Acoustic Monitoring of Rolling Element Bearings using a Sparse Microphone Array

Xian Wu, Jeroen Zegers, Hervé Denayer, Konstantinos Gryllias

Acoustic Monitoring of Rolling Element Bearings using a Sparse Microphone Array

Xian Wu[1,2], Jeroen Zegers[3], Hervé Denayer[1,2], Konstantinos Gryllias[1,2]

[1]KU Leuven, Department of Mechanical Engineering, Celestijnenlaan 300, B-3001, Heverlee, Belgium
[2]Flanders Make@KU Leuven, Belgium
[3]Flanders Make, 3001 Heverlee, Belgium
xian.wu1@kuleuven.be

Abstract

Acoustic monitoring of rolling element bearings in industrial environments can provide a non-contact solution for early detection of bearing failures and prevention of costly downtime. Nevertheless, extracting the bearing signature of interest from the other contributions in the acoustic signals remains a key challenge. A possible approach to improve the Signal-to-Noise Ratio (SNR) of bearing signatures is to exploit spatial information, obtained using acoustic imaging techniques. However, Conventional microphone arrays are rarely used for bearing condition monitoring in an industrial setting due to their extensive number of microphones, which leads to high complexity and cost. In this paper, we propose a sparse microphone array, exploiting the modulation characteristics of the bearing signatures. The process begins with the extraction of the envelope from the raw signals, followed by frequency domain conventional beamforming on the envelope signals to generate acoustic imaging maps of the bearing signals. These maps are then used to optimize an indicator within a predefined region, effectively separating the bearing signal. The validity of this methodology is confirmed through numerical simulations of point sources in a free field environment, using simple amplitude modulated signals and signals generated with a phenomenological rolling element bearing model. Finally, experimental results collected from an in-house test rig demonstrate the robustness and accuracy of the proposed methodology in industrial environments.

1 Introduction

Condition monitoring of machinery, a critical aspect of predictive maintenance that ensures the smooth operation of industrial processes, traditionally relies primarily on vibration analysis. In 2011, Randall and Antoni [12] provide a comprehensive tutorial on rolling element bearing diagnostics. Nonetheless, vibration-based fault diagnosis techniques possess inherent limitations. The accelerometers have to be mounted very close to the rotating component to be analyzed, which can be challenging due to accessibility constraints; Moreover, multiple accelerometers are required to track multiple bearings, which can be impractical in real-life scenarios [7].

In recent years, acoustic monitoring has emerged as a promising alternative for condition monitoring [6, 17, 19]. Unlike vibration analysis, acoustic monitoring does not require contact with the machine, making it a more flexible and less intrusive method. It is easier to install and allows for simultaneous monitoring of multiple components [8]. In spite of its advantages, acoustic monitoring also encounters challenges, particularly in noisy environments where background noise and interference from other sources can mask the useful signals. This motivates researchers to explore microphone array techniques as a means to enhance the signal quality and improve the accuracy of fault detection. Cardenas Cabada et al. [5] combine spectral kurtosis with acoustic imaging to detect impulsive sources linked to faults in rotating machinery. Another study by Mauricio et al. [9] employs a microphone array to monitor a bearing on a test rig, using envelope signals instead of raw signals for time-domain beamforming. However, these studies do not delve into the impact of microphone array geometry on the performance of the system, particularly when using envelope signals.

In a practical industrial setting, the use of fewer microphones would be highly beneficial as it would lower the costs associated with hardware for acoustic monitoring, thus giving rise to the concept of sparse arrays.

This concept of sparse arrays has also seen exploration in sound event classification within household environments [3]. In this household context, a sparse array was deployed, and various array processing techniques were tested. Conventional beamforming techniques like Delay-and-Sum (DAS) and Minimum Variance Distortionless Response (MVDR) showed susceptibility to aliasing effects due to inadequate spatial sampling points. However, the study demonstrated that the Multiple Signal Classification (MUSIC) algorithm could address this issue, thanks to its super-resolution capability. Despite its effectiveness, the MUSIC algorithm has a limitation—it relies on prior knowledge of the number of sources, which can be challenging in noisy environments. Nevertheless, we find inspiration in other fields. For instance, in radio electronics, lowering the sampling frequency to decrease the number of required antennas is a common practice [13]. In medical ultrasonic imaging, envelope beamforming, which involves demodulation before beamforming, is an established idea to decrease the required number of transducers [10].

In this paper, we seek to bridge this gap by focusing on the exploration of microphone array geometry and the employment of envelope signals in the context of acoustic machinery monitoring. We first delve into the signal characteristics of Rolling Element Bearings (REBs) in section 2 and provide a brief overview of beamforming (section 3) and envelope extraction concepts (section 4). We then propose an approach leveraging sparse arrays that encompasses four stages: *Subband Selection*, *Envelope Extraction*, *Beamforming*, and *Indicator Optimization* (section 5). The specific advantages and applications of our methodology will be discussed in detail with validation results obtained in simulation (section 6) and experiment (section 7). Finally, we conclude with key findings and future directions in section section 8.

2 Signal Characteristics of Rolling Element Bearing and Fault Indicator

Rolling element bearings (REBs) are integral to a wide variety of rotating machinery [12]. Comprising inner and outer rings, rolling elements (balls or rollers), and a cage, these components are designed for durability. However, they are not immune to faults such as surface defects, cracking, and wear. These faults manifest as distinct vibration signatures, which are crucial for diagnostic purposes. The vibration, and hence acoustic, signatures of REBs are typically a modulated blend of two frequencies: a lower frequency signal associated with the fault, and the bearing's inherent higher frequency signal [12]. One such fault-related frequency is the Ball Pass Frequency of the Outer race (BPFO). The BPFO, which denotes the frequency at which a rolling element passes a specific point on the outer raceway. A deviation in the BPFO signal can indicate a defect in the outer raceway.

In this work, we delve into the Squared Envelope Spectrum (SES) of these signals, a rich source of diagnostic information [1]. Particularly, we employed a key indicator of bearing faults, derived from the SES of the bearing signal. This indicator captures the energy at the BPFO, providing a quantitative measure of the severity of outer raceway defects. The indicator, as shown in Equation 1, is defined as the ratio of the peak amplitude at the BPFO frequency to the average signal floor level. The signal floor level is computed using a moving average, denoted by μ, and the median absolute deviation, denoted by v, of the SES at the BPFO frequency. To establish a robust threshold, we apply a factor of 2 in the signal floor computation. This threshold accommodates normal signal fluctuations while remaining sensitive to significant changes, thereby minimizing the risk of false positives.

$$I = \frac{P}{\text{mean}(\mu + 2v)} \tag{1}$$

Figure 1 illustrates the SES of a REB with an outer ring fault. The BPFO frequency and the average signal floor level are marked for reference. The computed indicator, designated as I, serves as a reliable metric for detecting outer raceway defects in REBs, thereby facilitating the development of effective diagnostic strategies.

3 Beamforming

Beamforming, fundamentally a method of spatial filtering, employs an array of sensors to focus on a specific direction. This technique operates under the assumption of plane waves, scanning a single direction of wave incidence at a time. When considering point sources or spherical waves, the method concentrates its attention towards one of these sources [11].

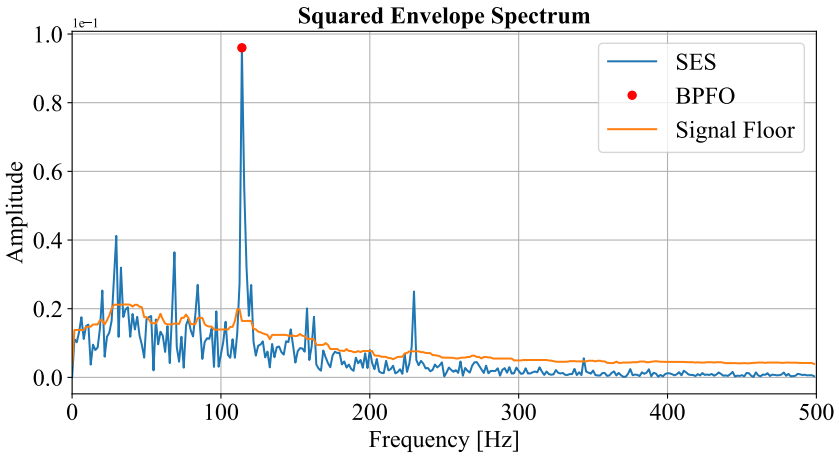


Figure 1: SES of a REB with an outer ring fault. The red dot represents the BPFO peak used for the indicator calculation, the blue line represents the SES of the signal, and the orange line illustrates the signal floor.

Note that in this work, we use bold capital letters to denote matrices, bold lower-case letters for vectors, non-bold lower-case letters for scalars in time domain, and non-bold upper-case letters for scalars in frequency domain. The core equation of beamforming expresses the pressure field $\mathbf{p}$ as a function of the sources $\mathbf{q}$ through the matrix $\mathbf{G}$, which can be written as:

$$\mathbf{p} = \mathbf{G}\mathbf{q} \tag{2}$$

The vector $\mathbf{p}$ is a $m \times 1$ vector representing the measurements from m microphones, $\mathbf{q}$ is a $n \times 1$ vector representing n sources and hence the transfer matrix $\mathbf{G}$ has a dimension of $m \times n$. In this study, we will focus on nearfield beamforming due to the close proximity of the array to the source. Therefore, we will utilize a monopole source model to establish the transfer functions [18]. Each element of the transfer function matrix $\mathbf{G}$ is calculated based on the corresponding element of the distance matrix $\mathbf{R}$, as described by the following equation:

$$G_{mn} = \frac{1}{4\pi r_{mn}} e^{-jkr_{mn}} \tag{3}$$

In the above equation, G_{mn} represents the element in the m-th row and n-th column of the transfer function matrix $\mathbf{G}$, and r_{mn} represents the corresponding element of the distance matrix $\mathbf{R}$. Here, k represents the wave number, which is related to the frequency and the speed of sound in the medium. The symbol j is the imaginary unit, which is used here to represent the phase shift of the wave due to propagation. Solving this equation using the least squares method, we obtain:

$$\mathbf{q} = \frac{\mathbf{w}(\mathbf{G}^H \mathbf{p})}{\|\mathbf{G}\|^2} \tag{4}$$

where $\mathbf{w}$ is a weight vector and $\mathbf{G}^H$ represents the Hermitian transpose of $\mathbf{G}$. To counterbalance the greater weight given to distant sources in Equation 4, the weight vector $\mathbf{w}$ can be defined as:

$$\mathbf{w} = \|\mathbf{G}\| \tag{5}$$

This normalizes the expectation of the identified sources with respect to their positions. In frequency domain beamforming, the signals $\mathbf{p}$ used in Equation 4 represent the coefficients of the discrete Fourier transform of the complete temporal signals. Once the source is identified, its temporal signal is obtained using an inverse Fourier transform.

Furthermore, the physical layout of the microphone array, particularly its size and microphone distribution, is a critical factor influencing the quality of beamforming. Specifically, the operational range of a microphone array in beamforming analysis is restricted to frequencies below the frequency where spatial aliasing starts to occur, which in turn is dependent on the distance between the microphones, D_m as well as the speed of sound c. This relationship can be expressed as:

$$f_{\max} = \frac{c}{2D_m} \tag{6}$$

For instance, assuming a speed of sound c of 343 m/s, for a BPFO around 120 Hz, the corresponding D_m would be about 1.43 m. For a carrier frequency of 1000 Hz, D_m would be roughly 0.17 m. This relationship is crucial for configuring the microphone array to accurately localise the acoustic sources of interest.

4 Use of the Envelope and its Extraction via Hilbert Transform

In the process of signal propagation, the signal P passes through a filter bank H. Each filtered signal, represented in the frequency domain, is given by:

$$P_i(f) = H_i(f) \cdot P(f) \tag{7}$$

The original signal, $P(f)$, is the result of the source signal $Q(f)$ being propagated through the transfer function $G(f)$, thus we have:

$$P(f) = G(f) \cdot Q(f) \tag{8}$$

Combining Equation 7 and Equation 8, we have:

$$P_i(f) = G(f) \cdot H_i(f) \cdot Q(f) = G(f) \cdot Q_i(f) \tag{9}$$

This equation indicates that the filtered signal in the frequency domain is the original signal passed through the filter i and modulated by the transfer function. We then approximate the transfer function around f_0 using a first-order Taylor series. This approximation is sufficient for small variations in frequency around f_0, leading to the equation:

$$P_i(f) \approx |G(f_0)| e^{j[\phi(f_0) + (f-f_0)\phi'(f_0)]} Q_i(f) \tag{10}$$

where $\phi(f_0)$ and $\phi'(f_0)$ represent the phase and the derivative of the phase of the transfer function at f_0, respectively. Applying the inverse Fourier transform to transition to the time domain, we obtain the representation of the filtered signal:

$$p_i(t) = |G(f_0)| e^{j\phi(f_0)} e^{-jf_0\phi'(f_0)} q_i(t + \frac{\phi'(f_0)}{2\pi}) \tag{11}$$

This equation introduces a time delay term $\frac{\phi'(f_0)}{2\pi}$, which arises due to the inverse relationship between frequency and time in the Fourier transform. We then consider the filtered signal as a sinusoid of frequency f_0, modulated by an envelope function $a_p(t)$. This approximation, given by:

$$p_i(t) = a_p(t) e^{j2\pi f_0 t} \tag{12}$$

is valid when the signal's amplitude varies slowly compared to the sinusoidal frequency. Substituting Equation 11 into the definition of the envelope function (Equation 12), we find:

$$a_p(t) \approx G(f_0) e^{-jf_0\phi'(f_0)} a_q(t + \frac{\phi'(f_0)}{2\pi}) \tag{13}$$

This equation elucidates the relationship between the envelope function of the output signal, $a_p(t)$, and that of the input signal, $a_q(t)$, showing that the output envelope is a delayed, scaled version of the input envelope. The modulus of the envelope function is given by:

$$|a_p(t)| \approx |G(f_0)||a_q(t - \tau_g(f_0))| \tag{14}$$

where $\tau_g = -\frac{\phi'(f_0)}{2\pi}$, is referred to as the group delay. The group delay represents the time delay of the envelope of the signal, which can be different from the delay of the signal itself, especially in dispersive media. For non-dispersive media, such as air, where all frequencies of signals propagate at the same speed c, we find that the envelope function is given by:

$$|a_p(t)| = \frac{|a_q(t - \frac{r}{c})|}{r} \tag{15}$$

where r is the distance between the source signal and the microphone. This relationship is pivotal as it allows us to consider the envelope $a_p(t)$ as an ordinary time signal for signal processing tasks such as beamforming. To extract the envelope, we apply the Hilbert transform to the subband component $p_i(t)$, yielding the analytical signal:

$$p_{a,i}(t) = p_i(t) + jp'_i(t) \tag{16}$$

This signal can be expressed in polar form as:

$$p_{a,i}(t) = a_{p,i}(t)e^{j\theta_i(t)} \tag{17}$$

where $a_{p,i}(t)$ represents the envelope of the signal and $\theta_i(t)$ is the instantaneous phase. The envelope can then be extracted from the analytical signal using the equation:

$$a_{p,i}(t) = \sqrt{p_i(t)^2 + p'_i(t)^2} \tag{18}$$

In summary, it has been shown that the envelope, with its spatial information, can be treated as a time signal for beamforming and outlined a method for its extraction.

5 Proposed Methodology

In this study, we propose a novel approach that leverages the benefits of a sparse array by operating it at its maximum allowable frequency. This interpretation of the Shannon condition, which traditionally emphasizes signal frequency, focuses on the envelopes of the signals. These envelopes have frequencies significantly lower than the signals themselves, allowing us to increase the distance between microphones and decrease the total number of microphones required.

Our proposed methodology involves a four-step process:

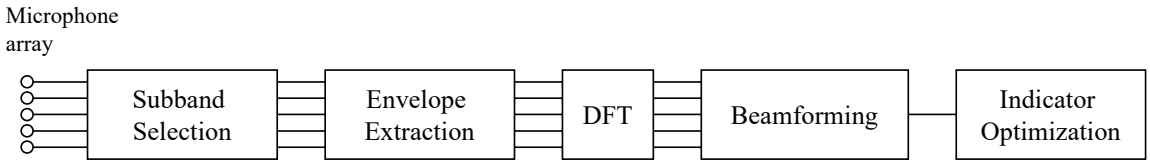


Figure 2: Proposed Methodology.

1. **Subband Selection:** This method starts with identifying the carrier band, which is accomplished through spectral coherence, a technique utilized in cyclostationary analysis [2].

2. **Envelope Extraction:** Once the carrier band is identified, we process the signal within this band using the Hilbert Transform to create an analytical signal. From this signal, we extract the envelope, which holds the signature of the bearing.

3. **Beamforming:** The extracted envelope is transformed to the frequency domain using the Discrete Fourier Transform (DFT). This sets the stage for frequency-domain beamforming, a process that scans a predefined area and generates an array of source signals.

4. **Indicator Optimization:** Each potential signal, transformed back to the time domain via an inverse DFT, is then evaluated using an indicator. The indicator quantifies the signal's quality and is defined as the ratio of the peak amplitude of the BPFO to a floor value in the envelope spectrum (computed using the moving median absolute deviation). The signal possessing the maximum indicator, presumably with the best SNR, is selected.

While this approach may impact spatial resolution due to the lower operating frequency, it retains a large aperture size, which is often crucial in industrial settings for monitoring large machinery. This methodology provides a different perspective within the field of beamforming research, offering a balance between spatial resolution and coverage area. It contrasts with alternative strategies, such as using a dense array with the same number of microphones but with a much smaller aperture size.

6 Application of the Proposed Methodology on Numerical Simulations

In this section, we aim to validate the proposed methodology by conducting a series of simulations. The synthetic signals, generated from a phenomenological model, are tailored to replicate a localized outer ring fault at a constant shaft speed. The effectiveness of the simulations is determined using the indicator detailed in section 2.

6.1 Simulation Configuration

The synthetic signals for this study are crafted using a tool that is designed for simulating mechanical signatures of faulted gears and bearings in non-stationary conditions [4]. In this study, we simulate a REB fault scenario with a constant shaft speed and a localized fault in the outer ring. The synthetic signal generation mirrors the process detailed in the referenced work, starting with an impulse train that represents the fault occurrence. This impulse train is then passed through a linear time-invariant (LTI) system modeled as a single degree of freedom (SDOF) system, as per the original model's methodology. This process generates an acceleration response from the displacement response, which can be interpreted as the acceleration signal of the bearing. We extend this process by first obtaining a velocity response through single differentiation of the displacement response. This velocity response is then used as an input to a vibroacoustic model to compute the corresponding acoustic pressure P. Assuming the source can be represented by a pulsating sphere with surface s and uniform radial velocity v, the acoustic radiation formula is given by:

$$P = \frac{j\omega\rho_0 sv}{4\pi r} e^{-jkr} \tag{19}$$

where ω is the angular frequency, ρ_0 is the air density, s is the vibrating surface area, v is the vibration velocity, r is the distance to the observer, and $k = \frac{\omega}{c}$ with c being the speed of sound. The simulation is set with the following parameters: the shaft speed is kept constant at 30 Hz, and three simulated acoustic sources are introduced in the environment. The first source is our target, while the other two sources are intended to serve as interference. The BPFO for the target source and the first interference source is set at 120 Hz, whereas for the second interference, it is reduced to 105 Hz. The natural frequencies for the target and the first interference source are set at 1000 Hz. However, the second interference source has a higher natural frequency, set at 2000 Hz. These parameters, especially the subtle frequency differences between the sources, are chosen to simulate a complex acoustic environment that closely mimics the realistic scenarios of encountering multiple faulty bearings.

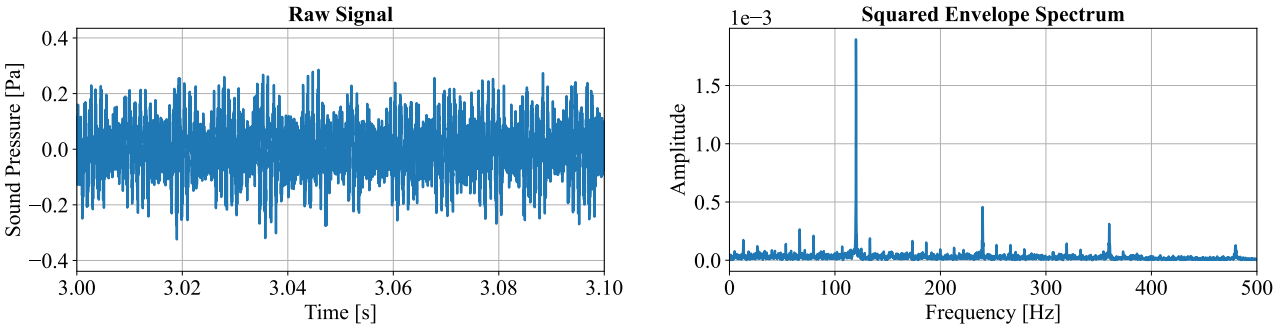


Figure 3: Simulated REB faulty signal of source 1.

In addition to the signal generation, the configuration of the microphone array geometry plays a critical role in the simulation setup. Positioned in the xy-plane, the microphone array is designed as a square grid,

with each side measuring 1.5 m. Depending on the specific requirements of data collection and resolution, the number of microphones in each direction can be adjusted. These microphones are placed at equidistant points and maintained at the same height of 1.65 m in the z direction. The three simulated acoustic sources, each representing a unique fault condition, are located at different positions in the xy-plane, all at the same height in the z direction. Table 1 presents the specific configuration of the simulated sources.

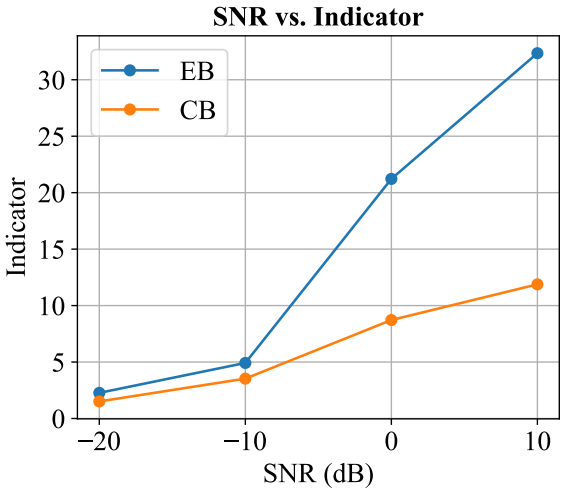
Source	Symbol	x [m]	y [m]	z [m]	Carrier [Hz]	Modulation [Hz]
1 (Target)	circle	1.1	1.2	1	1000	120
2 (Interference)	triangle	0.5	0.2	1	1000	105
3 (Interference)	star	0	1.2	1	2000	120

Table 1: Configuration of the simulated sources.

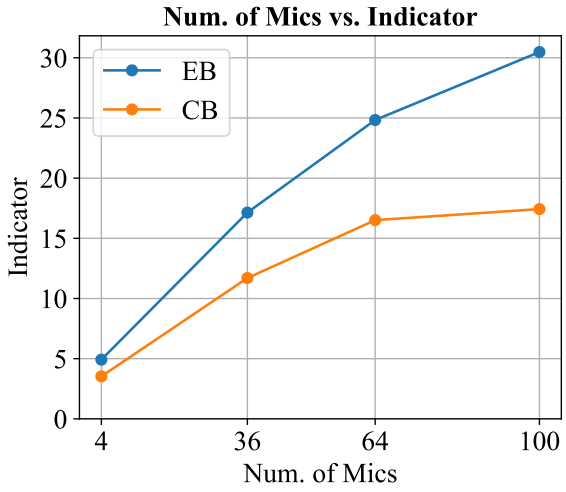
During the course of the study, we will vary the number of microphones per side from 2 to 10. This change translates into a total of 4 to 100 microphones in the array, thereby altering the distance between each microphone from 1.5 m to 0.17 m. This extensive range facilitates an in-depth investigation into the impact of microphone array density on the effectiveness of the proposed signal processing technique.

6.2 Analysis of Simulation Results

The performance of the proposed Envelope Beamforming (EB) methodology is compared against the Conventional Beamforming (CB) approach. The primary distinction lies in the sequence of operations: EB extracts the envelope before performing beamforming, while CB performs beamforming directly on the raw signal before envelope extraction. The comparison is conducted under varying SNR conditions and different numbers of microphones. In the simulation, the SNR is adjusted by adding Gaussian white noise to the signals, while the number of microphones is varied to observe the effect on the signal quality. The indicators in the following figures are calculated from the signal extracted within a spherical volume surrounding the target source. The sphere has a radius of 0.2 m centered around the source.



(a) Comparison of EB and CB for variable SNRs with four microphones.



(b) Comparison of EB and CB for variable microphone counts at -10 dB SNR.

Figure 4: Comparison of EB and CB.

The simulation results underscore the superior performance of the EB method. As illustrated in Figure 4a, the indicator values for EB consistently exceed those of CB across all tested SNR levels, particularly at higher SNRs, demonstrating EB's robustness in effectively extracting useful information from noisy signals. When varying the number of microphones (Figure 4b), EB continues to demonstrate its advantage. Even with a limited number of microphones, EB delivers a superior signal quality indicator. Conversely, CB, despite a significant increase in the number of microphones to 100, only achieves an indicator of 16.51 at -10 dB SNR, still less than what EB achieves with only 36 microphones, further emphasizing the efficacy of the proposed methodology.

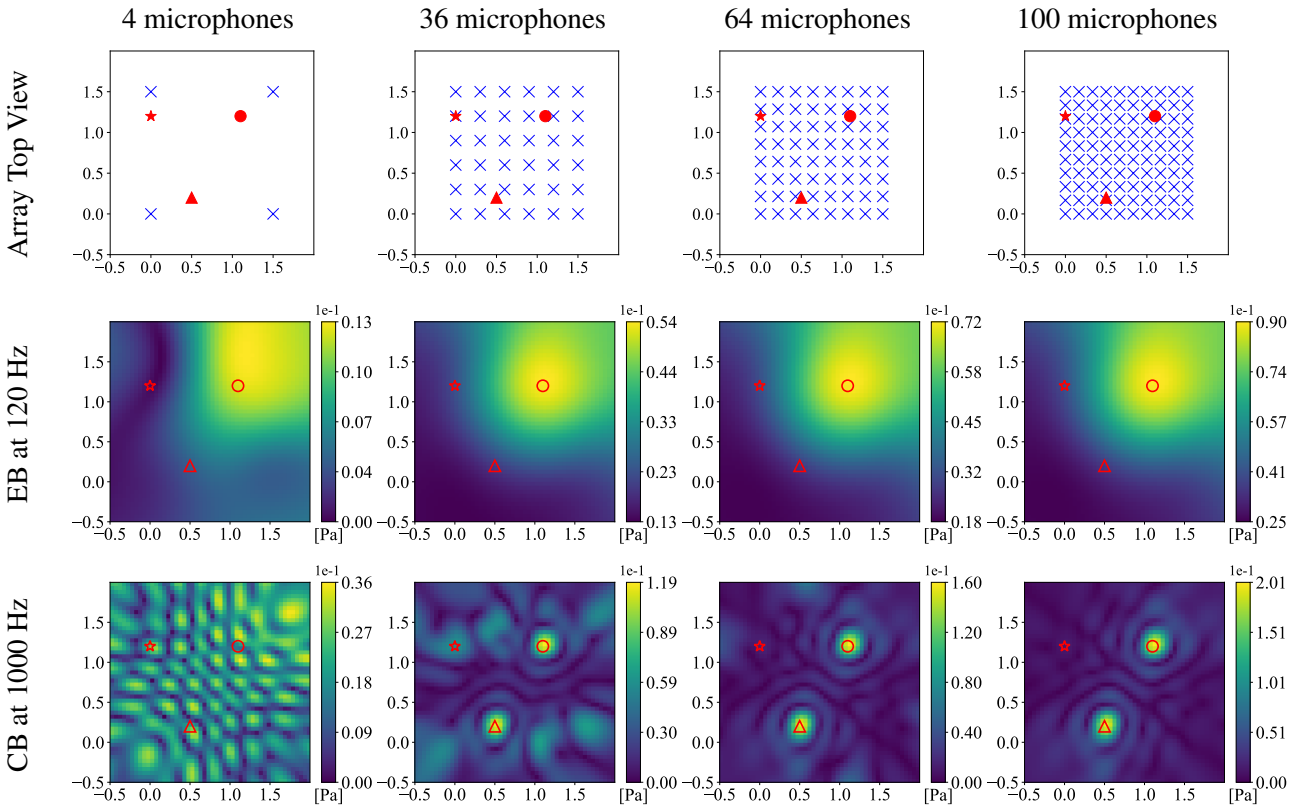


Figure 5: Acoustic Imaging Maps with Variable Microphone Counts (circle for Source 1, triangle for Source 2, star for Source 3 and the cross symbols denote positions of the array microphones in a plane 65 *cm* distant from the sources).

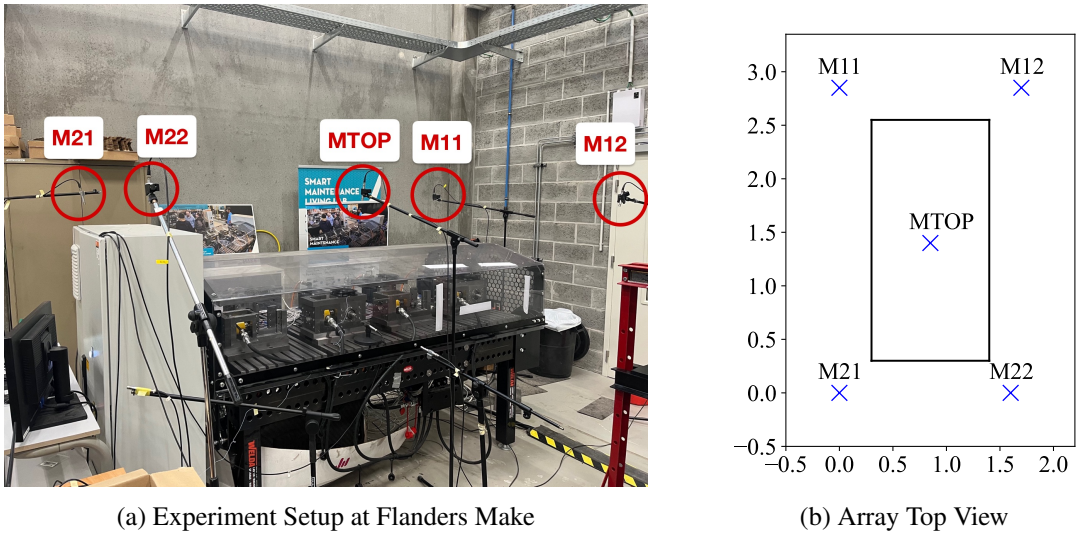
The acoustic imaging maps in Figure 5 produced by both methods reveal notable differences. The EB method at 120 Hz effectively isolates the target source (source 1) by filtering out the interfering sources 2 and 3. This is achieved through an initial filtering at the carrier frequency of 1000 Hz, which removes source 3, followed by beamforming at 120 Hz, which filters out source 2. Conversely, the CB method at 1000 Hz detects both the target source (source 1) and one of the interferences (source 2) because it performs beamforming directly on the raw signal at the carrier frequency of 1000 Hz, which both sources share. Nevertheless, source 3 is not detected due to its different carrier frequency of 2000 Hz. However, while EB can localize the source, it displays a large lobe (heat zone) due to beamforming at the target frequency of 120 Hz. On the other hand, CB begins to produce acceptable maps with at least 36 microphones, offering better resolution due to beamforming at the target frequency of 1000 Hz. Yet, if the carrier band frequency increases in practical scenarios, CB may necessitate an even larger array of microphones. In the EB images, we can observe different amplitudes on the map even though the imaging map is more or less the same. This is primarily due to the choice of weights described in Equation 5 in section 3. Despite this, the chosen steering vector formulation results in more precise localization and imaging maps [14]. It is important to note that in this context, our focus lies more on the precision of the localization and on relative values rather than on absolute values. Although the larger lobe suggested by the EB method may hint at less precise localization, performance can be enhanced using advanced beamforming techniques such as deconvolution (CLEAN) [15]. Furthermore, in some real-world scenarios, rough identification of the faulty machine may suffice, making precise localization less critical. In such cases, the noise robustness and fewer microphone requirements of EB present a more feasible solution.

7 Application of the Proposed Methodology on Real Measurements

In this section, the proposed methodology is applied to acoustic signals captured from an experimental setup featuring a REB with an outer ring fault. The methodology's application encompasses the utilization of a microphone array for signal acquisition, implementation of signal processing techniques for data analysis, and a validation procedure conducted to verify the effectiveness of the implemented methods.

7.1 Experiment Setup

The experimental setup, facilitated by Flanders Make, consisted of an array of five microphones and an accelerometer, specifically designed to capture acoustic signals from a FAG 6205-C-TVH bearing integrated into a test rig. The shaft within this rig was powered by an electric motor, maintaining a constant speed of 32 Hz. A hydraulic pump cycled every 20 seconds, applying a 9 *kN* load on the bearing for a duration of 5 seconds each time. This intermittent operation of the pump introduced significant noise to the acoustic environment, in addition to the background noise from machinery and human activities. The test rig was enclosed within an acrylic cover, causing the sound emitted from the bearing to travel to the cover before reaching the microphones (see Figure 6a). This resulted in the acoustic source behaving more like a vibrating surface than a point source. With this configuration, the environment was reflective and reverberant, and the typical assumptions of a free field and point source were not satisfied. To accurately localize the source, the microphone array was strategically arranged to achieve the maximum aperture. Figure 6b provides a visual representation of their placements. Furthermore, to counteract mismatches in the free field model due to the reflective setting, the microphones were calibrated using a Time Difference of Arrival (TDOA) based optimization algorithm [16].



(a) Experiment Setup at Flanders Make

(b) Array Top View

Figure 6: (a) Photographic depiction of the experimental setup, showing the positioning of the microphones around the test rig, enclosed within an acrylic cover. (b) Top view of the microphone array. The cross symbols denote positions of the array microphones and the rectangle illustrates the predefined scanning area on the enclosure's upper surface, located 30 *cm* under the microphone plane.

During the experiment, the acoustic signals collected by the microphones underwent several processing stages as outlined in section 5. Initially, the raw signals were filtered, and the carrier band was identified using spectral coherence analysis Figure 7. The signals were then demodulated and subjected to a beamforming technique. This method computed the acoustic energy distribution for each six-second frame, ensuring a sufficient frequency resolution for identifying the BPFO at 114.72 Hz. A scanning area was defined as the enclosure's upper surface as illustrated in Figure 6b, due to the invalidity of the point source assumption. The experiment transitioned from a state of a healthy bearing to an outer ring fault over a duration of eight hours. The indicator was computed for each frame of the demodulated signals, representing the ratio of the peak amplitude at BPFO to a floor value in the SES as introduced in section 2. The signal with the maximum indicator, expected to provide the best SNR, was selected. The performance of this EB technique was compared against a single microphone (MTOP) placed just above the test rig and an accelerometer. The solitary microphone in this setup simulated the reference case where a single microphone is positioned close to the source for optimal SNR.

7.2 Analysis of Experimental Results

Figure 8a visualizes the Point Spread Function (PSF) of the microphone array at the BPFO of 114.72 Hz. The focus is on the circle depicted in the figure. The main lobe is visible at the center of the map, with side lobes appearing at both ends of the y-axis. This is due to the y-axis being longer than the x-axis, leading to a

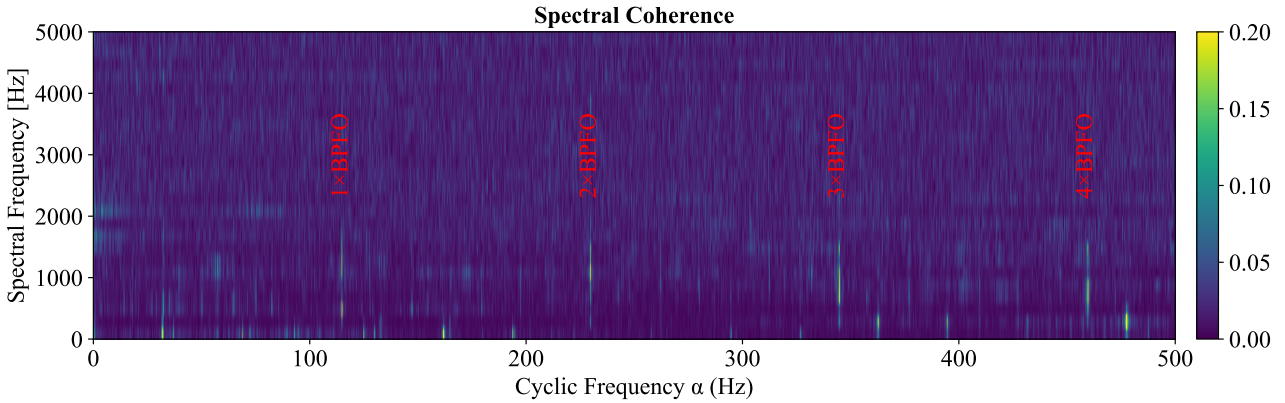
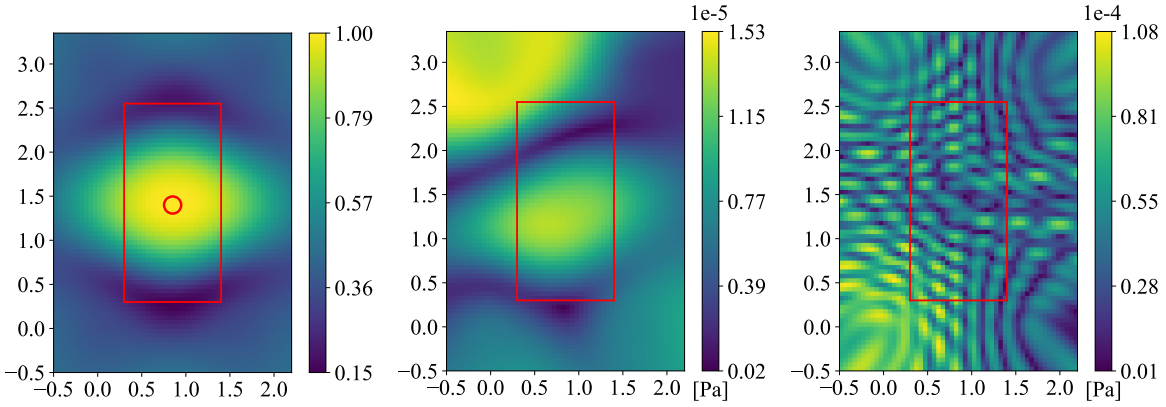


Figure 7: Spectral coherence analysis of the acoustic signals during the fault event, denoted with the BPFO at 114.72 Hz. The cyclic frequency α and the spectral frequency f are represented on the x-axis and y-axis, respectively. Vertical lines indicate the first four harmonics of the BPFO, evidencing significant periodic modulations in the signal.

mild aliasing effect within the scanned surface area. The PSF provides valuable insights into the spatial and frequency limitations of the array. In Figure 8b, the EB acoustic map is displayed during an outer ring fault. The map exhibits a lobe similar to the PSF, with a strong BPFO energy concentration at the center of the surface. It is important to note the heat zone in the upper left corner of the map, which corresponds to the location of the hydraulic pump. Contrarily, the CB map in Figure 8c presents an aliasing map at 1000 Hz, which does not provide any useful localization information.



(a) PSF at 114.72 Hz. (b) EB at 114.72 Hz. (c) CB at 1000 Hz.

Figure 8: Comparison of Acoustic Imaging Techniques.

The experimental results, presented in Figure 9, capture the transition of the bearing's state from healthy to faulty. The final 15% of the recorded frames in the dataset is plotted, with the time frame on the x-axis and the indicator value on the y-axis. For improved readability and a clearer visualization of the trend, the indicator values have been averaged over every 100 frames. We focus on this portion of the data as it crucially displays the onset and evolution of the bearing fault. The raw, unprocessed version of the indicator values, without the 100-frame averaging, is provided in the appendix for further reference. Figure 9 illustrates the performance of four different measurement techniques - EB, CB, ACC, and MTOP - over time as the bearing's condition evolves. Each technique is represented by a distinct curve: the EB by the blue curve, the CB by the orange curve, the ACC by the red curve as the ground truth, and the MTOP by the green curve. All signals, including the ACC, were subject to the same carrier band filtering and demodulation processes, with their respective indicator values computed thereafter. This consistent processing across all channels not only ensures comparability but also provides a more stable ground truth for the ACC signal.

An noteworthy observation from the results is the early increase of the indicator values for the EB, CB, and MTOP methods from frame 3700 to 3900, preceding the rise in the ACC indicator. While this is initially surpris-

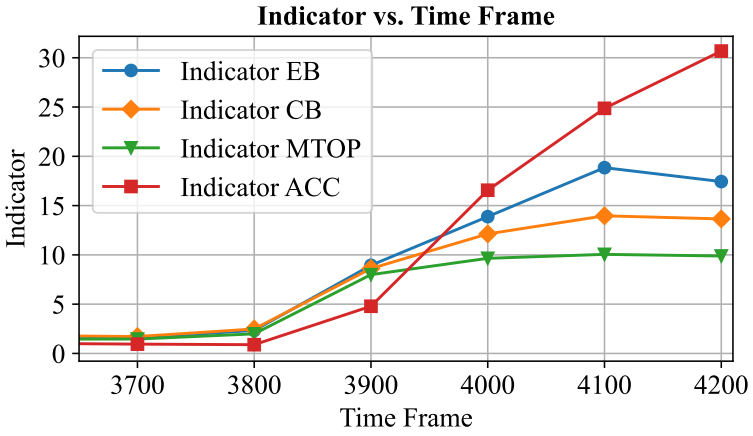


Figure 9: Experimental results depicting the progression of the bearing's condition over time. The four curves correspond to the Envelope Beamforming (EB, blue), Conventional Beamforming (CB, orange), single microphone (MTOP, green) methods, and accelerometer (ACC, red).

ing, an examination of the SES in Figure 10, where the indicators are computed as depicted in section 2, reveals the presence of the BPFO in the ACC signal as well. However, due to the presence of other peak frequencies in the SES of ACC, its indicator is not as prominent as those of the acoustic signals. A plausible explanation could be that the structure's natural frequencies are more efficient at transmitting the BPFO frequency. From frame 3800, all four channels (EB, CB, ACC, MTOP) start to exhibit elevated indicator values. The acoustic channels (EB, CB, and MTOP) experience a significant increase from an indicator value of around 5 to over 10, while the ACC indicator climbs from around 2 to over 5. It is crucial to note that following this period, the indicator values exhibit significant variability, likely due to noise and other interference during the experiment. Despite these fluctuations, the EB and CB methods consistently exhibit higher indicator values than MTOP, particularly noticeable after frame 3900. This provides strong evidence of the superior performance of the beamforming techniques in this experimental setup. This result showcases the potential of beamforming techniques for early fault detection in industrial bearings. It demonstrates the effectiveness of these methods over more traditional monitoring approaches, even in the presence of noise and interference. However, further research is needed to enhance the stability and consistency of these techniques in various operational scenarios. The figure of the raw data for the indicator values of each method are provided in the appendix for further reference.

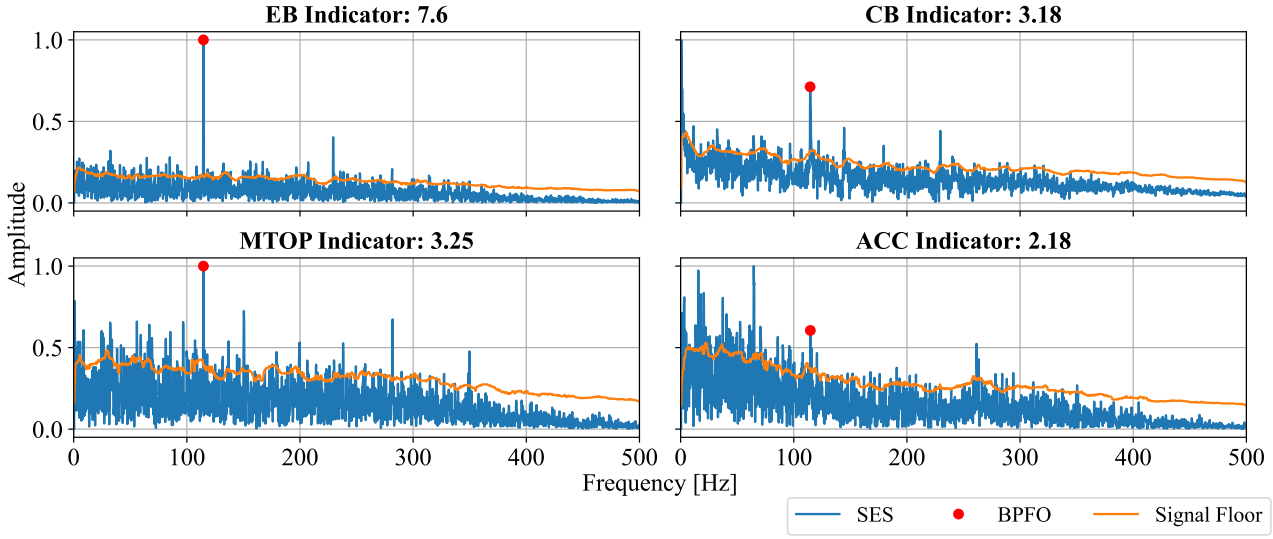


Figure 10: SES of the signals from EB, CB, MTOP and ACC methods for a representative frame where the phenomena of interest are clearly demonstrated. The respective indicator values for each method are included in the plot and the amplitude of the SES is normalized, emphasizing the relative values which are crucial for calculating the indicator. Similar patterns are observed across various frames.

8 Conclusion

In this study, we proposed using a sparse array beamforming technique for early fault detection in Rolling Element Bearings (REBs). We demonstrate the feasibility of employing Envelope Beamforming (EB), even with a reduced number of microphones, for successful fault detection based on (low) characteristic bearing frequencies.

Our approach exhibits better performance to Conventional Beamforming (CB) in terms of robustness and efficiency in a numerical simulation. Additionally, the experimental results confirmed our simulation findings, underscoring the promising performance of EB over a single microphone in identifying outer ring faults. While the vibration-based method may outperform in our chosen indicator, the non-contact nature and capability of simultaneous multiple bearing monitoring make acoustic techniques a compelling alternative. However, the broader lobe size at lower frequencies stands as a significant disadvantage of EB, making precise localization more challenging. This issue necessitates further exploration of advanced beamforming techniques to improve localization accuracy in future research.

To conclude, our study emphasizes the potential utility of sparse array beamforming techniques, specifically EB, for early fault detection in industrial bearings. The results indicate that this method could provide a more cost-effective and resource-efficient pathway towards improved operational performance and reduced maintenance costs in industrial machinery.

Acknowledgments

We gratefully acknowledge the support of the European Commission through the H2020 ETN ECO DRIVE project (GA 858018) as part of the Marie Skłodowska-Curie program. We extend our appreciation to the ACMON ICON project and Flanders Make, the strategic research centre for the manufacturing industry, which facilitated the measurements conducted in this study. This work also received support from VLAIO within the frames of the ACMON ICON project. We are deeply grateful for all the contributions that have made this research possible.

References

[1] D. Abboud, J. Antoni, S. Sieg-Zieba, and M. Eltabach. Envelope analysis of rotating machine vibrations in variable speed conditions: A comprehensive treatment. *Mechanical Systems and Signal Processing*, 84:200–226, Feb. 2017.

[2] J. Antoni, G. Xin, and N. Hamzaoui. Fast computation of the spectral correlation. *Mechanical Systems and Signal Processing*, 92:248–277, Aug. 2017.

[3] M. R. Bai, S.-S. Lan, J.-Y. Huang, Y.-C. Hsu, and H.-C. So. Audio enhancement and intelligent classification of household sound events using a sparsely deployed array. *The Journal of the Acoustical Society of America*, 147(1):11–24, Jan. 2020.

[4] M. Buzzoni, G. D'Elia, and M. Cocconcelli. A tool for validating and benchmarking signal processing techniques applied to machine diagnosis. *Mechanical Systems and Signal Processing*, 139:106618, May 2020.

[5] E. Cardenas Cabada, Q. Leclere, J. Antoni, and N. Hamzaoui. Fault detection in rotating machines with beamforming: Spatial visualization of diagnosis features. *Mechanical Systems and Signal Processing*, 97:33–43, Dec. 2017.

[6] L. Chen, Y.-S. Choy, K.-C. Tam, and C.-W. Fei. Hybrid microphone array signal processing approach for faulty wheel identification and ground impedance estimation in wheel/rail system. *Applied Acoustics*, 172:107633, Jan. 2021.

[7] A. Glowacz. Fault diagnosis of single-phase induction motor based on acoustic signals. *Mechanical Systems and Signal Processing*, 117:65–80, Feb. 2019.

[8] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mechanical Systems and Signal Processing*, 76-77:283–293, Aug. 2016.

[9] A. Mauricio, H. Denayer, and K. Gryllias. Time-domain beamformed envelope spectrum of acoustic signals for bearing diagnostics. In *Proceedings of the 30th International Conference on Noise and Vibration Engineering (ISMA 2022)*, 2022.

[10] L. Medina and E. Moreno. Envelope-beamformer location of inhomogeneities. In *Proceedings of the World Congress on Ultrasonics*, 2003.

[11] R. Merino-Martínez, P. Sijtsma, M. Snellen, T. Ahlefeldt, J. Antoni, C. J. Bahr, D. Blacodon, D. Ernst, A. Finez, S. Funke, T. F. Geyer, S. Haxter, G. Herold, X. Huang, W. M. Humphreys, Q. Leclère, A. Malgoezar, U. Michel, T. Padois, A. Pereira, C. Picard, E. Sarradj, H. Siller, D. G. Simons, and C. Spehr. A review of acoustic imaging methods using phased microphone arrays: Part of the "Aircraft Noise Generation and Assessment" Special Issue. *CEAS Aeronautical Journal*, 10(1):197–230, Mar. 2019.

[12] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—A tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, Feb. 2011.

[13] B. Razavi. *RF microelectronics*. Prentice Hall, Upper Saddle River, NJ, 2nd ed edition, 2012.

[14] E. Sarradj. Three-Dimensional Acoustic Source Mapping with Different Beamforming Steering Vector Formulations. *Advances in Acoustics and Vibration*, 2012:1–12, June 2012.

[15] P. Sijtsma, R. Merino-Martinez, A. M. Malgoezar, and M. Snellen. High-resolution CLEAN-SC: Theory and experimental validation. *International Journal of Aeroacoustics*, 16(4-5):274–298, July 2017.

[16] C. Vanwynsberghe, P. Challande, F. Ollivier, J. Marchal, and R. Marchiano. Geometric calibration of very large microphone arrays in mismatched free field. *The Journal of the Acoustical Society of America*, 145(1):215–227, Jan. 2019.

[17] R. Wang, F. Liu, F. Hou, W. Jiang, Q. Hou, and L. Yu. A Non-Contact Fault Diagnosis Method for Rolling Bearings Based on Acoustic Imaging and Convolutional Neural Networks. *IEEE Access*, 8:132761–132774, 2020.

[18] F. E. White. *Fundamentals of Acoustics* by Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, and James V. Sanders. *The Journal of the Acoustical Society of America*, 72(3):1090–1090, Sept. 1982.

[19] S. K. Yadav, K. Tyagi, B. Shah, and P. K. Kalra. Audio Signature-Based Condition Monitoring of Internal Combustion Engine Using FFT and Correlation Approach. *IEEE Transactions on Instrumentation and Measurement*, 60(4):1217–1226, Apr. 2011.

A Appendix

- Figure of the raw data for the indicator values of each method

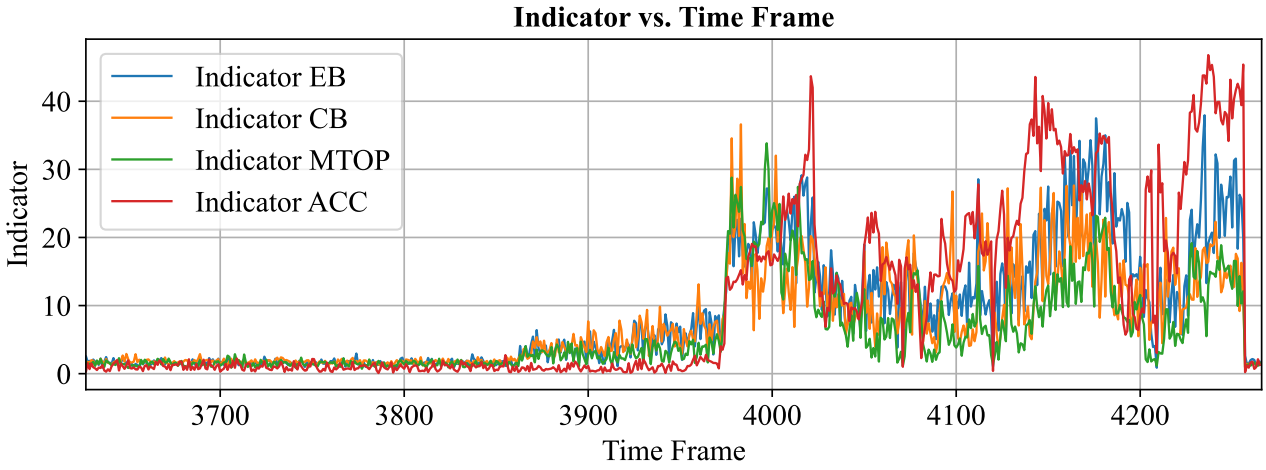


Figure 11: Experimental results depicting the progression of the bearing's condition over time. The four curves correspond to the Envelope Beamforming (EB, blue), Conventional Beamforming (CB, orange), single microphone (MTOP, green) methods, and accelerometer (ACC, red).