness of MCC against outliers.
- **Bandwidth Control**: The parameter σ_k dictates the width of the Gaussian kernel, balancing between sensitivity to small errors and robustness against large ones.

In the classical NMF method, the data matrix $\mathbf{S}$ is approximated by the product of two non-negative matrices $\mathbf{W}$ and $\mathbf{H}$, minimizing the distance between $\mathbf{S}$ and $\mathbf{WH}$, usually using Euclidean, Kullback-Leibler, or Itakura-Saito divergence. In optimization problems, MCC maximizes the correntropy between the original signal and the output one. In the MCC-NMF approach, the cost function is replaced by a measure of correntropy [48]. It can be written as follows:

$$\max_{W,H} V(\mathbf{S}, \mathbf{WH}) = \mathbb{E}\left[\exp\left(-\frac{(\mathbf{S}-\mathbf{WH})^2}{2\sigma_k^2}\right)\right]. \tag{5}$$

2.4. Sparse NMF with MCC Objective

Sparse NMF is an extension of the NMF method that introduces additional sparsity constraints during matrix decomposition [49]. The proposed approach uses the most popular method for introducing sparsity in NMF, i.e., adding a penalty term in the l_1 norm to the cost function.

$$\max_{W,H\geq 0} \sum_{i=1}^{m}\sum_{j=1}^{n} \exp\left(\frac{(\mathbf{S}_{ij}-[\mathbf{WH}]_{ij})^2}{2\sigma_k^2}\right) - \lambda_H \|\mathbf{H}\|_1, \tag{6}$$

where $\sigma_k > 0$ represents the bandwidth of the Gaussian kernel used in the MCC, $||\cdot||_1$ means the l_1 norm, λ_H is the regularization parameter controlling the sparsity of $\mathbf{H}$. The term $\exp\left(\frac{(\mathbf{S}_{ij}-[\mathbf{WH}]_{ij})^2}{2\sigma_k^2}\right)$ represents the Gaussian kernel applied to the residuals, while the regularization term $\lambda_H\|\mathbf{H}\|_1$ promotes sparsity in $\mathbf{H}$, encouraging solutions with fewer active components.

2.5. Algorithm of Sparse NMF MCC

The MCC weight $\mathbf{\Omega}$ is calculated based on the residual $\mathbf{E}_{ij} = \mathbf{S}_{ij} - \mathbf{(WH)}_{ij}$ and is given by:

$$\mathbf{\Omega}_{ij} = \exp\left(-\frac{(\mathbf{S}_{ij}-[\mathbf{WH}]_{ij})^2}{2\sigma_k^2}\right) = \exp\left(-\frac{\mathbf{E}^2}{2\sigma_k^2}\right). \tag{7}$$

This weight $\mathbf{\Omega}$ is used to modify the update rules of $\mathbf{W}$ and $\mathbf{H}$. They help adjust the importance of individual approximation errors, down-weight significant reconstruction errors, and provide robustness to outliers.

The algorithm implements Sparse NMF MCC using multiplicative update rules for $\mathbf{W}$ and $\mathbf{H}$. It consists of three main stages, namely initialization, update rules, and stopping condition.

- **Initialization:** Initialize the matrices $\mathbf{W}$ and $\mathbf{H}$ with the non-negative values. Set the parameters σ_k and λ. In presented case both matrices, i.e., $\mathbf{W}$ and $\mathbf{H}$ were initialize using NNDSVD.
- **Update Rules:** The update for $\mathbf{H}$ incorporates the MCC weights:

$$\mathbf{H} \leftarrow \mathbf{H} \odot \frac{\mathbf{W}^T(\mathbf{\Omega} \odot (\mathbf{S} \oslash (\mathbf{WH}+\epsilon)))}{1+\lambda}, \tag{8}$$

where $\mathbf{W}^T$ is the transposition of matrix $\mathbf{W}$, $\mathbf{\Omega}$ is the matrix of MCC weights, $\odot$ is the Hadamard product, $\oslash$ is the Hadamard division and ϵ is a small constant to prevent division by zero. The update for $\mathbf{W}$ is:

$$\mathbf{W} \leftarrow \frac{\mathbf{W} \odot (\Omega \odot (\mathbf{S} \oslash (\mathbf{WH} + \epsilon)))\mathbf{H}^T}{||H^T||_1} \tag{9}$$

Finally, the normalization of $\mathbf{W}$ is done by the columns of matrix $\mathbf{W}$:

$$\mathbf{W} \leftarrow \frac{\mathbf{W}}{||\mathbf{W}||_1}.$$

In this paper, two versions of the Sparse NMF MCC are discussed. In the first version, it is assumed that the kernel size (bandwidth) is a constant parameter. In the second version, the kernel size is calculated using a following formula [48,50]:

$$\sigma_k^2 = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (\mathbf{S}_{ij} - [\mathbf{WH}]_{ij})^2 \tag{10}$$

- **Stopping Condition:** The algorithm continues until a maximum number of iterations is reached or the optimality gap becomes sufficiently small.

Additionally, Algorithm 1 provides the pseudocode detailing the proposed methodology for reference.

The described Sparse NMF MCC is particularly effective in environments with heavy-tailed noise or outliers, which are commonly found in real-world applications, such as vibration signal analysis. By applying the Sparse NMF with MCC to the spectrogram matrix $\mathbf{S}$ of a vibration signal, the algorithm can isolate specific frequency components related to cyclic impulses while simultaneously suppressing non-cyclic noise and other irrelevant signal components. The sparsity constraint in $\mathbf{H}$ helps focus on the impulsive nature of the signal, making this method ideal for detecting cyclic impulsive signals in noisy environments.

For clarity and consistency, the list of all variables and parameters appearing in the formulas is provided in Appendix B.

2.6. Selecting the Parameters

Another important aspect of the proposed model is the selection of hyperparameters, such as the MCC bandwidth (σ_k), sparsity constraint (λ), and the number of filters (NMF rank r), all of which significantly influence the final results. The bandwidth of MCC (σ_k) plays a crucial role in balancing sensitivity and robustness—smaller σ_k values increase robustness by reducing the influence of outliers and heavy-tailed noise, while larger σ_k values make the method more sensitive to noise. The sparsity constraint (λ) dictates the emphasis placed on sparsity; a higher λ results in a more focused decomposition on key components, which is beneficial for detecting impulsive signals. However, excessive sparsity may overlook important details. The rank of NMF (r) determines the number of filters or components; a higher r provides more detail but may lead to overfitting, while a lower r simplifies the decomposition but risks missing nuances.

Balancing these hyperparameters is key to optimizing the performance of Sparse NMF MCC. The bandwidth σ_k controls noise management, λ regulates the sparsity level, and r defines the resolution of the decomposition—each influencing the quality of cyclic impulsive signal detection. In this work, these parameters were selected through trial and error, following the described rules. Still, there is considerable potential to apply constrained optimization algorithms for choosing the optimal values. The ENVSI (defined

in Section 2.7) could serve as the output metric to be maximized, with σ_k, λ, and r as input parameters to be optimized for the best performance. It is also worth mentioning that the effectiveness of MCC bandwidth σ_k has been investigated in several previous studies [51,52].

Algorithm 1: Sparse NMF MCC: Sparse Non-negative Matrix Factorization with Maximum Correntropy Criterion

Input: $\mathbf{S} \in \mathbb{R}_+^{m \times n}$ (non-negative input matrix),
r (rank of decomposition),
σ_k (initial Gaussian kernel bandwidth),
λ (sparsity regularization parameter),
ϵ (small constant to avoid division by zero),
maxIter (maximum number of iterations),
type (bandwidth selection mode: fixed or adaptive)
Output: $\mathbf{W} \in \mathbb{R}_+^{m \times r}$ (basis matrix),
$\mathbf{H} \in \mathbb{R}_+^{r \times n}$ (coefficient matrix)

Initialization: Initialize $\mathbf{W}^{(0)}$, $\mathbf{H}^{(0)}$ with non-negative values (e.g., NNDSVD);
Set parameters σ_k, λ, ϵ;
for $l = 0$ **to** *maxIter-1* **do**
  Compute residual: $\mathbf{E}^{(l)} \leftarrow \mathbf{S} - \mathbf{W}^{(l)}\mathbf{H}^{(l)}$;
  Compute MCC weight matrix:

$$\mathbf{\Omega}^{(l)} \leftarrow \exp\left(-\frac{(\mathbf{E}^{(l)})^2}{2\sigma_k^2}\right)$$

  Update $\mathbf{H}$ using multiplicative rule:

$$\mathbf{H}^{(l+1)} \leftarrow \mathbf{H}^{(l)} \odot \frac{(\mathbf{W}^{(l)})^T\left(\mathbf{\Omega}^{(l)} \odot \left(\mathbf{S} \oslash (\mathbf{W}^{(l)}\mathbf{H}^{(l)} + \epsilon)\right)\right)}{1 + \lambda}$$

  Update $\mathbf{W}$ using multiplicative rule:

$$\mathbf{W}^{(l+1)} \leftarrow \frac{\mathbf{W}^{(l)} \odot \left(\left(\mathbf{\Omega}^{(l)} \odot \left(\mathbf{S} \oslash (\mathbf{W}^{(l)}\mathbf{H}^{(l+1)} + \epsilon)\right)\right)(\mathbf{H}^{(l+1)})^T\right)}{\|(\mathbf{H}^{(l+1)})^T\|_1}$$

  Normalize columns of $\mathbf{W}^{(l+1)}$:

$$\mathbf{W}^{(l+1)} \leftarrow \frac{\mathbf{W}^{(l+1)}}{\|\mathbf{W}^{(l+1)}\|_1}$$

  if *type == 2* **then**
    Update bandwidth using:

$$\sigma_k^2 \leftarrow \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\mathbf{S}_{ij} - [\mathbf{W}^{(l+1)}\mathbf{H}^{(l+1)}]_{ij}\right)^2$$

  end
  if *convergence criterion met* **then**
    break
  end
end

2.7. Envelope Spectrum Indicator

The Envelope Spectrum Indicator, called ENVSI, is suitable due to its ability to emphasize fault-related spectral components in a clear and quantifiable manner. It compares the energy of the components corresponding to the fault to the energy of the whole signal at a given frequency range. In the proposed work, it has been utilized in two ways: to select the NMF class associated with the damage signature, and to compare the performance of the proposed methods with that of the benchmark techniques. It takes values between 0 and 1, resulting in simplicity of interpretation. A higher ENVSI value indicates a clearer representation of the fault frequency and its harmonics, confirming the relevance of the extracted features. In contrast, a low ENVSI value means that the method does not effectively isolate the informative frequency band, i.e., the damage cannot be detected. The ENVSI can be defined as follows [8]:

$$\text{ENVSI} = \frac{\sum_{i=1}^{H} A_{\text{fault},i}^2}{\sum_{j=1}^{N} \text{SES}_j}, \tag{11}$$

where $A_{\text{fault},i}$ denotes the amplitude of the envelope spectrum at the fault frequency and its corresponding harmonics, while H refers to the number of harmonics included in the analysis. In the presented analysis, the H is equal to 5. The denominator, N, represents the total number of frequency bins used to compute the overall energy in the squared envelope spectrum (SES).

3. Simulation

This section outlines the procedure for generating synthetic signals under various conditions, including Gaussian noise with different SNR levels and heavy-tailed noise with varying intensities. The proposed method is then applied to these signals to evaluate its robustness and effectiveness under diverse noise environments.

To benchmark the performance of the proposed approach, several established NMF techniques were employed for comparison. In particular, the Classic NMF method was implemented using the `nnmf` function available in MATLAB 2024a [53]. Additionally, the β-HALS NMF algorithm, as presented in [39], was considered.

3.1. Simulated Signal

In the following, the equations for the generated synthetic signal are presented. The observed signal, y, in a vibration-based fault detection scenario, can be modeled as the superposition of two components:

$$\mathbf{y} = \mathbf{s} + \mathbf{n}, \tag{12}$$

where **y** represents the total simulated signal, **s** denotes the SOI component, and **n** describes the noise. The noise **n** is classified into two categories: zero-mean white Gaussian noise and heavy-tailed noise. For the heavy-tailed distribution, the Student's t-distribution has been selected.

The Student's t distribution is a fundamental statistical tool to incorporate an essential parameter: degrees of freedom ν. This distribution provides a robust framework for modeling data, accommodating deviations from normality. Widely utilized in statistical inference, hypothesis testing, and modeling scenarios involving small sample sizes or unknown population standard deviation, Student's t distribution enhances the accuracy and reliability of statistical conclusions. The probability density function of the Student's t distribution is expressed as follows:

$$f(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \tag{13}$$

In this expression, x represents the random variable, and ν denotes the degrees of freedom. For a visual representation of the impact of the degrees of freedom parameter on the probability density function of the Student's t distribution, refer to Figure 2.

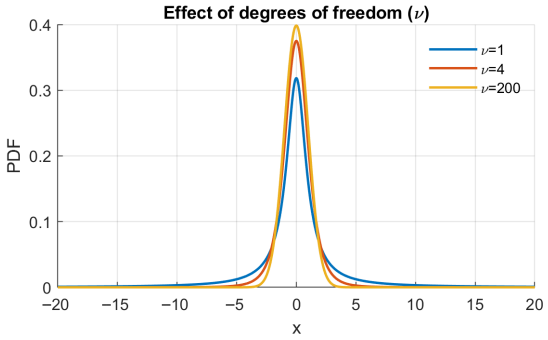

Figure 2. The effect of degree of freedom ν on Student t distribution.

The SOI, **s**, is characterized as a periodic impulse signal. It is modeled using a damped sinusoidal function that repeats periodically over time. The mathematical representation of **s** is given as:

$$s(t) = A \sum_{m=-\infty}^{\infty} \exp\{-\alpha_s t\} \sin(2\pi f_c t + \phi_s) * \delta(t - m(T_p + \Delta T_p(t))). \tag{14}$$

In this formulation, A denotes the magnitude of the sinusoidal component, while f_c specifies the frequency. The damping factor α_s indicates the rate at which the amplitude decays over time. The function $\delta(\cdot)$ is the Dirac delta function, which localizes the impulses. The term ϕ_s represents the initial phase shift. The convolution operation is indicated by the symbol $*$. The nominal period of the signal is denoted as T_p, and $\Delta T_p(t)$ describes the slow variations in the period due to jitter effects.

3.2. Signal in the Presence of Gaussian Noise

Figure 3 illustrates a synthetic signal contaminated with Gaussian noise at three different levels. The SOI is centered at 2500 Hz and includes a periodic impulsive component.

The top row presents the raw signals. In the first subplot, the noise level is relatively low ($\sigma = 0.3$), allowing the impulsive component to be partially visible despite a modest SNR. In the middle and right subplots, the noise levels increase to $\sigma = 1.4$ and $\sigma = 2.7$, respectively, resulting in increasingly challenging SNRs where the impulsive features become difficult to distinguish.

The middle row shows the corresponding spectrograms, computed using MATLAB's `spectrogram` function with parameters detailed in Table 1. For the low-noise case ($\sigma = 0.3$), the impulsive components around the 2500 Hz center frequency are clearly isolated in the time-frequency domain. However, as the noise level increases, the spectrograms become dominated by noise, and the periodic components are no longer discernible—highlighting the limitations of conventional spectral methods under low SNR conditions.

The bottom row presents the squared envelope spectrum (SES) for each case. For the low-noise case, clear periodic peaks appear in the SES, and the ENVSI value is relatively high, indicating strong spectral sparsity and the presence of modulation. Conversely, for $\sigma = 1.4$ and $\sigma = 2.7$, the SES becomes flatter and less structured, resulting in much lower ENVSI values. This underscores the need for more advanced signal processing methods to enhance and extract impulsive features under high-noise conditions.

Table 1. Parameters used for the spectrogram generation.

Parameter	Value
Window	`hamming(256)`
Overlap	180
FFT Points	512
Sampling Frequency (Fs)	25,000

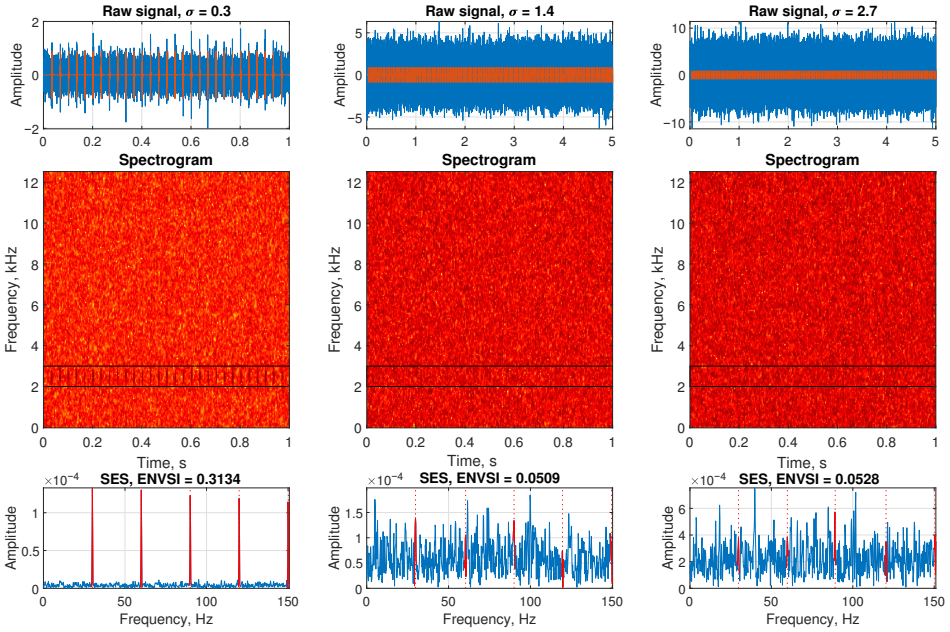

Figure 3. Simulated signals in the presence of Gaussian noise, spectrograms, and SES for three different σ parameters, respectively.

The Figure 4 is presented based on an exemplary signal contaminated by Gaussian noise with standard deviation $\sigma = 1.4$, corresponding to a low signal-to-noise ratio of SNR = -21.678 dB. The aim is to compare the performance of different NMF-based filtering techniques in extracting the impulsive component from this noisy signal.

In the first row, the the filter spectral profiles of the extracted components are shown. The first subplot illustrates the frequency response of the Sparse NMF MCC method, where the band-pass component around the frequency of interest (2500 Hz) is clearly visible. The second subplot shows the result of the Sparse NMF MCC with modified bandwidth, which still preserves good selectivity around the main spectral region. In contrast, the third subplot (Classic NMF) shows multiple wideband and less structured components, and in the last subplot the β-HALS NMF method detects the correct bandwidth; however, the amplitude of the baseline is relatively high, indicating that this type of filter tends to retain a significant amount of background noise.

In the second row, the corresponding filtered signals are presented in the time domain. The third row shows the SES and the corresponding ENVSI value calculated for each method. For the Sparse NMF MCC method and Sparse NMF MCC with modified bandwidth, the periodic peaks in SES are clearly visible, and ENVSI reaches values of 0.09303 and 0.0938, respectively. On the other hand, Classic NMF and β-HALS NMF methods fail to highlight the periodic structure, resulting in lower ENVSI values of 0.0648 and 0.0689. These results demonstrate that under low SNR conditions, Sparse NMF MCC-based methods provide superior filtering capability and improved impulsive feature extraction.

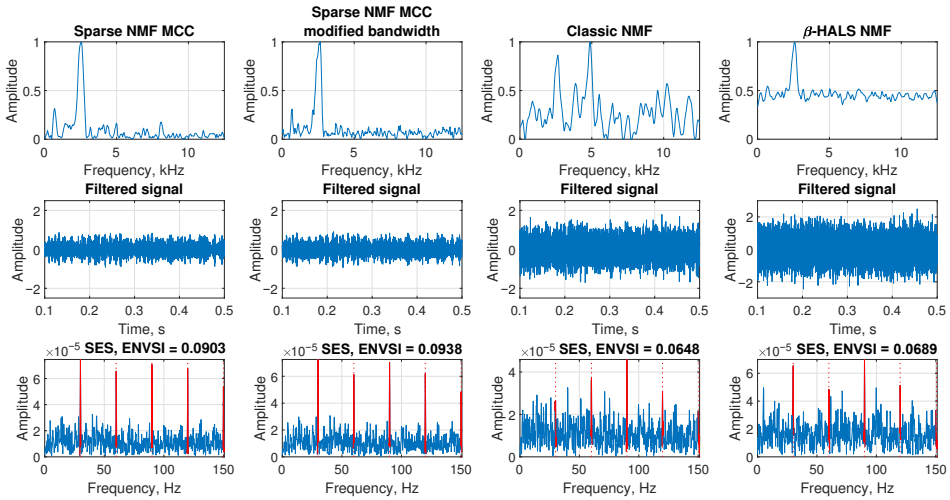

Figure 4. The comparison of the results obtained from the Sparse NMF MCC, Sparse NMF MCC with the modified bandwidth, classic NMF, and β-HALS NMF, respectively. The analysis is presented from the exemplary signal with Gaussian noise (σ = 1.4, SNR = −21.678).

The results shown in Figure 5 were calculated using the same signal as in Figure 4. This allows for a comparison of the outcomes obtained by the NMF-based approaches. The results are presented in the same scheme.

In the first row, the spectral kurtosis, Alpha selector, CV selector, and Spearman selector are presented, respectively. Unfortunately, none of these selectors provides the correct informative frequency band, making it impossible to determine whether the signal contains damage. The second row presents the corresponding filtered signals in the time domain, which also do not convey any information about the damage. The third row illustrates the SES and the corresponding ENVSI values calculated for each method. The peaks corresponding to the fault frequency and their harmonics are not visible, and the ENVSI values are low.

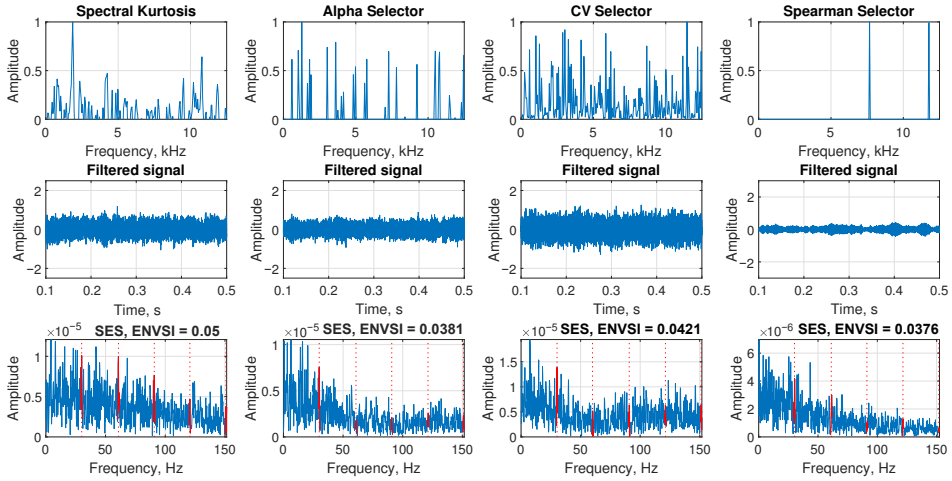

Figure 5. The comparison of the results obtained from the spectral kurtosis, Alpha selector, CV selector, and Spearman selector, respectively. The analysis is presented from the exemplary signal with Gaussian noise (σ = 1.4, SNR = −21.678).

3.3. Signal in the Presence of Heavy-Tailed Noise

The Figure 6 is presented to investigate the effect of heavy-tailed noise on the signal characteristics. In this case, the raw signal is contaminated with impulsive noise modeled

using a Student's t-distribution with varying degrees of freedom $\nu \in \{3, 5, 10\}$. The parameter ν controls the tail behavior of the distribution: lower values correspond to heavier tails and more extreme noise outliers, while higher values approach Gaussian behavior.

In the first row, the raw signals are shown. As we can see, when $\nu = 3$, the signal is heavily corrupted by high-amplitude fluctuations due to the impulsive nature of the noise. As ν increases to 5 and then to 10, the impulsiveness of the noise decreases, and the amplitude range becomes more consistent with Gaussian-like behavior. However, in all three cases, the impulsive components of interest are difficult to observe directly from the time domain due to the presence of strong noise.

The second row presents the spectrograms calculated for each signal using MATLAB's `spectrogram` function with the same parameters described in Table 1. In all three subplots, the time-frequency representation is dominated by broadband noise, and there is no clear evidence of the periodic impulsive structure. This highlights the challenge of detecting fault-related features under heavy-tailed, non-Gaussian noise conditions.

In the third row, the SES is calculated for each case, and the corresponding ENVSI values are also shown. As expected, for $\nu = 3$, the SES is highly irregular and noisy, with no significant peaks, and the ENVSI value is low (0.0522). Similarly, for $\nu = 5$ and $\nu = 10$, the ENVSI values remain low (0.0514 and 0.068, respectively), and the periodic structure is not captured. These results confirm that in the presence of heavy-tailed impulsive noise, traditional energy-based spectral features such as SES and ENVSI are ineffective for detecting weak, periodic fault signatures, motivating the need for more robust and adaptive signal processing techniques.

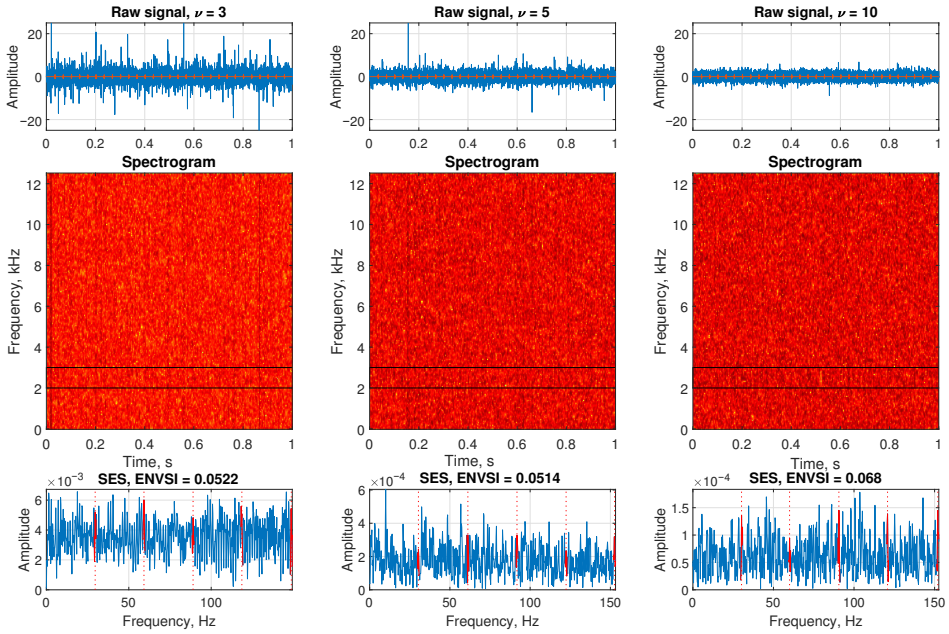

Figure 6. Simulated signals in the presence of heavy-tailed noise (Student's t with three different ν parameter), spectrograms and SES.

The Figure 7 presents the performance comparison of different NMF-based filtering methods when applied to a signal contaminated with heavy-tailed noise modeled by a Student's t-distribution with degrees of freedom $\nu = 5$. This case represents a strong impulsive noise environment that challenges traditional filtering approaches.

In the first row, the filter spectral profiles of the extracted components are shown. For the Sparse NMF MCC method, the frequency band around the target component is well identified. Similarly, in the second subplot, the modified bandwidth version of Sparse NMF

MCC also captures the relevant band, although it exhibits a broader shape. The Classic NMF method (third subplot) produces the chaotic filter spectral profile. In the last subplot, the β-HALS NMF detects a narrow region of interest, but with a high baseline amplitude, which suggests that it retains a considerable amount of background noise.

The second row shows the time-domain filtered signals. In both Sparse NMF MCC methods, impulsive features are partially preserved despite the high noise level, and the background noise is relatively reduced. However, Classic NMF and β-HALS NMF fail to suppress the heavy-tailed noise sufficiently, making it difficult to distinguish impulsive components from the output signal.

In the third row, the SES and the ENVSI values are calculated. The highest ENVSI is obtained using Sparse NMF MCC (0.2042), followed closely by its modified bandwidth version (0.1841), both showing distinguishable periodic peaks. On the contrary, Classic NMF and β-HALS NMF result in noisy SES plots with weak or invisible periodic structure, and their ENVSI values drop to 0.0614 and 0.053, respectively. These findings suggest that under the influence of heavy-tailed noise, Sparse NMF MCC methods are more robust in isolating modulated fault-related components.

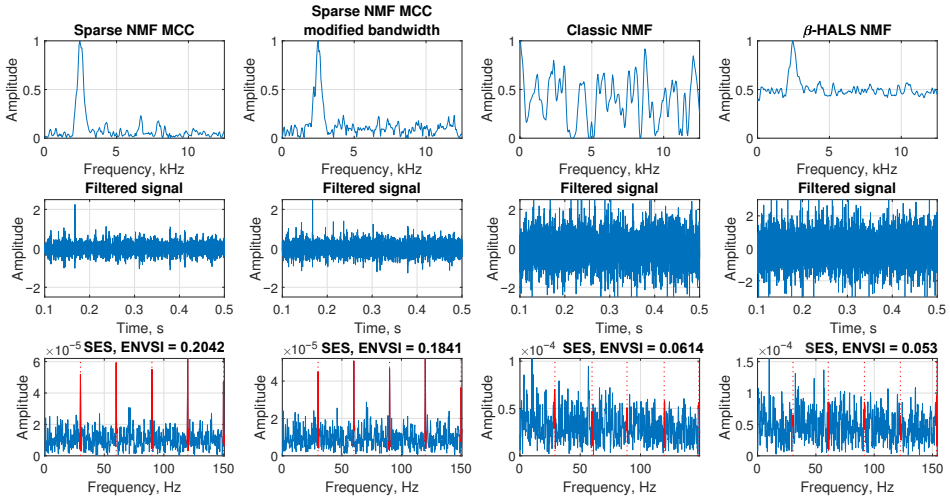

Figure 7. The comparison of the results obtained from the Sparse NMF MCC, Sparse NMF MCC with the modified bandwidth, classic NMF, and β-HALS NMF, respectively. The analysis are presented from the exemplary signal with heavy-tailed noise ($\nu = 5$).

The results presented in Figure 8 were obtained using the same signal as in Figure 7. This enables a comparison of the outcomes achieved by the NMF-based approaches. In the first three subplots, which correspond to spectral kurtosis, the Alpha selector, and the CV selector, the filters obtained appear random and do not indicate any damage. However, the Spearman selector shows a partially correct band. The second row displays the time-domain filtered signals, while the third row presents the SES and ENVSI values. In the first three subplots, where the filters do not provide the correct IFB, the SES are noisy, and the periodic component cannot be visible. The ENVSI values are low. In contrast, the SES corresponds to the Spearman selector has a high noise level; however, the periodic impulses corresponding to the fault frequency and its harmonics, marked in red, are visible. This results in the ENVSI value (0.1455) being the highest among all the compared selectors. However, it is still lower than the one obtained based on the Sparse NMF MCC and its modified bandwidth version.

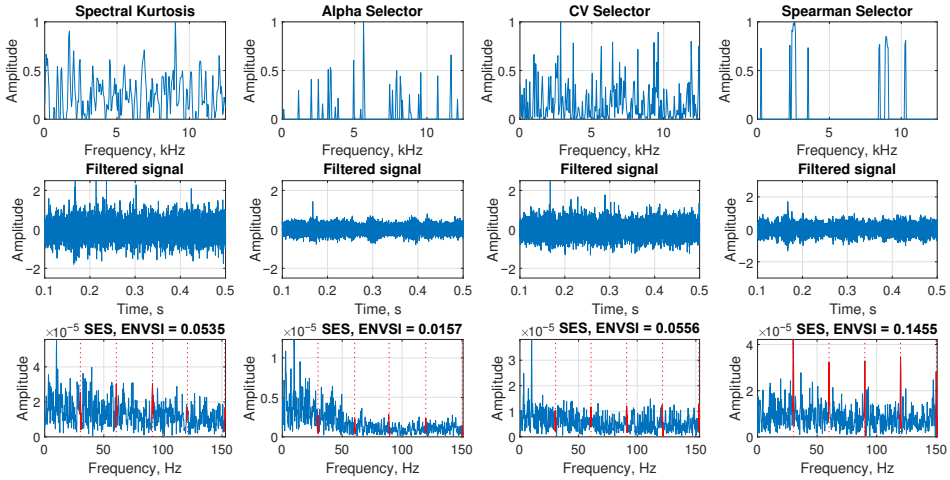

Figure 8. The comparison of the results obtained from the spectral kurtosis, Alpha selector, CV selector, and Spearman selector, respectively. The analysis is presented from the exemplary signal with heavy-tailed noise ($\nu = 5$).

3.4. Monte-Carlo Analysis for Simulated Signal

In this section, a Monte Carlo analysis is carried out to evaluate the robustness and repeatability of the proposed filtering approaches under various noise conditions. By performing multiple independent simulations with different realizations of noise, both Gaussian and heavy-tailed (Student's *t*-distributed), we aim to assess the consistency of frequency localization and fault feature extraction across different NMF-based methods.

3.5. Gaussian Case

The Figure 9 presents a Monte Carlo analysis conducted over 50 independent trials to assess the robustness and consistency of the proposed filtering methods under different levels of Gaussian noise. Three different values of the noise standard deviation are considered: $\sigma = 0.3$, $\sigma = 1.2$, and $\sigma = 2.0$, corresponding to SNR levels of -8.2993 dB, -20.3439 dB, and -24.7757 dB, respectively. The objective is to evaluate how each method behaves across repeated runs and whether it can consistently extract the impulsive component under varying noise conditions.

3.6. Non-Gaussian Case

Each row corresponds to a different noise level. The columns compare the four filtering approaches: Sparse NMF MCC, Sparse NMF MCC with modified bandwidth, Classic NMF, and β-HALS NMF. The vertical axis shows the normalized amplitude of the extracted components, plotted across frequency (horizontal axis) and Monte Carlo iterations (depth axis).

In the top row ($\sigma = 0.3$, high SNR), all methods can localize the frequency component of interest. The Sparse NMF MCC and its modified version show sharp and consistent frequency localization. The Classic NMF shows more variation, and β-HALS NMF produces a flatter distribution with lower amplitude (the background level is slightly higher).

In the middle row ($\sigma = 1.2$), the increased noise degrades the performance of all methods; however, the Sparse NMF MCC variants still preserve the dominant frequency structure across most runs. The classic NMF exhibits more scattering across frequencies. The β-HALS NMF allows for detecting the proper frequency; however, the background of the noise also increased significantly.

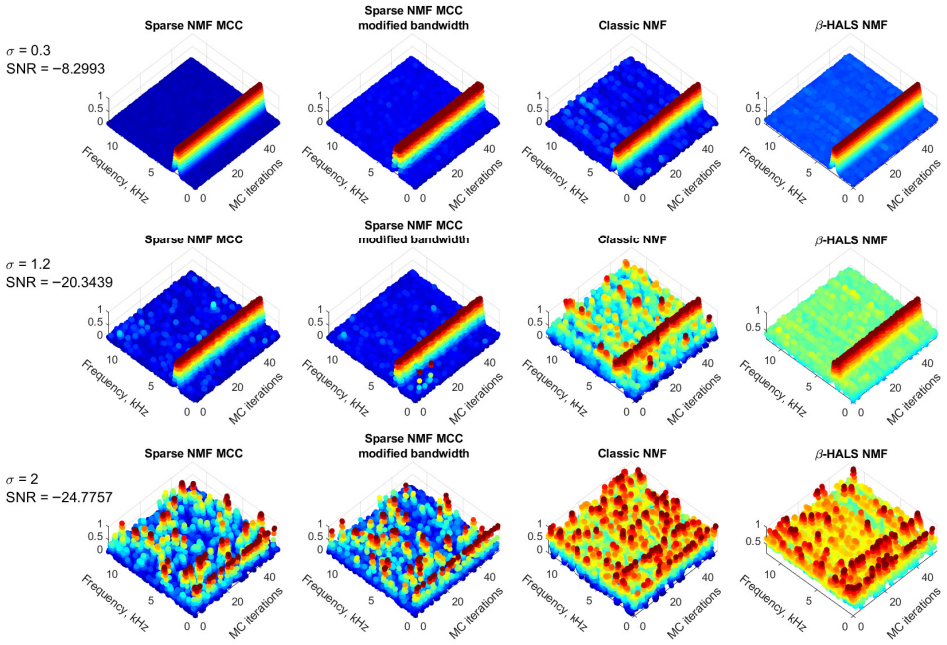

Figure 9. 50 Simulated signals in the presence of a different range of Gaussian noise.

In the bottom row ($\sigma = 2.0$), corresponding to the lowest SNR, all methods struggle to extract meaningful components due to the extremely noisy environment. In the Sparse NMF MCC method, the frequency localization is unclear or inconsistent across iterations. Similarly, both Classic NMF and β-HALS NMF result in highly scattered and noisy outputs with no dominant spectral focus. These results indicate that, although the proposed Sparse NMF MCC methods are relatively more robust, all methods experience performance degradation under very low SNR conditions.

The Figure 10 presents the Monte Carlo analysis for 100 independent iterations under heavy-tailed noise modeled by the Student's t-distribution with different degrees of freedom: $\nu = 3, 5, 10$. As ν increases, the tail behavior becomes lighter and the distribution approaches Gaussian. This analysis is designed to evaluate the robustness and consistency of various NMF-based filtering methods in the presence of increasingly impulsive noise.

Each row corresponds to a different value of ν, representing various levels of impulsiveness. The columns show the four filtering techniques: Sparse NMF MCC, Sparse NMF MCC with modified bandwidth, Classic NMF, and β-HALS NMF. The vertical axis shows the normalized spectral amplitude, while the horizontal axes represent frequency and Monte Carlo iterations.

In the top row ($\nu = 3$), the noise is strongly heavy-tailed (highly impulsive). All methods struggle to maintain frequency localization, and the spectral responses are scattered across iterations. Although Sparse NMF MCC and its modified variant show slightly better concentration compared to Classic NMF and β-HALS NMF, none of the methods achieve stable or reliable frequency tracking under such conditions.

In the middle row ($\nu = 5$), the impulsiveness is moderate. The Sparse NMF MCC methods begin to show more consistent localization around the frequency of interest, while Classic NMF displays scattered outputs and poor reproducibility. Although the β-HALS NMF shows a correct and consistent bandwidth, the noise level is significantly increased, which substantially degrades its performance.

In the bottom row ($\nu = 10$), the noise distribution is much closer to Gaussian. The results show more precise and consistent frequency patterns for all methods; however, the

performance of Sparse NMF MCC remains the most robust, with concentrated frequency tracking across all iterations. These findings confirm that while all methods benefit from reduced impulsiveness, the proposed Sparse NMF MCC framework offers the best trade-off between robustness and consistency under heavy-tailed noise conditions.

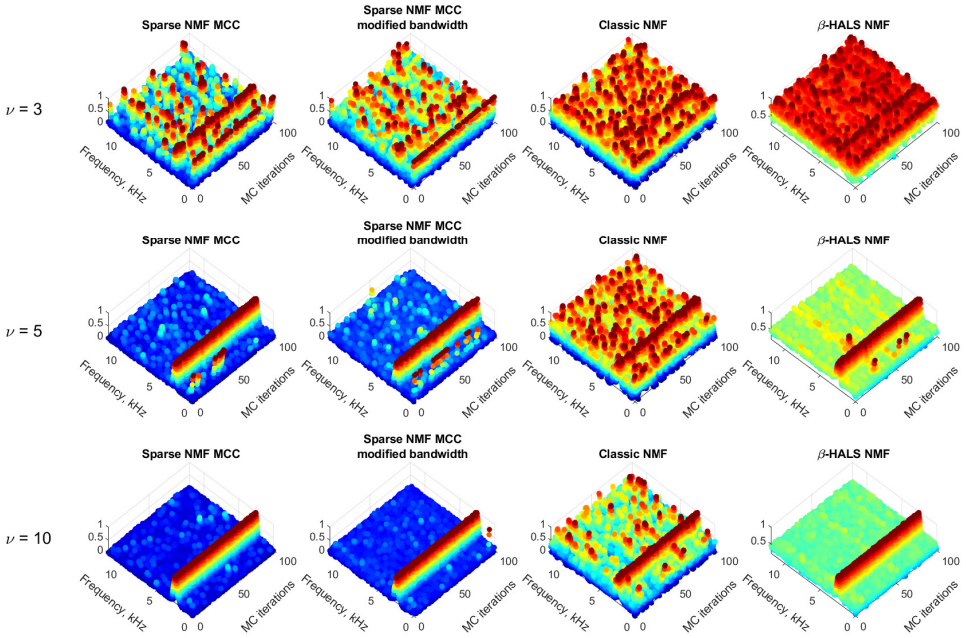

Figure 10. 100 Simulated signals in the presence of a different range of heavy-tailed noise (Student's t).

The Monte Carlo simulations indicate that the Sparse NMF MCC method and those with the modified bandwidth, offers consistent and proper results under both Gaussian and impulsive noise conditions. It achieves better frequency localization and lower variability of the background level compared to the other tested methods.

In addition, the efficiency calculations have been carried out. Simulations were conducted for the 36 equally spaced values of the parameter ν, ranging from 3 to 10.

To determine the threshold for distinguishing whether a given signal comes from the healthy machine, healthy signals were simulated values (36 × 150 signals), and the ENVSI were computed (denoted $ENVSI_H$). A threshold value is defined as $1.25 \times \max(ENVSI_H)$.

In Figure 11, an efficiency value of 1 indicates that the fault was correctly detected in 100% of simulations. The performance of the proposed methods, Sparse NMF MCC and Sparse NMF MCC with modified bandwidth, was compared to Classic NMF and β-HALS NMF. Although β-HALS NMF is designed for non-Gaussian signals, it exhibits a high background level, which results in low ENVSI values. To compare, the selected base feature in β-HALS NMF was normalized by subtracting its minimum value.

3.7. Sensitivity Analysis

The Figure 12 presents the results of the sensitivity analysis of the NMF-based algorithm with respect to its two model parameters: λ and σ_k. The algorithm's performance is assessed using the ENVSI. As can be seen, 19 σ_k values from the range between 0 and 2 and 14 λ parameters from the range between 0 and 1 were taken into account. As in the previous results, each column corresponds to test signals generated with a different parameter ν, namely 3, 5, and 10. The upper row of plots shows the median ENVSI values obtained from 30 Monte Carlo simulations, while the lower row illustrates the corresponding standard deviations. For $\nu = 3$, where the simulated signals exhibit heavy tails and high variability,

the algorithm performs inconsistently, and ENVSI values remain low regardless of parameter choice. In general, larger values of λ lead to improved algorithm efficiency. Moreover, the σ_k in the range between 0.3 and 0.7 allows for obtaining the most promising results. This demonstrates that properly tuned λ and σ_k significantly enhance the robustness of the NMF-based method, especially when the input data deviates from Gaussian behavior.

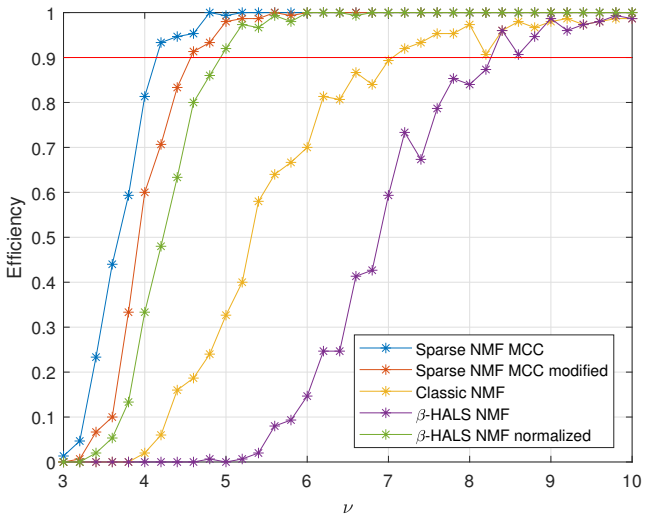

Figure 11. Efficiency calculation. The red line indicates an efficiency level of 90% correct results.

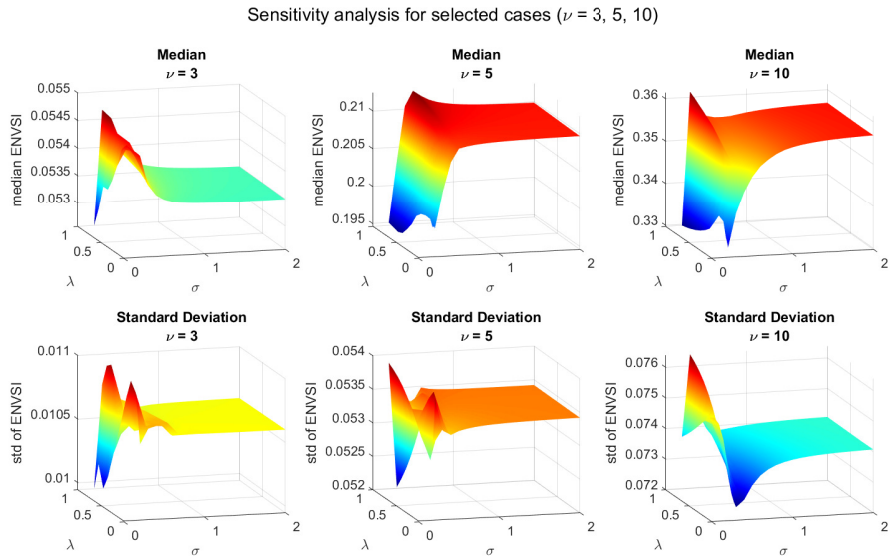

Figure 12. Parameter sensitivity of the NMF-based method under a different range of heavy-tailed noise (Student's t).

4. Real Data Analysis

To validate the proposed method under real-world conditions, three distinct case studies are presented. These include: (1) acoustic monitoring of belt conveyor idlers in an operational industrial environment, (2) vibration analysis of a copper ore crusher subjected to heavy-duty loads, and (3) controlled fault diagnosis using a laboratory-scale test rig. Each case offers a unique combination of signal characteristics, fault types, and measurement setups, providing a comprehensive assessment of the robustness and adaptability of the proposed approach to both acoustic and vibration signals under practical constraints.

4.1. Case 1: Belt Conveyor Acoustic Signal

Idlers are rotating components of belt conveyors designed to support the moving belt (see Figure 13a). Each idler consists of a shaft, two bearings, and a protective coating. Given the large number of idlers in operation, a fast, contactless method to evaluate the condition of their bearings is crucial. To address this need, an experiment was conducted on a fully operational conveyor system. Due to the practical limitations, in-depth visual inspection was not possible, but the characteristics of the signal indicated the presence of localized damage.

For the experiment, audio recordings were taken from each idler using a standard smartphone (see Figure 13b). Each recording lasted approximately 10 s and had a sampling frequency of 48 kHz. The sound was extracted from video recordings using MATLAB's built-in function. The captured signals exhibited characteristics indicative of faults, making them suitable for further analysis.



(**a**) (**b**)

Figure 13. Conveyor belt and signal acquisition setup. (**a**) Conveyor belt system. (**b**) Acquiring acoustic signal from conveyor belt.

4.2. Case 2: Copper Ore Crusher

An ore crusher used in the mining industry (see Figure 14) serves as a case study to demonstrate the advantages of the proposed approach. The crusher operates under variable load and speed conditions due to time-varying load fluctuations imposed by the unknown volume and size of the material pieces that impose resistance on the machine. A vibration signal was measured on the casing of the rolling element bearing that supports the crusher's main shaft. The signal was recorded using Endevco accelerometers, model 751-10, (New York, United States) and a data acquisition system managed through LabVIEW Signal Express 2013 software. Vibration signals were collected during normal operation, and an example was selected in which the crusher was fully loaded, including oversized pieces of material. The parameters of the signal are as follows: a sampling frequency of 25 kHz and a total duration of approximately 2 min, although only a 10-s segment is analyzed in this paper.

In the industry, it is challenging to obtain representative signals from the machine with a faulty bearing. The special case of the crusher underscores a critical engineering issue: its impulsive loading characteristics mean that even minimal damage can quickly escalate, potentially causing catastrophic system failure. To ensure a controlled experiment, data from a healthy bearing was used as a baseline, while a cyclic impulsive signal (representing the SOI) was added to simulate a local fault. This approach allows for testing various fault severities, including scenarios where the SOI is either fully obscured by background noise or just barely visible in the raw signal but remains difficult to detect. The fault frequencies for the bearing (SKF 23264) are listed in Table 2, and the analysis focuses on an inner race fault within one of the bearings.



Figure 14. Copper ore crusher machine.

Table 2. Characteristic frequencies of 23264 CCK/W33 bearing.

Description	Value
Shaft speed frequency	3 Hz
Cage defect frequency (FTF)	1.3 Hz
Ball spin frequency (BSF)	10.6 Hz
Inner race defect frequency (BPFI)	30.7 Hz
Outer race defect frequency (BPFO)	23.3 Hz
Rolling element defect frequency	21.1 Hz

4.3. Case 3: Test Rig

The experimental data were acquired using a dedicated test rig consisting of an electric motor, gearbox, couplings, and two bearings, as shown in Figure 15. To simulate realistic fault conditions, one of the bearings was intentionally damaged. Acoustic signals were recorded at a sampling rate of 50 kHz using a Brüel & Kjær 4189 microphone (Nærum, Denmark) along with a Kistler LabAmp 5165 (Sindelfingen, Germany) A data acquisition system. Throughout the experiment, the rotational speed was maintained at a constant 1041 rpm. The bearings used were SKF 1205 EKTN9 models, with one exhibiting a localized defect on the outer race. The characteristic frequencies associated with the bearing components are summarized in Table 3, where the outer race defect frequency is highlighted at 91.11 Hz.

Table 3. Characteristic frequencies of 1205 EKTN9 bearing.

Description	Value
Shaft speed frequency	17.35 Hz
Cage defect frequency (FTF)	7 Hz
Ball spin frequency (BSF)	42.75 Hz
Inner race defect frequency (BPFI)	134.44 Hz
Outer race defect frequency (BPFO)	91.11 Hz
Rolling element defect frequency	85.5 Hz



Figure 15. Test rig used in the experiment.

4.4. Results for Real Dataset

In the following section, preliminary analyses using well-established approaches are presented for the three real-world case studies shown in Figure 16. The Figure 16 illustrates the baseline signal analysis for the three real-world case studies introduced earlier. Each column corresponds to one case: Case 1 (belt conveyor), Case 2 (copper ore crusher), and Case 3 (test rig with seeded fault). The figure is structured into three rows to present a comparative overview of the time-domain signals, spectrograms, SES.

In the first row, the raw acoustic or vibration signals are shown for each case. Across all three cases, the impulsive component is not clearly distinguishable in the time domain due to the influence of background noise and operating variability. In particular, Case 2 and Case 3 exhibit relatively stronger amplitude fluctuations, likely resulting from load or structural resonance effects.

The second row presents the spectrograms of the signals, computed using MATLAB's `spectrogram` function with parameters appropriate for each sampling rate. While energy concentration across frequency bands can be observed, there is no visible evidence of repetitive impulsive signatures or modulation bands in any of the cases. This highlights the difficulty of detecting bearing-related faults in raw measurements without pre-processing or enhancement, especially under operational noise.

The third row shows the SES along with the calculated ENVSI values. For Case 1, several periodic peaks are weakly visible and the ENVSI reaches a moderate value of 0.101, indicating a low level of modulation sparsity. In Case 2, the SES is almost flat and noisy, with an ENVSI of just 0.0227, making fault-related frequency components indistinct. Case 3 shows even lower sparsity (ENVSI = 0.0061), and no clear periodic components are observed. These results confirm that raw signal energy analysis and conventional envelope techniques alone are insufficient for reliable fault diagnosis in complex industrial and experimental environments.

The Figures 17, 19, and 21 present the performance of four NMF-based filtering methods applied to the signals from Case 1, Case 2, and Case 3, respectively. Each column corresponds to one method: Sparse NMF MCC, Sparse NMF MCC with modified bandwidth, Classic NMF, and β-HALS NMF. Similarly, the Figures 18, 20, and 22 present the performance of four selectors selected for comparison applied to the signals from Case 1, Case 2, and Case 3, respectively. Each column corresponds to one selector: spectral kurtosis, alpha selector, CV selector, and Pearson selector. The figures are organized into three rows that visualize the selected informative frequency band, the filtered time-domain signal, and the resulting SES with the ENVSI.

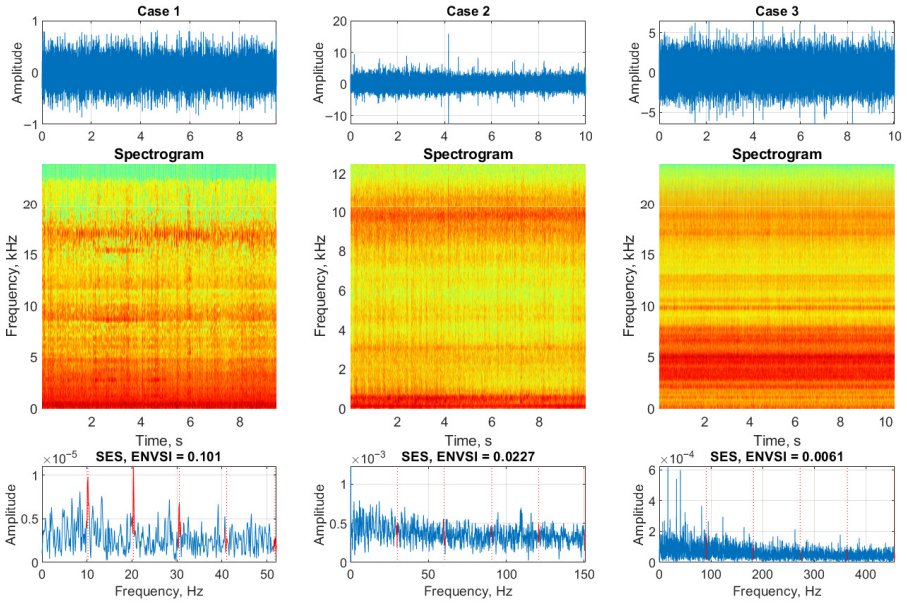

Figure 16. The time series, spectrogram, and SES for Case 1, Case 2, and Case 3, respectively.

The Figure 17 presents the performance of four NMF-based filtering methods applied to the acoustic signal from Case 1 (belt conveyor idler). In the first row, the extracted frequency component is shown. All methods identify a band around 15–18 kHz. The modified bandwidth version presents the narrowband selectivity. In contrast, Classic NMF and β-HALS NMF exhibit sharper peaks but with higher values in low-frequency bands, which results in higher noise levels in the filtered signals. The Sparse NMF MCC has lower selectivity in the 15–18 kHz range; however, it obtains values around zero in the low-frequency band. It gives a lower noise level in the filtered signal and a higher ENVSI value.

The second row shows the corresponding time-domain filtered signals. All methods recover an amplitude-modulated impulsive structure, though with varying levels of noise suppression. Sparse NMF MCC results in a cleaner impulsive pattern, while the other three methods produce noisier outputs with lower peak-to-background contrast.

The third row displays the SES for each filtered signal, along with the corresponding ENVSI value. All four methods successfully reveal a periodic pattern in the frequency domain. Sparse NMF MCC achieves the highest ENVSI (0.3111), followed by Classic NMF (0.305), β-HALS NMF (0.2899), and Sparse NMF MCC (mod. BW) (0.2887). Although the ENVSI values are relatively close, the SES generated by Sparse NMF MCC provides clearer periodic peaks. These results demonstrate that all methods are capable of detecting the impulsive modulation associated with the fault, but the proposed Sparse NMF MCC offers a slight edge in terms of clarity and sparsity.

The Figure 18 presents the results of four selectors calculated for the signal for Case 1. In the first row, the shapes of the obtained selectors are presented. The second row shows the resulting time-domain signals after filtering. The third row presents the SES of the filtered signals along with the corresponding ENVSI values. The alpha selector and Pearson selector present the highest ENVSI value equal to 0.2991 and 0.291, respectively. Note that it is still a lower value than the ENVSI value calculated for the signal filtered by the proposed Sparse NMF MCC.

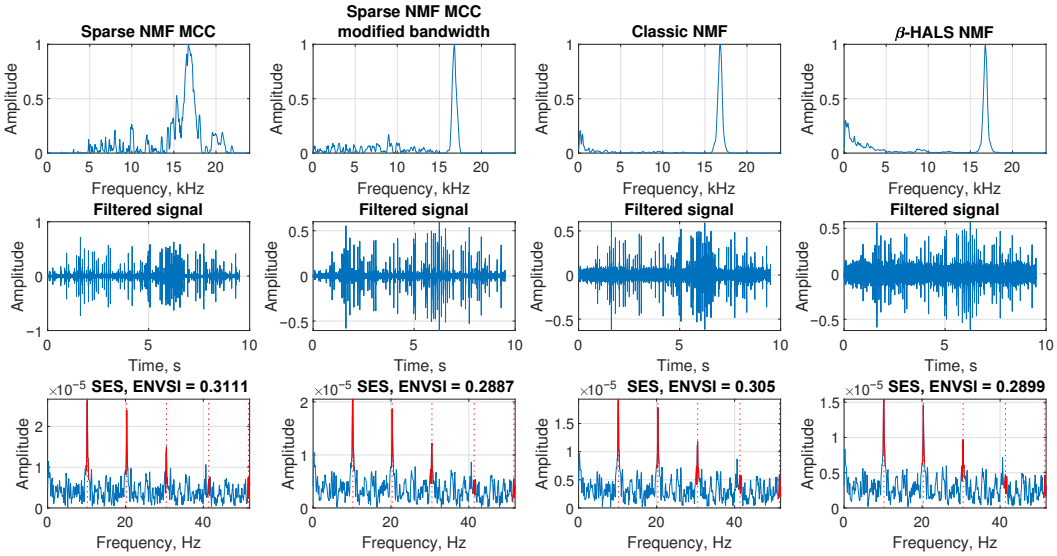

Figure 17. Filtering results for Case 1. Comparison of NMF-based methods using frequency response, filtered signal, and SES with ENVSI values.

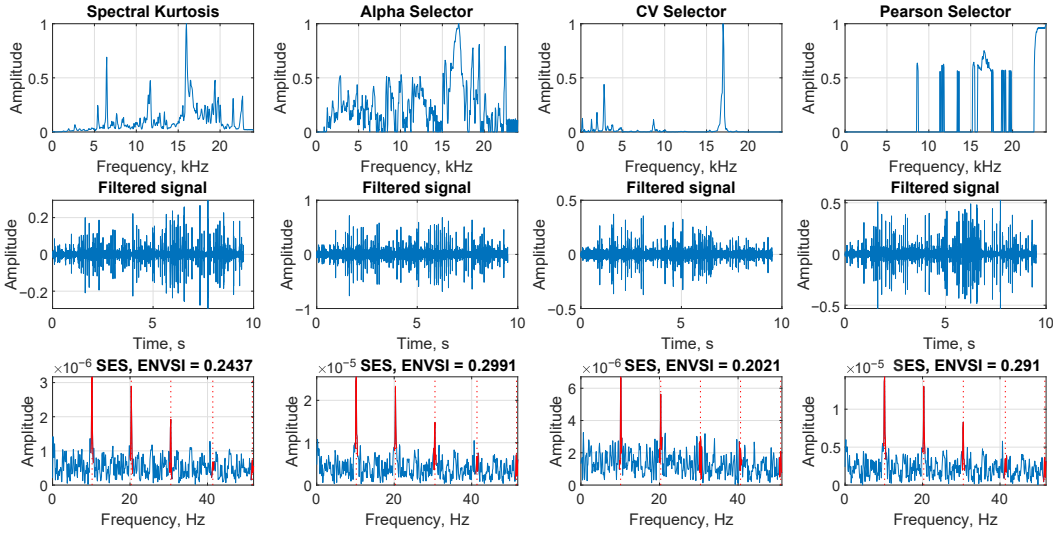

Figure 18. Filtering results for Case 1. Comparison of four selectors, filtered signal, and SES with ENVSI values.

The Figure 19 presents the results of applying four NMF-based filtering methods to the vibration signal collected from the copper ore crusher (Case 2). In the first row, the filter characteristics of the extracted components are displayed. Classic NMF and β-HALS NMF identify the informative band as the one in which non-cyclic impulses occur. In the case of Sparse NMF MCC with modified bandwidth, the most informative band is indicated as being below 1 kHz. Finally, the Sparse NMF MCC indicates the band in which cyclic impulses associated with damage occur. Despite the appearance of values in higher bands, it is the only one that correctly shows the information band.

The second row shows the time-domain signals after filtering. The impulsive content remains buried in noise, especially in Classic NMF and β-HALS NMF, where the filtered signals exhibit large amplitude spikes that do not follow a periodic pattern. Sparse NMF MCC method suppresses background variation more effectively, though the recovered signal is still far from clean, it allows damage to be detected.

In the third row, the SES plots for each method are shown. Sparse NMF MCC achieves the highest ENVSI value (0.1258), with faint periodic peaks corresponding to fault-related frequencies. The other three methods produce noisy SES plots with little distinguishable

structure and low ENVSI values: 0.0201 for Sparse NMF MCC (mod. BW), 0.021 for Classic NMF, and 0.0207 for β-HALS NMF. These results demonstrate that while the environment is too noisy for reliable extraction using conventional methods, the proposed Sparse NMF MCC still provides the most promising outcome in terms of modulation sparsity and fault feature enhancement.

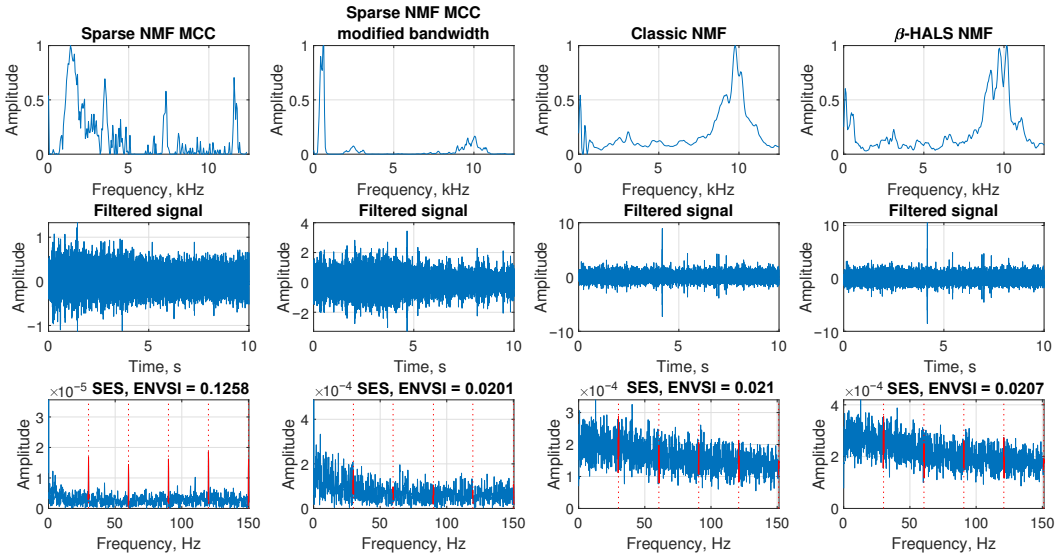

Figure 19. Filtering results for Case 2. Comparison of NMF-based methods using frequency response, filtered signal, and SES with ENVSI values.

The Figure 20 presents the results of four selectors calculated for the signal for Case 2. As can be observed in the first row, none of the presented selectors cannot select the proper informative frequency band. Both spectral kurtosis and methods dedicated to impulsive noise data were unable to indicate the IFB correctly. As a result, both the filtered signals and the SES failed to represent the fault component effectively, and the ENVSI values obtained are very low.

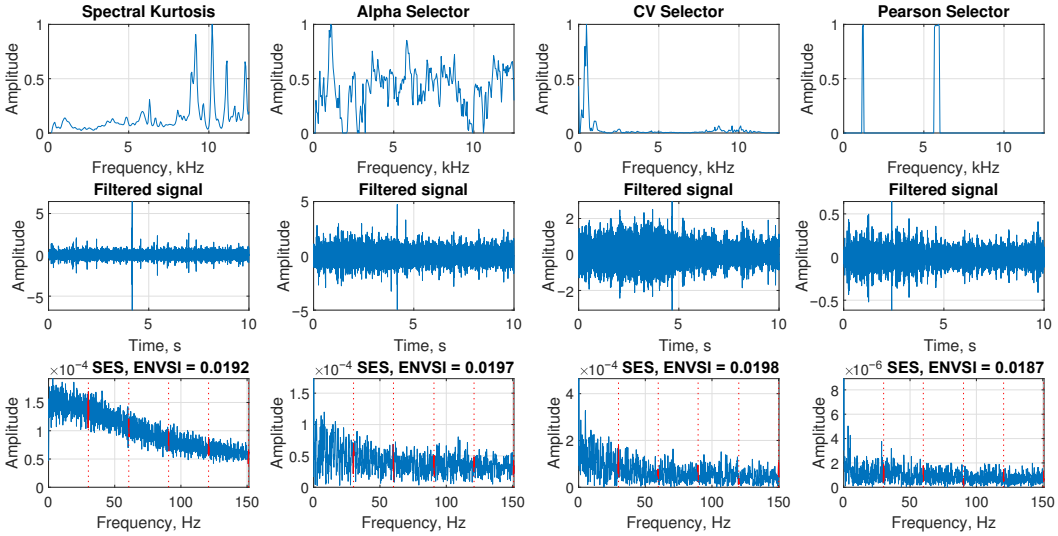

Figure 20. Filtering results for Case 2. Comparison of four selectors, filtered signal, and SES with ENVSI values.

The Figure 21 presents the filtering results for Case 3, which involves acoustic measurements from a test rig containing a bearing with an outer race defect. In the first row, the spectral responses extracted by each method are shown. Both Sparse NMF MCC methods

capture a narrow and focused band around 20–23 kHz, corresponding to the fault-induced modulation. In contrast, Classic NMF and β-HALS NMF identify much broader and noisier components, resulting in ineffective separation of the impulsive content.

In the second row, the filtered signals in the time domain are visualized. Sparse NMF MCC produces a clean signal with clearly distinguishable impulsive events. The modified bandwidth version also reveals regular impulsive patterns but includes slightly higher background noise. Classic NMF and β-HALS NMF produce filtered signals that are more contaminated with noise, lacking any distinct repetitive features.

The third row shows the SES and ENVSI values. Sparse NMF MCC yields the highest ENVSI (0.4737), followed by its modified bandwidth version (0.3636), both showing sharp peaks at the fault frequency and its harmonics. Classic NMF and β-HALS NMF achieve significantly lower ENVSI values (0.0118 and 0.0087, respectively), and the periodic peaks are barely visible. These results confirm that Sparse NMF MCC is the most effective method for isolating fault-related modulations in controlled laboratory conditions.

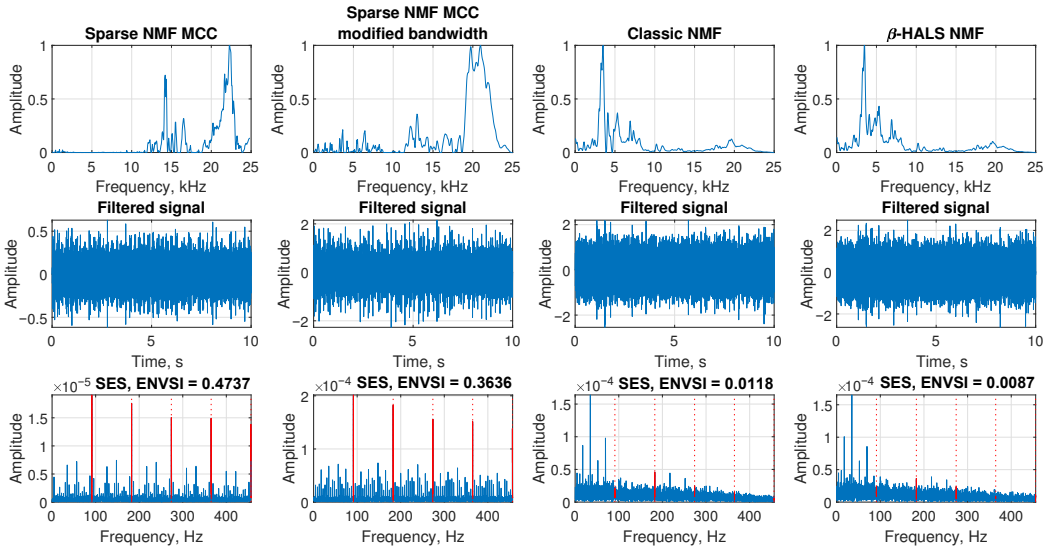

Figure 21. Filtering results for Case 3. Comparison of NMF-based methods using frequency response, filtered signal, and SES with ENVSI values.

The Figure 22 presents the results of four selectors calculated for the signal for Case 3. Spectral kurtosis, alpha selector, and CV selector do not allow the correct IFB to be obtained. They focus on other frequency bands. The Pearson selector, in contrast, accurately indicates the proper bandwidth and is highly selective. This is evident in both the filtered signal, where cyclic impulses are visible, and on the SES. The ENVSI value obtained in this case is 0.3973.

Figure 23 presents a comparative analysis of the ENVSI values obtained for all evaluated methods across three different cases. The ENVSI serves as an efficiency indicator, quantifying how well each approach enhances the visibility of fault-related components in the squared envelope spectrum.

In Case 1, all methods yield relatively similar ENVSI values, enabling the detection of damage; however, the Sparse NMF MCC and Classic NMF slightly outperform the others. In Case 2, the Sparse NMF MCC has the highest value of the ENVSI, and it is the only one that allows for the detection of damage. In Case 3, both Sparse NMF MCC and its modified version obtain a high value of ENVSI. In this case, the Pearson selector also enables the detection of damage. The results from the analysis of real data show that the proposed Sparse NMF MCC consistently achieves better diagnostic performance in all

cases, confirming its robustness and efficiency in extracting fault-related information from complex vibration signals.

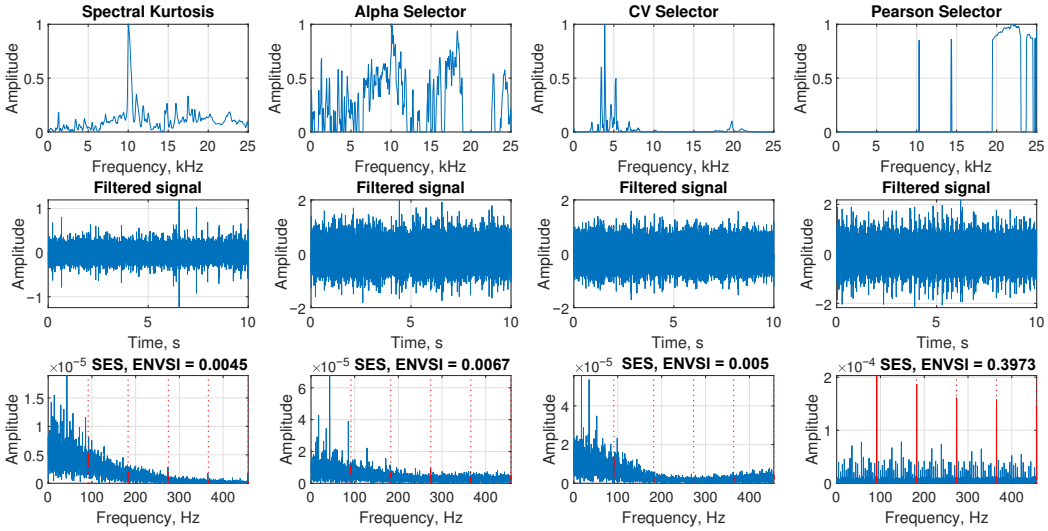

Figure 22. Filtering results for Case 3. Comparison of four selectors, filtered signal, and SES with ENVSI values.

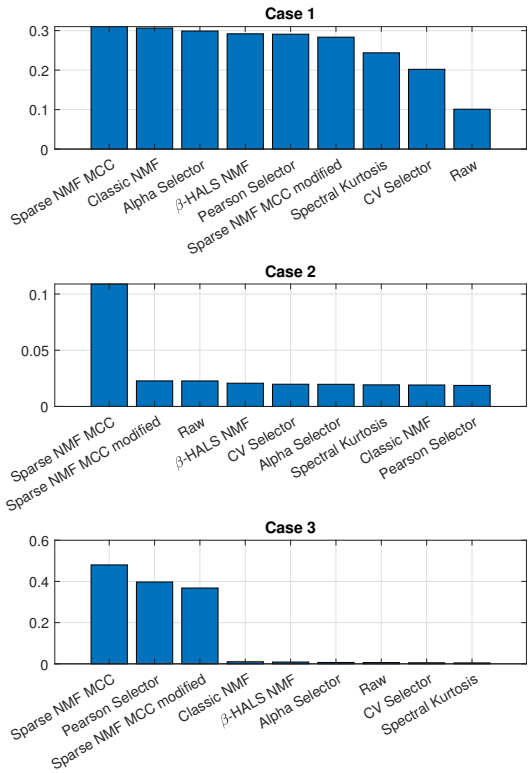

Figure 23. Comparison of the obtained ENVSI values for all analyzed methods for signals corresponding to the Case 1, Case 2, and Case 3, respectively.

Moreover, to ensure fair comparison, all NMF-based methods were tested both with their default initialization schemes (as provided in their standard implementations) and with NNDSVD initialization.

The comparative results for Cases 1, 2, and 3 are summarized in Table 4. The comparison has been made with the MATLAB implementation of the classical NMF (default) and with the random initialization in β-HALS NMF. The use of NNDSVD initialization did not

significantly affect the quality of the obtained results for Classical NMF and β-HALS (the ENVSI parameter values changed by at most 0.01). However, this significantly affects the results obtained with sparse MCC NMF for Case 2, where NNDSVD initialization enables damage detection. The performance of the contribution of each component of the proposed algorithm is presented in Appendix A.

Table 4. Comparison of NMF with different initializations and selector results for Cases 1–3, respectively. The highest value in each column has been bolded.

Method	Type	Case 1	Case 2	Case 3
NMF Sparse MCC	Random	0.2803	0.0196	0.4650
	NNDSVD	**0.3111**	**0.1258**	**0.4737**
Sparse NMF MCC	Random	0.3011	0.0228	0.3188
Modified Bandwidth	NNDSVD	0.2887	0.0201	0.3636
Classic NMF	Default	0.3051	0.0197	0.0136
	NNDSVD	0.3050	0.0210	0.0118
β-HALS NMF	Random	0.2927	0.0208	0.0088
	NNDSVD	0.2899	0.0207	0.0087
Spectral Kurtosis	–	0.2437	0.0192	0.0045
Alpha Selector	–	0.2991	0.0197	0.0067
CV Selector	–	0.2021	0.0198	0.0050
Pearson Selector	–	0.2910	0.0187	0.3973

The main hyperparameters and their empirically selected ranges for the three experimental case studies are summarized in Table 5. The NMF rank r was chosen based on the number of components to include in the signal: one class for damage, one for noise, and two for non-Gaussian noise. The sparsity coefficient and the MCC kernel bandwidth were tuned by a trial-and-error process guided by the ENVSI and visual inspection of the reconstructed components, also based on the rule in Section 2.6.

Table 5. Key hyperparameters and empirical selection ranges used in the Sparse NMF-MCC algorithm for the three experimental case studies.

Parameter	Symbol	Case 1 (Conveyor)	Case 2 (Ore Crusher)	Case 3 (Bearing Rig)
Sparsity coefficient	λ	1	1	1
MCC kernel bandwidth	σ_k	0.3	0.3	0.3
Factorization rank	r	4	4	4

4.5. Discussion, Limitations, and Future Work

The proposed Sparse NMF-MCC demonstrates strong robustness under both Gaussian and heavy-tailed noise conditions and achieves effective separation of fault-related components in both simulated and real-world datasets. However, as with any decomposition-based method, certain limitations and potential failure cases should be acknowledged. The performance may degrade when impulsive fault signatures overlap strongly with broadband background noise, when the decomposition rank is not appropriately chosen, or when fault impulses are extremely weak relative to dominant structural resonances. These conditions can lead to partial mode mixing or reduced separability among NMF components.

Moreover, the current method requires manual selection of hyperparameters such as the MCC bandwidth, sparsity coefficient, and decomposition rank, which can influence performance across different datasets. In future work, these parameters can be optimized automatically using optimization or learning-based algorithms to adapt the model to varying signal and noise characteristics. The present implementation also relies on the

STFT representation; extending it to adaptive time–frequency or tensor-based frameworks could further enhance its capability to capture complex non-stationary patterns.

While the current study intentionally focuses on the NMF family to analyze and improve its robustness under heavy-tailed and non-cyclic impulsive noise, deep-learning frameworks represent an exciting complementary direction. In our future research, we plan to integrate the proposed Sparse NMF-MCC within hybrid or deep-learning-based architectures to combine the interpretability and noise robustness of model-based decomposition with the adaptability of data-driven learning. Finally, real-time and online implementations represent promising directions for practical industrial monitoring and fault-diagnosis systems.

5. Conclusions

In this paper, we propose a robust sparse NMF method based on the maximum correntropy criterion for the automatic identification of signals of interest in bearing fault detection. The proposed approach addresses key limitations of traditional NMF-based decomposition techniques by introducing robustness against outliers and heavy-tailed noise via the correntropy measure. In addition, the inclusion of an ℓ_1-norm sparsity constraint promotes the isolation of impulsive features typically associated with mechanical faults.

Comprehensive Monte Carlo simulations under various noise conditions, including both Gaussian and heavy-tailed distributions, demonstrated the superior ability of the proposed method to extract cyclic impulsive components embedded in noisy environments. These results were further validated through real-world case studies involving conveyor belt systems, a copper ore crusher, and data from a controlled test rig. In all cases, the method successfully identified and enhanced fault-related frequency bands, even under challenging noise scenarios.

Overall, the proposed Sparse NMF MCC framework represents a significant advancement in automated fault detection for rotating machinery. Its ability to robustly extract informative components in the presence of noise makes it particularly well-suited for industrial applications where conventional techniques often fail.

Author Contributions: Conceptualization, H.S. and A.M.; methodology, H.S.; software, H.S. and A.M.; validation, H.S. and A.M.; formal analysis, H.S. and A.M.; investigation, H.S.; resources, A.M.; data curation, H.S. and A.M.; writing—original draft preparation, H.S.; writing—review and editing, H.S. and A.M.; visualization, H.S. and A.M.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The measurement data presented in this study are not publicly available due to restrictions of privacy.

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

Appendix A

The performance of the contribution of each component of the proposed algorithm are presented. The results are presented in four columns, each corresponding to the impact of a different component of the algorithm. The first column shows the results for Sparse NMF MCC with NNDSVD initialization, the second for Sparse NMF MCC with random initialization, the third for Sparse NMF with Euclidean distance, and the fourth for MCC NMF without sparsity.

In Figure A1, all methods yield the correct result, and the differences in ENVSI values are not significant; however, the proposed methods produce the highest ENVSI value.

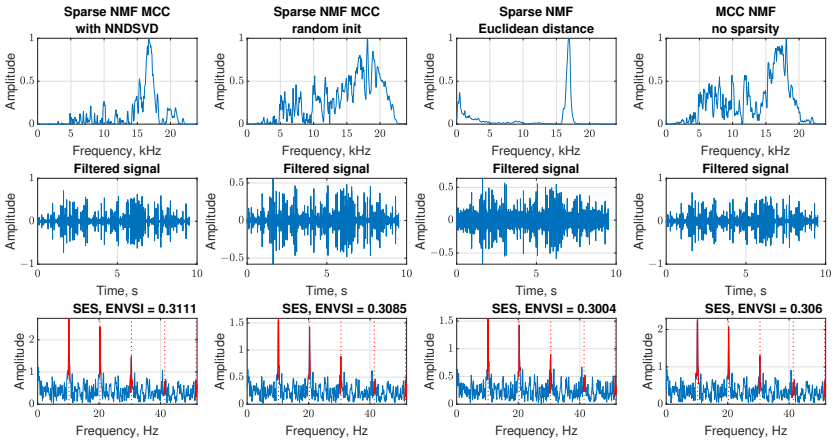

Figure A1. Filtering results for Case 1.

In Figure A2, the Sparse NMF with the Euclidean distance does not allow for obtaining the correct result. It suggests that the use of MCC is reasonable.

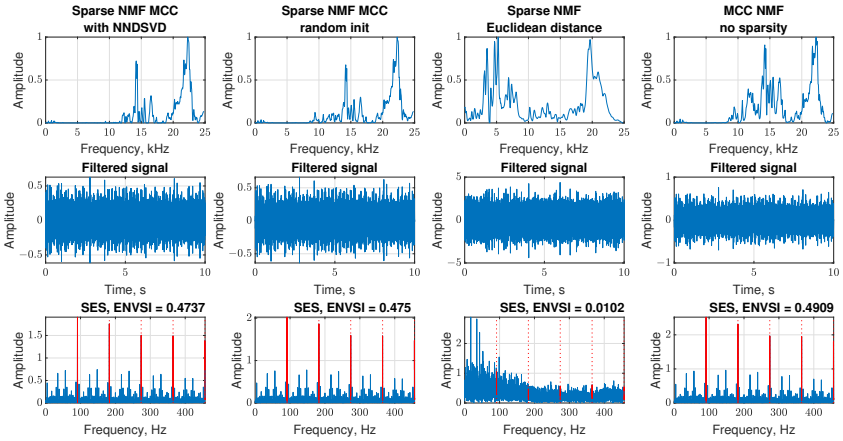

Figure A2. Filtering results for Case 3.

Finally, as shown in Figure A3, only the proposed procedure employing Sparse MCC NMF with NNDSVD initialization yields the correct result. The other methods in this comparison fail to produce filter characteristics conducive to obtaining a valid outcome. In summary, each element of the approach is important.

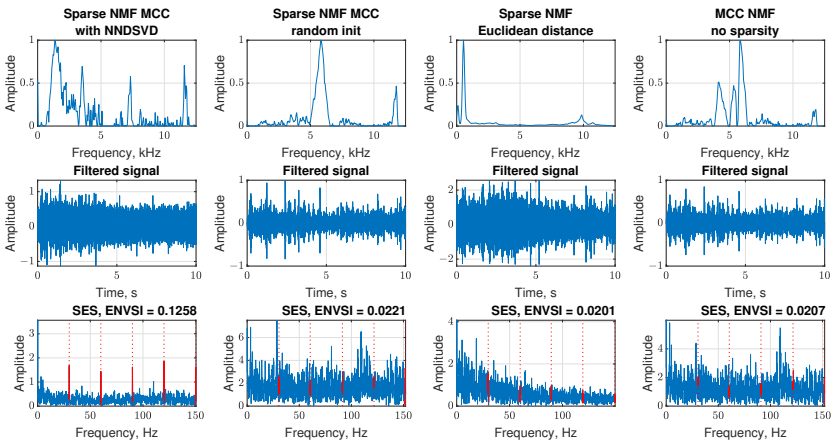

Figure A3. Filtering results for Case 2.

Appendix B

The list of key variables and parameters used in the proposed methodology has been presented in Table A1.

Table A1. Variables and Parameters Used in the Proposed Methodology.

Symbol	Description
$\mathbf{y}$	One-dimensional discrete data (raw input signal)
$STFT_{t,f}$	Short-Time Fourier Transform at time t and frequency f
w	Window function
L	Number of samples in the analyzed segment (window length)
N	Number of points used for Discrete Fourier Transform (DFT)
j	Imaginary unit
F_s	Sampling Frequency
$\mathbf{S} \in \mathbb{R}_+^{m \times n}$	Spectrogram (absolute value of $STFT$), the non-negative input matrix,
m, n	Dimensions (rows, columns) of the input matrix S
r	Rank of factorization
λ	Sparsity regularization parameter
$maxIter$	Maximum number of iterations in the Sparse NMF MCC algorithm
$\mathbf{W} \in \mathbb{R}_+^{m \times r}$	Basis matrix
$\mathbf{H} \in \mathbb{R}_+^{r \times n}$	Coefficient matrix
$\mathbf{E}$	Error matrix
V	Correntropy
$\kappa_{\sigma_k}(\cdot)$	Gaussian kernel function in Maximum Correntropy Criterion (MCC)
σ_k	Kernel bandwidth parameter ($\sigma_k > 0$) in MCC
$\|\cdot\|_1$	l_1 norm
Ω	MCC weight matrix
$\odot$	Hadamard product (element-wise multiplication)
$\oslash$	Hadamard division (element-wise division)
ϵ	Small constant introduced to prevent division by zero
$ENSVI$	Envelope Spectrum Indicator
SES	Squared Envelope Spectrum (SES)
$\mathbf{s}$	Signal of Interest (SOI) component
$\mathbf{n}$	Noise component
ν	Degrees of freedom (for Student's t-distribution, modeling heavy-tailed noise)
A	Magnitude/Amplitude of the sinusoidal component in SOI
f_c	Frequency of the sinusoidal component in SOI
α_s	Damping factor
$\delta(\cdot)$	Dirac delta function
T_p	Nominal period of the signal
$\Delta T_p(t)$	Slow variations in the period due to jitter effects
ϕ_s	Initial phase shift
$*$	Convolution operation
σ	standard deviation in Gaussian noise

References

1. Antoni, J. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mech. Syst. Signal Process.* **2006**, *20*, 282–307.
2. Antoni, J.; Randall, R.B. The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mech. Syst. Signal Process.* **2006**, *20*, 308–331.
3. Barszcz, T.; Randall, R.B. Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine. *Mech. Syst. Signal Process.* **2009**, *23*, 1352–1365.
4. Wang, Y.; Xiang, J.; Markert, R.; Liang, M. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications. *Mech. Syst. Signal Process.* **2016**, *66*, 679–698.

5. Liu, S.; Hou, S.; He, K.; Yang, W. L-Kurtosis and its application for fault detection of rolling element bearings. *Measurement* **2018**, *116*, 523–532.
6. Zhong, J.; Wang, D.; Li, C. A nonparametric health index and its statistical threshold for machine condition monitoring. *Measurement* **2021**, *167*, 108290.
7. Jaworski, P.; Pitera, M. The 20-60-20 rule. *Discret. Contin. Dyn. Syst. Ser. B* **2016**, *21*, 1149–1166.
8. Hebda-Sobkowicz, J.; Zimroz, R.; Pitera, M.; Wyłomańska, A. Informative frequency band selection in the presence of non-Gaussian noise—A novel approach based on the conditional variance statistic with application to bearing fault diagnosis. *Mech. Syst. Signal Process.* **2020**, *145*, 106971.
9. Nowicki, J.; Hebda-Sobkowicz, J.; Zimroz, R.; Wyłomańska, A. Dependency measures for the diagnosis of local faults in application to the heavy-tailed vibration signal. *Appl. Acoust.* **2021**, *178*, 107974.
10. Żak, G.; Wyłomańska, A.; Zimroz, R. Application of alpha-stable distribution approach for local damage detection in rotating machines. *J. Vibroengineering* **2015**, *17*, 2987–3002.
11. Randall, R.B.; Antoni, J.; Chobsaard, S. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mech. Syst. Signal Process.* **2001**, *15*, 945–962.
12. Borghesani, P.; Shahriar, M.R. Cyclostationary analysis with logarithmic variance stabilisation. *Mech. Syst. Signal Process.* **2016**, *70*, 51–72.
13. Bonnardot, F.; Randall, R.; Guillet, F. Extraction of second-order cyclostationary sources—Application to vibration analysis. *Mech. Syst. Signal Process.* **2005**, *19*, 1230–1244.
14. Marsick, A.; André, H.; Khelf, I.; Leclère, Q.; Antoni, J. Restoring cyclostationarity of rolling element bearing signals from the instantaneous phase of their envelope. *Mech. Syst. Signal Process.* **2023**, *193*, 110264.
15. Borghesani, P.; Antoni, J. A faster algorithm for the calculation of the fast spectral correlation. *Mech. Syst. Signal Process.* **2018**, *111*, 113–118.
16. Wodecki, J.; Michalak, A.; Wyłomańska, A.; Zimroz, R. Influence of non-Gaussian noise on the effectiveness of cyclostationary analysis–Simulations and real data analysis. *Measurement* **2021**, *171*, 108814.
17. Flandrin, P.; Rilling, G.; Goncalves, P. Empirical mode decomposition as a filter bank. *IEEE Signal Process. Lett.* **2004**, *11*, 112–114.
18. Arifin, M.S.; Wang, W.; Uddin, M.N. A modified EMD technique for broken rotor bar fault detection in induction machines. *Sensors* **2024**, *24*, 5186.
19. Wang, W.; Yuan, H. Bearing Fault Feature Extraction Method Based on Adaptive Time-Varying Filtering Empirical Mode Decomposition and Singular Value Decomposition Denoising. *Machines* **2025**, *13*, 50.
20. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2013**, *62*, 531–544.
21. Jiang, W.; Qi, Z.; Jiang, A.; Chang, S.; Xia, X. Lightweight network bearing intelligent fault diagnosis based on VMD-FK-ShuffleNetV2. *Machines* **2024**, *12*, 608.
22. Liu, G.; Ma, Y.; Wang, N. Rolling Bearing Fault Diagnosis Based on SABO–VMD and WMH–KNN. *Sensors* **2024**, *24*, 5003.
23. Shiri, H.; Wodecki, J. Analysis of the sound signal to fault detection of bearings based on Variational Mode Decomposition. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *942*, 012020.
24. Smith, J.S. The local mean decomposition and its application to EEG perception data. *J. R. Soc. Interface* **2005**, *2*, 443–454.
25. Shiri, H.; Wodecki, J.; Ziętek, B.; Zimroz, R. Inspection robotic UGV platform and the procedure for an acoustic signal-based fault detection in belt conveyor idler. *Energies* **2021**, *14*, 7646.
26. Lei, Y.; Lin, J.; He, Z.; Zuo, M.J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech. Syst. Signal Process.* **2013**, *35*, 108–126.
27. Wang, Y.; Liu, F.; Jiang, Z.; He, S.; Mo, Q. Complex variational mode decomposition for signal processing applications. *Mech. Syst. Signal Process.* **2017**, *86*, 75–85.
28. Li, Y.; Si, S.; Liu, Z.; Liang, X. Review of local mean decomposition and its application in fault diagnosis of rotating machinery. *J. Syst. Eng. Electron.* **2019**, *30*, 799–814.
29. Randall, R.B.; Antoni, J. Why EMD and similar decompositions are of little benefit for bearing diagnostics. *Mech. Syst. Signal Process.* **2023**, *192*, 110207.
30. Zhang, F.; Li, Z.; Cheng, Y.; Yang, Y.; Han, Y.; Yi, C.; Li, T.; Yan, J.; Dong, L.; Zhang, W. MHSNet: A Multi-Scale Hidden State Interaction Network for Fault Diagnosis of Rotating Machinery. *Tsinghua Sci. Technol.* **2025,** [CrossRef]
31. Wang, H.; Song, Y.; Yang, H.; Liu, Z. Generalized Koopman Neural Operator for Data-driven Modelling of Electric Railway Pantograph-catenary Systems. *IEEE Trans. Transp. Electrif.* **2025,** [CrossRef]
32. Yan, J.; Cheng, Y.; Zhang, F.; Li, M.; Zhou, N.; Jin, B.; Wang, H.; Yang, H.; Zhang, W. Research on multimodal techniques for arc detection in railway systems with limited data. *Struct. Health Monit.* **2025**, [CrossRef]
33. Xu, H.; Pan, H.; Zheng, J.; Tong, J.; Zhang, F.; Chu, F. Intelligent fault identification in sample imbalance scenarios using robust low-rank matrix classifier with fuzzy weighting factor. *Appl. Soft Comput.* **2024**, *152*, 111229.

34. Wodecki, J.; Kruczek, P.; Bartkowiak, A.; Zimroz, R.; Wyłomańska, A. Novel method of informative frequency band selection for vibration signal using Nonnegative Matrix Factorization of spectrogram matrix. *Mech. Syst. Signal Process.* **2019**, *130*, 585–596.
35. Liang, L.; Ding, X.; Liu, F.; Chen, Y.; Wen, H. Feature extraction using sparse kernel non-negative matrix factorization for rolling element bearing diagnosis. *Sensors* **2021**, *21*, 3680.
36. Hou, W.; Liu, X.; Wang, J.; Chen, C.; Xu, X. Multispectral Land Surface Reflectance Reconstruction Based on Non-Negative Matrix Factorization: Bridging Spectral Resolution Gaps for GRASP TROPOMI BRDF Product in Visible. *Remote Sens.* **2025**, *17*, 1053.
37. Wang, M.; Zhang, W.; Shao, M.; Wang, G. Separation and Extraction of Compound-Fault Signal Based on Multi-Constraint Non-Negative Matrix Factorization. *Entropy* **2024**, *26*, 583.
38. Gabor, M.; Zdunek, R.; Zimroz, R.; Wylomanska, A. Bearing Damage Detection With Orthogonal and Nonnegative Low-Rank Feature Extraction. *IEEE Trans. Ind. Inform.* **2023**, *20*, 2944–2955.
39. Michalak, A.; Zdunek, R.; Zimroz, R.; Wyłomańska, A. Influence of α-Stable Noise on the Effectiveness of Non-Negative Matrix Factorization—Simulations and Real Data Analysis. *Electronics* **2024**, *13*, 829.
40. Boutsidis, C.; Gallopoulos, E. SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognit.* **2008**, *41*, 1350–1362.
41. Allen, J. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Trans. Acoust. Speech Signal Process.* **1977**, *25*, 235–238.
42. Lee, D.; Seung, S. Algorithms for non-negative matrix factorization. In Proceedings of the Advances in Neural Information Processing Systems Denver, CO, USA, 11 May 2001; MIT Press: Cambridge, MA, USA, 2001; pp. 556–562.
43. Cichocki, A.; Zdunek, R.; Phan, A.H.; Amari, S.i. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
44. Cichocki, A.; Phan, A.H.; Caiafa, C. Flexible HALS algorithms for sparse non-negative matrix/tensor factorization. In Proceedings of the 2008 IEEE Workshop on Machine Learning for Signal Processing, Cancun, Mexico, 16–19 October 2008; pp. 73–78.
45. Cichocki, A.; Phan, A.H. Fast local algorithms for large scale nonnegative matrix and tensor factorizations. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **2009**, *92*, 708–721.
46. Chen, B.; Zhu, Y.; Hu, J.; Principe, J.C. *System Parameter Identification: Information Criteria and Algorithms*; Elsevier: London, UK, 2013.
47. Peng, S.; Chen, B.; Sun, L.; Ser, W.; Lin, Z. Constrained maximum correntropy adaptive filtering. *Signal Process.* **2017**, *140*, 116–126.
48. Wang, J.J.Y.; Wang, X.; Gao, X. Non-negative matrix factorization by maximizing correntropy for cancer clustering. *BMC Bioinform.* **2013**, *14*, 107.
49. Peng, S.; Ser, W.; Lin, Z.; Chen, B. Robust sparse nonnegative matrix factorization based on maximum correntropy criterion. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018; IEEE eXpress Conference Publishing: Red Hook, NY, USA, 2018; pp. 1–5.
50. Du, L.; Li, X.; Shen, Y.D. Robust nonnegative matrix factorization via half-quadratic minimization. In Proceedings of the 2012 IEEE 12th International Conference on Data Mining, Brussels, Belgium, 10–13 December 2012; IEEE Computer Society Conference Publishing Services (CPS): Danvers, MA, USA, 2012; pp. 201–210.
51. Shiri, H.; Zimroz, P.; Wodecki, J.; Wyłomańska, A.; Zimroz, R.; Szabat, K. Using long-term condition monitoring data with non-Gaussian noise for online diagnostics. *Mech. Syst. Signal Process.* **2023**, *200*, 110472.
52. Shiri, H.; Zimroz, P.; Wyłomańska, A.; Zimroz, R. Estimation of machinery's remaining useful life in the presence of non-Gaussian noise by using a robust extended Kalman filter. *Measurement* **2024**, *235*, 114882.
53. Berry, M.W.; Browne, M.; Langville, A.N.; Pauca, V.P.; Plemmons, R.J. Algorithms and applications for approximate nonnegative matrix factorization. *Comput. Stat. Data Anal.* **2007**, *52*, 155–173.

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.