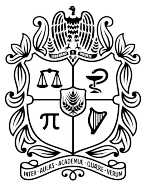


UNIVERSIDAD
NACIONAL
DE COLOMBIA

Characterization of faults in rotating machines using multivariate component analysis

Anderson Alberto Ruales Torres

Universidad Nacional de Colombia
Faculty of Engineering and Architecture
Department of Electrical, Electronic and Computing Engineering
Manizales, Colombia
2020

Characterization of faults in rotating machines using multivariate component analysis

Anderson Alberto Ruales Torres

Thesis presented as a partial request to apply for degree of:
Master in Engineering - Industrial Automation

Advisor:
Ph.D., Germán Castellanos Domínguez
Co-Advisor:
Ph.D., Óscar Cardona Morales

Research Subject:
Machine Diagnostics and Vibration Analysis
Research Group:
Signal Processing and Recognition Group

Universidad Nacional de Colombia
Faculty of Engineering and Architecture
Department of Electrical, Electronic and Computing Engineering
Manizales, Colombia
2020

Caracterización de fallas en máquinas rotativas usando análisis de componentes multivariado

Anderson Alberto Ruales Torres

Tesis presentada como requisito parcial para optar al título de:
Magister en Ingeniería - Automatización Industrial

Director:
Ph.D., Germán Castellanos Domínguez
Co-Director:
Ph.D., Óscar Cardona Morales

Línea de Investigación:
Diagnóstico de máquina y análisis de vibraciones
Grupo de Investigación:
Grupo de Control y Procesamiento Digital de Señales

Universidad Nacional de Colombia
Facultad de Ingeniería y Arquitectura
Departamento de Ingeniería Eléctrica y Electrónica
Manizales, Colombia
2020

Este trabajo está dedidaco a mi familia que fue mi fuente de apoyo constante en ésta travesía de conocimiento.

A mis compañeros y amigos de grupo que me brindaron su respaldo y conocimiento en momentos de oscuridad.

Equipado con sus cinco sentidos, el hombre explora el universo que lo rodea y a sus aventuras las llama ciencia.

Edwin Powell Hubble

Agradecimientos

Agradezco a mi padre Luis Alberto Ruales, a mi madre Fanny Torres y a mi hermana Luisa María que siempre me brindaron su apoyo incondicional y la motivación necesaria para continuar. Gracias a todas las personas del *Grupo de Control y Procesamiento Digital de señales*, en especial a su director el profesor Germán Castellanos Domínguez, quien brindo su guía y experiencia para el desarrollo de esta tesis. Quisiera expresar mi mas sincero agradecimiento a mí Co-Director y amigo Óscar Cardona Morales que compartió su conocimiento, experiencia y me guió con paciencia y dedicación durante el desarrollo de esta investigación. Agradezco a mis amigos que me acompañaron en este arduo camino. De igual manera, a todas las personas que de una u otra forma contribuyeron e hicieron posible la finalización de esta tesis.

Debo agradecer además a las instituciones que hicieron posible la realización de este trabajo por su apoyo económico: A COLCIENCIAS (año 2017 y 2018), entidad que financió el proyecto de investigación *Desarrollo de un sistema de monitoreo de condición y diagnóstico de fallas en línea de sistemas de generación de energía hidroeléctrica empleando una red de sensores inalámbricos de datos de alta resolución*. Finalmente, agradezco a la FUNDACIÓN CEIBA (año 2019), que a través del proyecto *Formación del talento humano de alto nivel para el fortalecimiento de necesidades estratégicas de CTeI del departamento de nariño*, me brindaron el apoyo financiero e hicieron posible la culminación de este trabajo de maestría.

Resumen

La presente tesis pretende desarrollar un conjunto de metodologías que permitan la extracción de características y la separación de fuentes ocultas, con el fin de diagnosticar los distintos tipos de fallos en cajas de engranajes y rodamientos. Primero, se propone la implementación de algoritmos para la detección de fallas en una caja de engranajes con 10 tipos de fallas, donde se obtuvo porcentajes de acierto por encima del 85 %, así, los resultados de clasificación muestran que la metodología empleada para la extracción de características es significativa. Segundo, se realiza un análisis de diferentes métodos de separación de fuentes ocultas, los cuales logran resaltar las frecuencias características de la falla, evidenciando que son una herramienta útil en la identificación. Finalmente, se propone un algoritmo de Análisis de Componentes Independientes basado en múltiples restricciones (*mcICA*), aprovechando la información codificada por la envolvente de la señal para localizar la falla. Por lo tanto, la metodología tratada en este documento contribuye tanto a la evaluación del estado de salud como al proceso de mantenimiento de la máquina.

Palabras clave: Rodamiento de elementos rodantes, Señales de vibración, Localización de fallas, Análisis de componentes independientes, Extracción de características, Análisis de vibraciones .

Abstract

This thesis aims to develop a set of methodologies that allow the feature extraction and blind source separation, to diagnose the different types of faults in gearboxes and bearings. First, it is proposed the implementation of fault detection algorithms in a gearbox with 10 fault types, where it is obtained success percentages above 85 %. Thus, the results of the classification show that feature extraction methodology is significant. Second, an analysis of different methods of blind source separation is realized, which highlights the characteristic frequencies of the fault, showing that they are a useful tool in the identification. Finally, multiple-constrained Independent Component Analysis (*mcICA*) is proposed, taking advantage of the information encoded by the signal envelope to fault localization. Therefore, the methodology discussed in this document provides both the evaluation of the state of health and the machine maintenance process.

Keywords: Rolling element bearing; Vibration signals; Fault localization; Independent component analysis, Feature extraction, Vibration analysis

Contents

List of Figures

List of Tables

List of symbols

Symbols with Latin letters

Symbols	Description
$\boldsymbol{A}$	Mixing matrix
$\mathcal{C}$	Cumulant matrices
$\boldsymbol{D}$	Unitary transform
$\boldsymbol{E}$	Eigen matrix
$\bar{\boldsymbol{K}}$	Kernel matrix
$\boldsymbol{P}$	Permutation matrix
$\boldsymbol{S}$	Sources matrix
$\mathcal{S}$	Short-Time Fourier Transform
$\tilde{\mathcal{S}}$	Frequency-marginal
$\boldsymbol{U}$	Rotation matrix
$\hat{\boldsymbol{U}}$	Unitary matrix
$\boldsymbol{W}$	De-mixing matrix
$\boldsymbol{X}$	Signal vibration matrix
$\hat{\boldsymbol{X}}$	Envelope signal vibration matrix
$\tilde{\boldsymbol{X}}$	Features matrix
$\boldsymbol{Y}$	Reconstructed matrix
f	Frequency
k	Discrete frequency
l	Label class
t	Continuous time
$\boldsymbol{r}$	Reference signal
$\boldsymbol{s}$	Sources

Symbols	Description
$\boldsymbol{w}$	De-mixing vector
$\boldsymbol{x}$	Signal vibration in the time
$\hat{\boldsymbol{x}}$	Envelop signal vibration in the time
$\boldsymbol{y}$	Reconstructed vector
$\boldsymbol{z}$	Whitened vector

Symbols with Greek letters

Symbols	Description
Σ	Covariance matrix
φ	Features time-domain
$\hat{\varphi}$	Features frequency-domain
$\tilde{\varphi}$	Features time-frequency domain
τ	Delay time

Subscript

Symbols	Description
f_k	k-th frequency bin
m	m-th sensors
p	p-th reference signal
q	q-th Record

Superscript

Symbols	Description
C	Features
C'	Features extraction

Symbols	Description
M	Sensors
N	Hidden sources
l	Label

Abbreviation

Abbreviation	Description
$BPFI$	Ball pass frequency, inner race
$BPFO$	Ball pass frequency, outer race
BPS	Ball spin frequency
$cICA$	Constrained Independent Component Analysis
CKA	Centered Kernel Alignment
FTF	Fundamental train frequency
GMF	Gear Mesh Frequency
IC	Independent Component
ICA	Independent Component Analysis
$mcICA$	Multiple-Constrained Independent Component Analysis
MSE	Mean Squared Error
PCA	Principal Component Analysis
REB	Rolling Element Bearings
$STFT$	Short-Time Fourier Transform

1 Preliminaries

1.1. Introduction

1.1.1. Motivation

Continuous progress in sensing capabilities coupled with the need for monitoring processes has changed the paradigm in the industry, where companies are migrating from traditional preventive maintenance strategies to programming conservative maintenance tasks and incorporating predictive maintenance concepts, which carried out only when required. Prognosis and health management systems in machinery are the foundation of any effective predictive maintenance scheme and are commonly known as condition monitoring. The main objective of this is to determine precisely the remaining useful life of the machine and its components, to provide the maintenance engineer with sufficient time to schedule a repair and purchase the components to be replaced before a catastrophic failure occurs, allowing to extend the production capacity and expand system reliability. However, there are still no systematic methodologies of how to accurately predict the remaining useful life of the machine and support decision making in spare parts management. Then, the task falls on the knowledge and experience of the maintenance engineer on the health of the machine. Therefore, there is a growing need to effectively develop and improve maintenance management systems based on automatic diagnostic and prognostic systems, that provide more reliable support to the personnel in charge.

The industrial sector in Colombia has seen the need to implement predictive maintenance plans to optimize the operation and lifetime of the machines and improve the safety of the operational personnel in charge. To this end, vibration analysis for fault identification has generated all kinds of expectations, thanks to its high cost/benefit profitability. Nevertheless, the analysis carried out does not consider the transient states of the machine where the first signs of damage are generated without considering that they depend mainly on the ability of an expert to interpret the results and whose fees are quite high. For this reason, the development of a methodology for diagnosis and prognosis of faults allows the optimization of preventive and predictive maintenance in the industry, which in turn translates into an increase in cost/benefit ratio and thus boost production. This optimization provides greater competitiveness in terms of processes and products, allowing industries to generate higher dividends in their niche market and reducing costs for unnecessary machine stops. It expects that an improvement in production processes would increase the labor supply, as a result of

the investment of companies in new markets and products.

1.1.2. Problem statement

Condition-based monitoring uses the information collected from the machine to recommend maintenance decisions that are called *Diagnostics*, which deals with the detection, identification, and localization of faults whenever it occurs. In a supervised system, error detection indicates if a breakdown occurs and if it exists, fault identification determines the nature of the damage. The vibration analysis is the most used technique in diagnostics because it allows finding the different sources that excite the machine. However, non-stationary vibration signals generated by the changes in load and/or speed of the machine complicate the identification of the source. At present, existing techniques that require considerable experience to achieve success in its implementation (Haidong et al., 2018). There is a need for the development of simple approaches to diagnose the state of the machine quickly. Commonly, we recognize three necessary steps to fault diagnosis, such as *i)* fault detection that means identifying the health state of the machine; *ii)* fault identification, i.e., given that the bearing or gearbox has a fault, we due identifying the fault type, and *iii)* fault prognosis, that estimate the current remaining useful life of the equipment (Jardine et al., 2006).
The differentiation of the faults is complicated since the raw signal in a time segment concentrates the entire vibration input of the machine. In fault detection, multiple techniques have been developed for the characterization of the vibration signals that come from bearings and gearboxes. We find many features in the time and frequency domain that allow extracting information of the fault, but these features have different advantages and limitations. For example, Root Mean Square is continuously growing with the development of bearing failures. However, it is challenging to discover impulsive defects. At the initial stage of fault, impulses tend to decrease with fault development are presented. In these cases, the kurtosis factor is sensitive to such impulses, but it shows poor performance to determine the progressive growth of fault (Shen et al., 2012). Among the characterization techniques in the frequency domain is the Fast Fourier Transform (FFT), which differentiates the types of faults from the amplitude and frequency of each spectral component (Atoui et al., 2013). Nevertheless, due to the variability of machine operating conditions and non-stationarity of signals, characterizations are used in the time-frequency domain. Wigner-Ville distribution, short-time Fourier transform, and Wavelet analysis are the time-frequency methods. In Cerrada et al. (2019), it is proposed to use the Wavelet Packet Transform (WPT), where the characteristics are the wavelet coefficients obtained from the vibration signal. Also, we find continuous wavelet transform (CWT) and discrete wavelet transform (DWT)that allow decomposing the signs at several levels. However, the wavelet basis must be selected appropriately because an inappropriate wavelet basis will directly influence the decomposition (Chen et al., 2016; Tao et al., 2019). Besides, there are statistical features in time and frequency domain that includes statistical moments such as mean, variance, skewness, and

kurtosis that extract more information about faults (Lei et al., 2008, 2010). In Singh and Zhao (2016) decomposes the raw signal into intrinsic mode function (IMF) using empirical mode decomposition (EMD) to identify fault information that can contain multiple sources of vibration and noise. This method requires a priori knowledge of the faults. However, a broad feature set includes irrelevant or redundant features directly affecting the performance of the classifier. Then, a feature extraction stage becomes necessary to improve the classification accuracy. Exists different methods of feature extraction; among them, we find Isomap, Maximum Variance Unfolding, Locally Linear Embedding, Laplacian Eigenmaps, among other methods (Van Der Maaten et al., 2009). In Jing et al. (2017) is present the Deep learning, also known as deep neuronal networks (DNN), its extracts information by a hierarchical structure with multiple neural layers but the tuning for free parameters and computational cost is complicated. In many cases, the fault characteristics are masked by a noise or other signal components presented in the standard machine operation. Therefore, it is important the study of latent variable models.

A latent variable is just an unobservable random variable. Thus, latent variable models allow extracting hidden patterns that may be related to fault behavior. There are several latent variable models. In our case, we will focus on models that will enable blind source separation (BSS). BSS is one of the popular methods for separating sources of interest. As is explain in Hyvärinen et al. (2004), several methods are very closely related to BSS. The most commonly used method is Independent Component Analysis (ICA), given that it does not need the prior knowledge of the problem, and it is enough to assume statistical independence between the sources. Thus, we can identify fault characteristics by source separation, where sources represent a force that influences a machine component known as a fault in the mechanical systems (Yang and Nagarajaiah, 2014). Several ICA-based methods have been developed. Hyvärinen and Oja (2000) proposed the FastICA algorithm, and it is based on a fixed-point iteration scheme for finding a maximum of the non-Gaussianity. In Žvokelj et al. (2016) has utilized Ensemble Empirical Mode Decomposition and ICA for bearing fault detection, but this method requires a stage of characterization, increasing computational cost. In Wu and Xiong (2019) presented a technique based on Cubic spline interpolation Intrinsic Time scale Decomposition (CITD) algorithm and FastICA algorithm, where the FastICA serves to reduce the noise of the reconstructed signal. Other authors proposed the SOBI algorithm based on the joint approximate diagonalization of multiple time-delayed correlation matrices (Belouchrani et al., 1997; Choi et al., 2002), and Cardoso (1999) developed the JADE algorithm, which is specifically a statistic based technique. JADE uses the joint diagonalization of a maximal set guarantees source identifiability. Miao et al. (2020) proposed a fault separation method which it is base on the median filter and JADE, but the filter can remove components from the fault and generate an error in reconstructs the characteristic signal. Besides, Popescu (2010) presented a combination of JADE and change detection in source signals, where the raw signals are transformed into the space of independent sources for the reduced number of components and applied the change detection methods on the

estimated sources. Nevertheless, ICA has two ambiguities (scaling and permutation) that limit its usage, it means that the IC amplitudes change on each algorithm iteration, and the appropriate IC order cannot be determined.

Mechanical signals are characterized by excessive complexity so that when using BSS methods, we must take into account some difficulties such as a mixture of vibrations of the convolutive type, an unknown number of individual sources in the mixture, among others, that compromise their viability (Antoni, 2005). In that regard, some authors had suggested modifying the BSS paradigm into extracting a single signal or component that encloses the bearing fault dynamic, also called *blind signal extraction* (Cardona-Morales et al., 2018; Smith and Randall, 2015a). Then, BSS does not look for the sources themselves but identifies the different components excited by the different sources from a blind deconvolution to identify the sources. In contrast, BSE seeks to separate the different components that exhibit some behavior in the vibration signal. Among the multiple methods adopted in the state-of-the-art related to BSE, constraint independent component analysis (cICA) or ICA with reference (Lu and Rajapakse, 2006; Zhang, 2008) has been designed to extract a single component (from the mixture) with a characteristic pattern. In Wang et al. (2011), the authors, using cICA for fault detection in Rolling element bearing (REB) given that it is one of the most critical mechanisms in rotating machinery, and it is failure implies, in many cases, substantial economic losses and catastrophic incidents in the industry. Jing et al. (2014) used the properties of the cICA algorithm to extract the fault feature from the vibration signal with interference noise, given that fault signal present periodic impact and a non-Gaussian signal, and in Yang et al. (2018), cICA was employed over non-stationary vibration signals where the machine speed is fluctuating. Other authors proposed to extract the fault feature in a single-channel signal using cICA (Leng et al., 2018). This method is based on ensemble empirical mode decomposition (EEMD), which meets the requirement of cICA algorithm, but the inclusion of EEMD significantly increases signal processing time. Nonetheless, the bearing fault localization is still an open issue that raises its relevance nowadays, since technology development allows think into sensor networks that provide simultaneous monitoring.

1.2. Objetives

Based on the problems mentioned above, this work is developed based on a general objective, which is divided into three specific objectives, as follows:

1.2.1. General Objetive

Develop a fault diagnostic methodology in rotating machinery based on latent variable models that allow failure identification and localization.

1.2.2. Specific Objetive

- Implement a methodology for multiple fault detection based on latent variable methods, which includes several machine speed conditions and multi-channel data.
- Analyze fault identification methods through the comparison of several types of blind source separation techniques applied to rotating machines with multi-channel data.
- Develop a methodology for fault localization by using a blind source separation technique that considers non-stationary failure patterns.

1.3. Scope and Structure of the thesis

This section provides an overview of the main terms and concepts within the framework of vibration analysis to give the reader the limitations and scopes of this work. For the proper development of the document, the following is the definition of some of the essential concepts.

- **Fault** is defined as an atypical behavior of at least one characteristic property of a variable. Therefore, the fault can cause a malfunction or system failure (Isermann, 2005).
- **Fault Diagnosis** consists of determining the characteristics of fault, with the most details such as the fault size, location, and time of detection, among others (Isermann, 2006).
- **Fault Detection** determines the occurrence of a fault in the monitored system, and it allows taking preventative action and avoiding possible system failure (Isermann, 2005).
- **Fault Identification** determines the magnitude (size) and type of fault. We can compare the nominal values with the estimated values and thus isolate the faults (Isermann, 2006).
- **Fault Localization** is to select among a finite number of potential locations. This problem is considering as a classification problem that precise damage localization over infinite possible damage locations (Kopsaftopoulos and Fassois, 2013).
- **Multi-channel Data** are mixed signals from an unknown multiple-input-multiple-output linear system (channel) (Yellin and Weinstein, 1994).
- **Sources**, in this case, will be vibration sources that it is generated by machine operation. These events are forces and motion such as shaft imbalance, impacts (e.g., due to bearing faults), fluctuating forces during gear mesh (as a result of small variations

in the load and/or speed of the machine), electromagnetic forces and pressure- and flow-related sources (Popescu, 2010).

- **Blind Source Separtion** consists of the recovering of the various independent sources exciting a system given only the measurements of the outputs of that system (Popescu, 2010).

In this work is proposing several methodologies based on different latent variable models for fault identification and localization in the rotative machine. The first approach is to extract features using Principal Component Analysis (PCA) and Center Kernel Alignment (CKA) on different domains (time, frequency, and time-frequency). Besides, we use the nearest neighbor classifier to discriminate signal states normal and fault. The multiclass classification problem seeks to differentiate between all possible faults on gear fault vibration signals. The second approach is related to the identification of fault characteristics. Thus, we analyze a set of independent component analysis methods to identify the presence of a fault pattern. The third approach develops a methodology that allows identifying the type of fault by using a multiple-constrained independent component analysis algorithm. Additionally, we present a fault localization stage using proximity measurements to associate the reconstructed independent component (IC) with bearing signals, which turn out to be efficient at locating the fault.

2 Multiple fault detection based on latent variable methods

2.1. Introduction

Diagnosis of rotatory machines and early fault detection has a significant impact on the industry because of reducing costs and accidents. Maintenance procedures should be performed following strict preventive schedules or earlier when a fault is detected (Medina et al., 2017). The gearboxes are components involved in mechanical transmission. Besides, they play a crucial role in the power transmission system, as they are designed for speed and torque conversion. These type of machines, they have complex configurations given that connecting two or more (parallel) shafts through coupled gears with different sizes. In consequence, detection and diagnosis of faults are challenging, and it usually requires analysis of vibration signals (Jiang et al., 2018; Igba et al., 2016).

The vibration analysis is used for fault diagnosis since its possible from the vibration signals to obtain important information on the faults. At present, existing techniques that require considerable experience to achieve success in its implementation (Haidong et al., 2018). There is a need for the development of simple approaches to diagnose the state of the machine quickly. Commonly, we recognize three basic steps to fault diagnosis such as *i)* fault detection, i.e., identify the health state of the machine; *ii)* fault classification, i.e., given that the gearbox has a fault, to identify the type of fault, and *iii)* fault prognosis, i.e., to estimate the current remaining useful life of the equipment. Therefore, we will concentrate on the first two steps due to the high demand for reliably diagnosis techniques on the health of the machine (Jardine et al., 2006).

For fault diagnosis, we find many features in the time, frequency, and time-frequency domains (Li et al., 2015; Chen et al., 2016; Tao et al., 2019). However, a broad feature set contains irrelevant or redundant features directly affecting the performance of the classifier. Then, a feature extraction stage becomes necessary to improve the classification accuracy. There are different methods of feature extraction; among them, we find Isomap, Maximum Variance Unfolding, Locally Linear Embedding, Laplacian Eigenmaps, among other methods (Van Der Maaten et al., 2009). In (Jing et al., 2017) is present the Deep learning, also known as deep neuronal networks (DNN), its extracts information by a hierarchical structure with multiple neural layers but the tuning for free parameters and cost computational is complex.

This chapter discusses two methods of feature extraction on different domains (time, frequency, and time-frequency) for gear fault diagnosis on vibration signals. These signals were collected under various operating loads and different speeds. Then, we estimate a pool of statistical features commonly used to characterize faults, and we use Principal Component Analysis (PCA) and Center Kernel Alignment (CKA) for feature extraction. As a result, we obtained an increase in the classification accuracy, using a k nearest neighbors classifier.
The agenda is as follows: Firstly, Section 2.2 describes the mathematical background of features estimation in the different domains, and it presents the methods used for the feature extraction. The developed experiments and achieved results are described in Section 4.3. Finally, the discussion is provided in Section 2.4.

2.2. Methods

2.2.1. Feature estimation

Time-domain and frequency-domain features

In Lei et al. (2010), are generated multidimensional feature sets including time-domain and frequency-domain to reveal gear health conditions. In gearbox, the time-domain signal usually changes when damage occurs in gear. Then, the standard gear signal may be different both its amplitude and distribution of the fault signal (Lei et al., 2008). We calculate several features in the time domain, such as the mean signal (φ_1) is the average value of the signal. Standard deviation (φ_2) is the root square of the dispersion of a signal around their reference mean value, and root mean square (φ_4) reflects the vibration amplitude and energy in time-domain. Moreover, we estimate the maximum values (φ_5) that are the peak value of the signal, skewness (φ_6) quantifies the asymmetry behavior of signal and kurtosis (φ_7) quantifies the peak value of the probability density function (Tao et al., 2019). The crest factor (φ_8) calculates how much impact occurs, and shape factor (φ_{10}) is a value that is affected by a shape object, among others, may be used to represent the time series distribution of the signal in the time-domain (Caesarendra and Tjahjowidodo, 2017). Besides, the frequency-domain feature extraction may contain relevant information that is not present in the time-domain. We estimate the mean frequency ($\hat{\varphi}_1$) that indicates the vibration energy. Besides, we calculate the variance frequency ($\hat{\varphi}_2$), skewness ($\hat{\varphi}_3$), kurtosis ($\hat{\varphi}_4$). Also, central frequency ($\hat{\varphi}_5$) shows the position changes of the main frequencies, standard deviation frequency ($\hat{\varphi}_6$) describe the convergence degree of the spectrum power and root mean square frequency ($\hat{\varphi}_7$), among others statistical features in the frequency-domain (Yan and Jia, 2018).
We estimated the suggested statistical features directly from the time series vector $\boldsymbol{x}_m^l \in \mathbb{R}^T$

where $m = [1, \ldots, M]$ is the register collect vibration signals from sensors installed on gearbox, $l = [1, \ldots, L]$ is the label for the different fault conditions and T corresponds to whole time period. In addition, we compute the features in frequency-domain by the Fourier Transform of the time series, $\boldsymbol{\Psi}(k) = \mathcal{F}\{\boldsymbol{x}(t)\}$. The estimated features are shown in Table 2-1, being f_k, $(k = 1, \ldots, K)$, the k-th frequency bin.

Table 2-1: Extracted statistical features in time (left) and frequency domain (right), thereafter Lei et al. (2008, 2010).

$\varphi_1 = \frac{\sum_{t=1}^{T} \boldsymbol{x}(t)}{T}$	$\hat{\varphi}_1 = \frac{\sum_{k=1}^{K} \boldsymbol{\Psi}(k)}{K}$
$\varphi_2 = \sqrt{\frac{\sum_{t=1}^{T} (\boldsymbol{x}(t)-\varphi_1)^2}{T-1}}$	$\hat{\varphi}_2 = \frac{\sum_{k=1}^{K} (\boldsymbol{\Psi}(k)-\hat{\varphi}_1)^2}{K-1}$
$\varphi_3 = \left(\frac{\sum_{t=1}^{T} \sqrt{\lvert\boldsymbol{x}(t)\rvert}}{T}\right)^2$	$\hat{\varphi}_3 = \frac{\sum_{k=1}^{K} (\boldsymbol{\Psi}(k)-\hat{\varphi}_1)^3}{K(\sqrt{\hat{\varphi}_2})^3}$
$\varphi_4 = \sqrt{\frac{\sum_{t=1}^{T} (\boldsymbol{x}(t))^2}{T}}$	$\hat{\varphi}_4 = \frac{\sum_{k=1}^{K} (\boldsymbol{\Psi}(k)-\hat{\varphi}_1)^4}{K\hat{\varphi}_2^2}$
$\varphi_5 = \max \lvert\boldsymbol{x}(t)\rvert$	$\hat{\varphi}_5 = \frac{\sum_{k=1}^{K} f_k \boldsymbol{\Psi}(k)}{\sum_{k=1}^{K} \boldsymbol{\Psi}(k)}$
$\varphi_6 = \frac{\sum_{t=1}^{T} (\boldsymbol{x}(t)-\varphi_1)^3}{(T-1)\varphi_2^3}$	$\hat{\varphi}_6 = \sqrt{\frac{\sum_{k=1}^{K} (f_k-\hat{\varphi}_5)^2 \boldsymbol{\Psi}(k)}{K}}$
$\varphi_7 = \frac{\sum_{t=1}^{T} (\boldsymbol{x}(t)-\varphi_1)^4}{(T-1)\varphi_2^4}$	$\hat{\varphi}_7 = \sqrt{\frac{\sum_{k=1}^{K} f_k^2 \boldsymbol{\Psi}(k)}{\sum_{k=1}^{K} \boldsymbol{\Psi}(k)}}$
$\varphi_8 = \frac{\varphi_5}{\varphi_4}$	$\hat{\varphi}_8 = \sqrt{\frac{\sum_{k=1}^{K} f_k^4 \boldsymbol{\Psi}(k)}{\sum_{k=1}^{K} f_k^2 \boldsymbol{\Psi}(k)}}$
$\varphi_9 = \frac{\varphi_5}{\varphi_3}$	$\hat{\varphi}_9 = \frac{\sum_{k=1}^{K} f_k^2 \boldsymbol{\Psi}(k)}{\sqrt{\sum_{k=1}^{K} \boldsymbol{\Psi}(k) \sum_{k=1}^{K} f_k^4 \boldsymbol{\Psi}(k)}}$
$\varphi_{10} = \frac{\varphi_4}{\frac{1}{T}\sum_{t=1}^{T} \lvert\boldsymbol{x}(t)\rvert}$	$\hat{\varphi}_{10} = \frac{\hat{\varphi}_6}{\hat{\varphi}_5}$
$\varphi_{11} = \frac{\varphi_5}{\frac{1}{T}\sum_{t=1}^{T} \lvert\boldsymbol{x}(t)\rvert}$	$\hat{\varphi}_{11} = \frac{\sum_{k=1}^{K} (f_k-\hat{\varphi}_5)^3 \boldsymbol{\Psi}(k)}{K\hat{\varphi}_6^3}$
	$\hat{\varphi}_{12} = \frac{\sum_{k=1}^{K} (f_k-\hat{\varphi}_5)^4 \boldsymbol{\Psi}(k)}{K\hat{\varphi}_6^4}$
	$\hat{\varphi}_{13} = \frac{\sum_{k=1}^{K} (f_k-\hat{\varphi}_5)^{1/2} \boldsymbol{\Psi}(k)}{K\sqrt{\hat{\varphi}_6}}$

Time-frequency domain features

We can consider that the vibration signal has non-stationary character, due to load and variable speed regimes so that it is necessary to broaden the spectrum of analysis techniques.

We seek to capture the dynamics of the vibration signal in its frequency components time-domain; for this, we use time-frequency representations. There are different types of time-frequency representations, which offer different interpretations. For this work, we select the spectrogram that allows an adequate interpretation of the energy concentration based on time and frequency. The spectrogram is a quadratic time-frequency representation that is part of the Short-Time Fourier Transform (STFT), then defined as:

$$\mathcal{S}(t,k) = \left| \int_{-\infty}^{\infty} \boldsymbol{x}(\tau) g(\tau - t) e^{-j2\pi\tau f} d\tau \right|^2 \tag{2-1}$$

where $g(t)$ represent a narrow window function centered at $t = 0$ and τ is the delay times in terms of the STFT. Now, we computed the frequency-marginal to find the frequencies that have the most energy over time (seen Eq. (2-2)).

$$\tilde{\mathcal{S}}(k) = \mathbb{E}\left\{ |\mathcal{S}(t,k)|^2 \right\} \tag{2-2}$$

where $\mathcal{S} \in \mathbb{R}^K$ is the frequency-marginal; thus, we used the following features statistical of Table **2-2** on the frequency-marginal, the mean ($\tilde{\varphi}_1$), standard deviation ($\tilde{\varphi}_2$), kurtosis ($\tilde{\varphi}_3$) and root mean square ($\tilde{\varphi}_4$).

Table 2-2: Extracted statistical features in time-frequency domain.

$\tilde{\varphi}_1 = \frac{\sum_{k=1}^{K} \tilde{\mathcal{S}}(k)}{K}$
$\tilde{\varphi}_2 = \sqrt{\frac{\sum_{k=1}^{K} (\tilde{\mathcal{S}}(k) - \tilde{\varphi}_1)^2}{K-1}}$
$\tilde{\varphi}_3 = \sqrt{\frac{\sum_{k=1}^{K} (\tilde{\mathcal{S}}(k))^2}{K}}$
$\tilde{\varphi}_4 = \frac{\sum_{k=1}^{K} (\tilde{\mathcal{S}}(k) - \tilde{\varphi}_1)^4}{(K-1)\tilde{\varphi}_2^4}$

Finally, we build a feature matrix $\tilde{\boldsymbol{X}} \in \mathbb{R}^{M \times C}$ that holds the feature vector for all M, where $\tilde{\boldsymbol{X}} = \left[\tilde{\boldsymbol{x}}_1^1, \ldots, \tilde{\boldsymbol{x}}_m^l, \ldots, \tilde{\boldsymbol{x}}_M^L \right]^\top$ comprises all the concatenated estimated features vectors, such that $\tilde{\boldsymbol{x}}_m^l = [\varphi_1, \ldots, \varphi_{11}, \ldots, \hat{\varphi}_1, \ldots, \hat{\varphi}_{13}, \ldots, \tilde{\varphi}_1, \ldots, \tilde{\varphi}_4] \in \mathbb{R}^C$.

2.2.2. Latent-based relevance analysis of estimation feature

Unsupervised dimension reduction

Provided a feature matrix, $\tilde{\boldsymbol{X}}$, unsupervised linear reduction can be carried out through a basic latent variable model that assumes an approximate decomposition of linearly uncorre-

lated variables $\hat{\boldsymbol{X}} \simeq \tilde{\boldsymbol{X}}\boldsymbol{A}$, where $\boldsymbol{A} \subset \mathbb{R}^{C \times C'}$ is the premultiplier, projecting onto an orthogonal subspace with a lower dimensionality $C' \leq C$, as below:

$$\boldsymbol{A}^* = \min_{\boldsymbol{A}} \left\{ \|\hat{\boldsymbol{X}} - \tilde{\boldsymbol{X}}\boldsymbol{A}\|_2^2 \right\} = \max_{\boldsymbol{A}} \left\{ \boldsymbol{A}^\top \boldsymbol{\Sigma}^{\boldsymbol{X}} \boldsymbol{A} \right\} \tag{2-3a}$$

$$\text{solving: } \boldsymbol{\Sigma}^{\boldsymbol{X}} \boldsymbol{A} = \boldsymbol{\nu} \boldsymbol{A} \tag{2-3b}$$

where $\boldsymbol{\Sigma}^{\tilde{\boldsymbol{X}}} = \mathbb{E}\left\{ \tilde{\boldsymbol{X}} \tilde{\boldsymbol{X}}^\top \right\}$ is the covariance matrix and $\boldsymbol{\nu} \subset \mathbb{R}^C$ is the vector, holding all eigenvalues retrieved from the characteristic equation of $\boldsymbol{\Sigma}^{\tilde{\boldsymbol{X}}} \subset \mathbb{R}^{C \times C}$ in Eq. (2-3b) (*Principal Component Analysis* – PCA). Notation $\|\cdot\|_p$ stands for L_p-norm.

Supervised dimension reduction

In this case, we assess the relevance analysis by matching of the labeled feature set $\{\tilde{\boldsymbol{X}}^l\}$ and its provided label set Λ, namely, we measure the projected similarity between the corresponding kernel matrices $\boldsymbol{K}^{\tilde{X}}$ and $\boldsymbol{K}^{\Lambda}$. Therefore, instead of using the optimizing rule in Eqs. (2-3a) and (2-3b) that neglects the label information, the supervised transformation matrix $\boldsymbol{A}^*$ is computed through the kernel-based learner as developed in Alvarez-Meza et al. (2017):

$$\boldsymbol{A}^* = \arg \max_{\boldsymbol{A}} \log \left(\mu(\bar{\boldsymbol{K}}^{\tilde{X}}(\boldsymbol{A}), \bar{\boldsymbol{K}}^{\Lambda}) \right), \tag{2-4}$$

where the matching metric $\mu(\cdot,\cdot;) \in \mathbb{R}^+$ (termed *Centered Kernel Alignment* - CKA) is the normalized inner product of both kernel matrices, estimated as follows:

$$\mu(\bar{\boldsymbol{K}}^{X}, \bar{\boldsymbol{K}}^{\Lambda}) = \langle \bar{\boldsymbol{K}}^{X}, \bar{\boldsymbol{K}}^{\Lambda} \rangle_{\mathtt{F}} \Big/ \sqrt{\langle \bar{\boldsymbol{K}}^{X}, \bar{\boldsymbol{K}}^{X} \rangle_{\mathtt{F}} \langle \bar{\boldsymbol{K}}^{\Lambda}, \bar{\boldsymbol{K}}^{\Lambda} \rangle_{\mathtt{F}}} \tag{2-5}$$

where notation $\langle \cdot, \cdot \rangle_{\mathtt{F}}$ stands for the Frobenius-based matrix inner product, $\bar{\boldsymbol{K}} = \tilde{\boldsymbol{I}} \boldsymbol{K} \tilde{\boldsymbol{I}}$ is the centered version of the kernel matrix $\boldsymbol{K}^{\Pi} = [\kappa^{\Pi}(\boldsymbol{\pi}, \boldsymbol{\pi}') : \forall \boldsymbol{\pi}, \boldsymbol{\pi}']$ (with $\Pi = \{\boldsymbol{X}, \Lambda\}$) that holds the pairwise similarity between vector samples $\boldsymbol{\pi}, \boldsymbol{\pi}' \subset \Pi$ embedded in a Hilbert space, $\tilde{\boldsymbol{I}} = \boldsymbol{I} - N^{-1}\boldsymbol{1}^\top \boldsymbol{1}$ is the empirical centering matrix, and $\boldsymbol{1} \subset \mathbb{R}^N$ is the all-ones vector.

2.3. Experimental Set-Up

The purpose comprises the following stages (see Fig. **2-1**): *i)* computing the statistical features of measured signal vibrations; *ii)* estimated latent components using PCA and CKA; And *iii)* classification of faults is computed through a 10-fold cross-validation procedure with partition 80 % of training and 20 % of testing, using the k nearest neighbor (k-nn) classifier. We use as baseline the performance obtained from the estimated features.

The proposed method is validated on a gearbox fault diagnosis test-rig fabricated by the lab of the Universidad Politécnica Salesiana, Ecuador. The system includes a 1.1 kW motor powered by three-phase 220 V at 60 Hz. The torque motion is transmitted into a gearbox, where

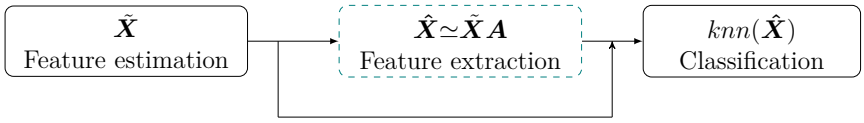


Figure 2-1: Proposed approach for fault classification using PCA and CKA.

several gear fault configuration are assembled. The output shaft of the gearbox was connected with an electromagnetic torque break ROSATI, 8.83 kW through a belt transmission. The load of the torque break can be manually adjusted by a controller GEN 100-15-IS510, which used three different load (No load, 10 V, 30 V). Variable-frequency drive was used to generate constant speeds (8 Hz, 12 Hz, 15 Hz) and variable speed (5-12 Hz, 12-18 Hz, 8-15 Hz). Additionally, an accelerometer PCB ICP 353C03 was mounted on the top of the gearbox to collect vibration signals, which were sent to a laptop via a data acquisition box NI cDAQ-9234. The data sampling rate was adjusted to 50 kHz, and collecting one vibration with a length of 10 second. Two spur gears (number of teeth $Gear1 = 53$, and $Gear2 = 80$) were installed on the input and the output shafts of the gearbox (Cerrada et al., 2016). The gearbox was configured with 10 different fault modes f1-f10 as shown in Table **2-3**. For each fault are measured 90 vibration signals.

Table 2-3: Gear fault conditions (Cerrada et al., 2019).

Label	Description
f1	Healthy pinion, healthy gear
f2	Pinion tooth chaffing, healthy gear
f3	Pinion tooth wear, healthy gear
f4	25 % pinion tooth breakage, healthy gear
f5	50 % pinion tooth breakage, healthy gear
f6	100 % pinion tooth breakage, healthy gear
f7	Healthy pinion, 25 % gear crack
f8	Healthy pinion, 100 % gear crack
f9	Healthy pinion, 50 % gear chaffing
f10	25 % pinion tooth breakage, 25 % gear crack

An incipient fault is an essential condition to be diagnosed in an industrial application since it is a fault just beginning to show symptoms (f2 and f3). Fault f1 is the normal condition, f4, f5, f7, f9 are moderate faults, f6 and f8 are severe faults, and f10 is multiple faults (Cerrada et al., 2016).

To capture the intrinsic dynamics of the signal under non-stationarity conditions for load and variable speed regimes, we use the spectrogram (Cardona-Morales, 2011). In Figs. **2-2** and **2-3** displays the time-frequency representation of downsampled version computed by the spectrogram with a 8192 frequency bins, 50 % overlapping and a Hamming window of 512 samples for ten fault conditions with speeds variable. In the upper right is presented to zoom on spectrogram which allows to visualize the low frequencies. In addition, the marginal frequency is shown that reveals the changes in the amplitudes along the frequencies considered.

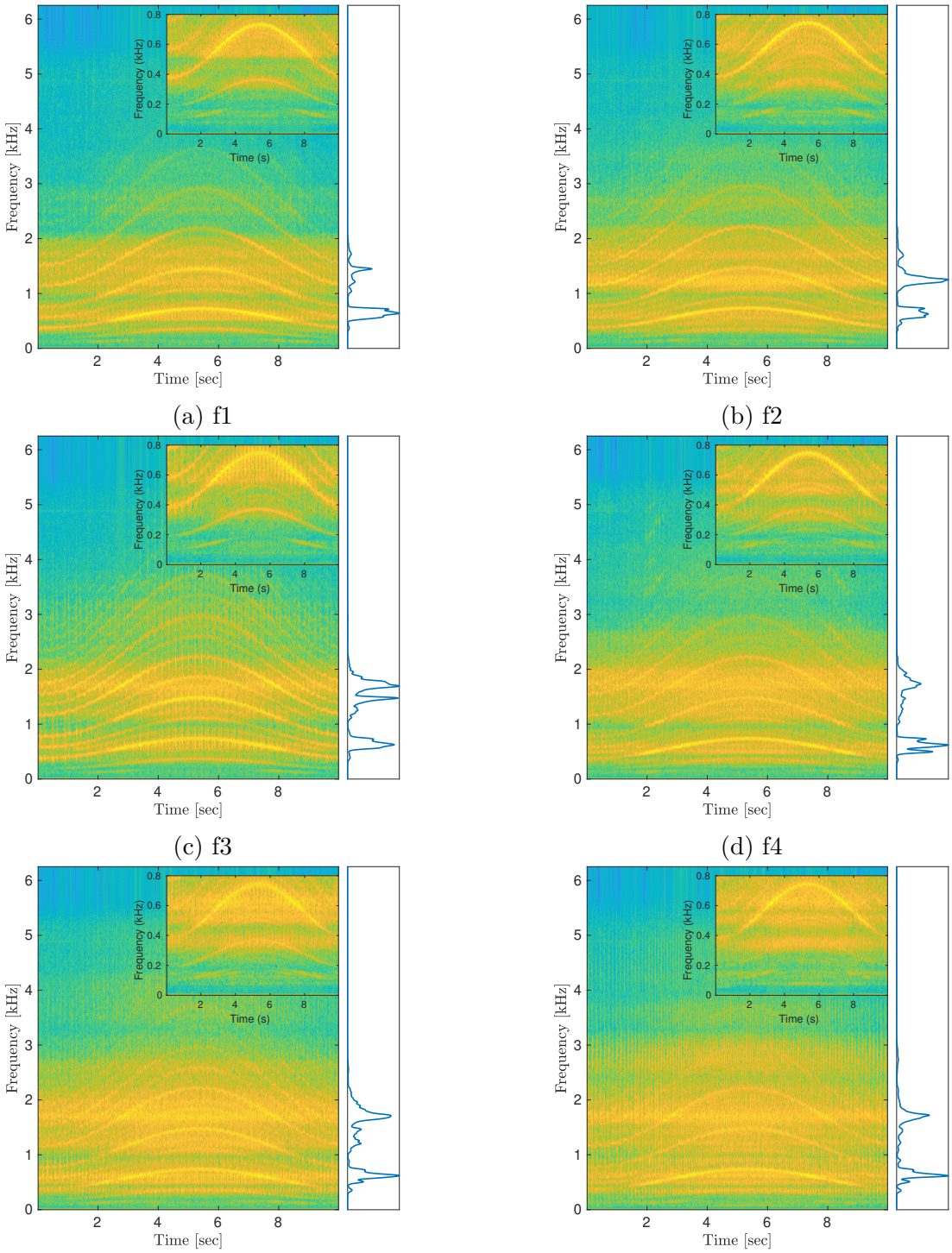
A gearbox generates quite prominent peaks in the gear mesh frequency (GMF), which is the number of teeth of the gears by the speed of rotation of the gear. In general, We note that GMF is contaminated by the noise that can be generated by the power grid. The GMF is $Gear1 = 399.1\,kHz$, which is seen with greater intensity in the zoom because the constant noise is mixed and generates a signal with greater power. It should be said that these components are not easy to separate and can lead to errors in the fault classification. Also, the constant noise components are present in the marginal frequency. For f3, f4 f5 and f6, the harmonics of the gear frequency are increasingly evident because of the degradation of the fault increases. In particular, f3 has a cyclic fault behavior because it is where a gear tooth fault begins to occur. However, for f4 and f5 the blows fade and reappear in f6. This fault process is particular to rolling element bearing since it presents a physical clearance that makes the fault very noticeable when it occurs for the first time, and towards the end of the bearing's useful life where the damage is fatal (Smith and Randall, 2015a). Besides, f7 to f10 present combinations of faults (gear and pinion). Nevertheless, the marginal frequency responds directly to the behavior of the constant noise and not of the frequencies of the gears as seen in Fig. **2-3**.

In this work, we divide the experiment in #1 for the speeds constant, #2 for the speeds variable, and #3 both speeds. We selected the number of neighbors heuristically, where we tested the values of 3, 5, 7 and 9 neighbors for the estimated characteristics. In Table **2-4**, we show that the highest performance is achieved by $k = 3$. Therefore, we set-up the classifier with 3 nearest neighbors for the development of the following experiments.

Table 2-4: Accuracy of classification for different neighbors.

Experiment (%)	$k = 3$	$k = 5$	$k = 7$	$k = 9$
#1	**91.3 ±1.9**	85.9 ±2.8	77.2 ±2.6	69.7 ±3.1
#2	**84.9 ±4.0**	82.7 ±3.2	74.2 ±3.3	70.1 ±2.4
#3	**88.7 ±1.5**	83.3 ±2.2	76.5 ±2.2	71.9 ±1.5

As a result, CKA-knn allows achieving the best performance for all experiments, decreasing

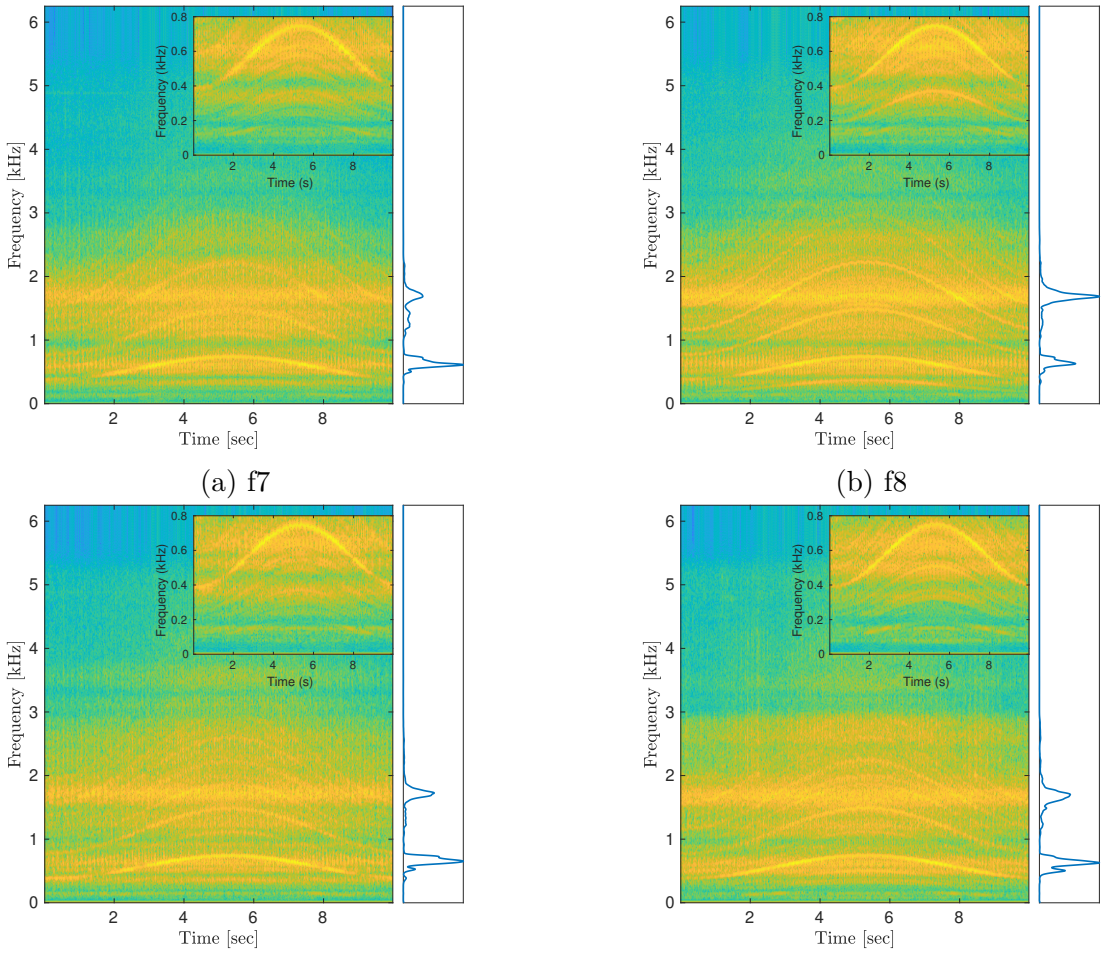


(a) f1 (b) f2

(c) f3 (d) f4

(e) f5 (f) f6

Figure 2-2: Time-frequency representation (spectrogram with hamming window, 50 % overlap, and 8192 bins) for f1 to f6 with load 30 V and speed 8-15 Hz.



(a) f7 (b) f8

(c) f9 (d) f10

Figure 2-3: Time-frequency representation (spectrogram with hamming window, 50 % overlap, and 8192 bins) for f7 to f10 with load 30 V and speed 8-15 Hz.

the standard deviation in #1 and #2. However, note that for #3 the knn and PCA-knn achieve comparable accuracy to CKA-knn as seen in Table **2-5**. We obtained the best performance for the speeds constant, assuming that the vibration signal is stationary; the time and frequency domains provide excellent interpretation capability. In contrast, the worst performance is presented for the experiment #2, due to the vibration signal is no-stationary.

Table 2-5: Best accuracy of gear fault classification by averaging of 10-folds for each methods.

Experiment (%)	knn	PCA-knn	CKA-knn
#1	91.3 ±1.9	88.4 ±2.3	**94.4 ±1.7**
#2	84.9 ±4.0	83.0 ±3.2	**86.7 ±2.9**
#3	88.7 ±1.5	86.2 ±1.3	**89.9 ±2.1**

In Fig. 2-4 is observing the median confusion matrices calculated on the accuracy of 10-folds. As shown in Table 2-5, we can appreciate that CKA-knn achieves the best performance. We note that for the experiment #1, the CKA-knn method shows high percentages of hits for the fault f7 (100 %), in contrast to knn and PCA-knn that are prone to fail (72.2 %). On the other hand, the experiment #2 with CKA-knn achieves acceptable sorting f7 (83.3 %), but its performance is low for f8 (66.7 %). Furthermore, both PCA-knn (88.9 %) as CKA-knn (94.4 %) improve performance for f10 over baseline (61.1 %) but have low accuracy in f8. In the experiment #3, CKA-knn obtained a high performance for faults f2 (100 %) and f3 (94.4 %), such faults are important since they have an early condition of fault. Additionally, it manages to improve the classification for f9, reaching 88.9 % of accuracy in comparison to knn 72.2 % and PCA-knn con 69.4 %.
We calculated the relevance vector for PCA and CKA methods, as evidenced by in Fig. 2-5. For visualization purposes, we perform normalization on the maximum relevance value for each method. In general, we can appreciate that PCA highlights many features; however, the performance is the lowest. Also, we note that the most relevant feature for all experiments performed is $\tilde{\varphi}_3$ calculated on the frequency-marginal. For its part, the CKA method highlights fewer features but the performance is the best in the 3 experiments (seen Table 2-5), where the highest values correspond to φ_{10} para #2 y #3, $\hat{\varphi}_2$ en #1 y #2, y $\hat{\varphi}_{10}$ for all the experiments.

2.4. Discussion

The methodology proposed in this chapter was applied to vibration signals with different types of faults, covering the typical cases of gear faults. Besides, they allow us to verify the robustness of the methodology because the database has variable operating conditions, such as different loads and speeds with which the data acquisition is configured.
For each experiment carried out in the development of this work, we tested with unsupervised learning method such as PCA and also with a supervised learning method such as CKA, as expected the best accuracy was obtained with CKA because the targets play an essential

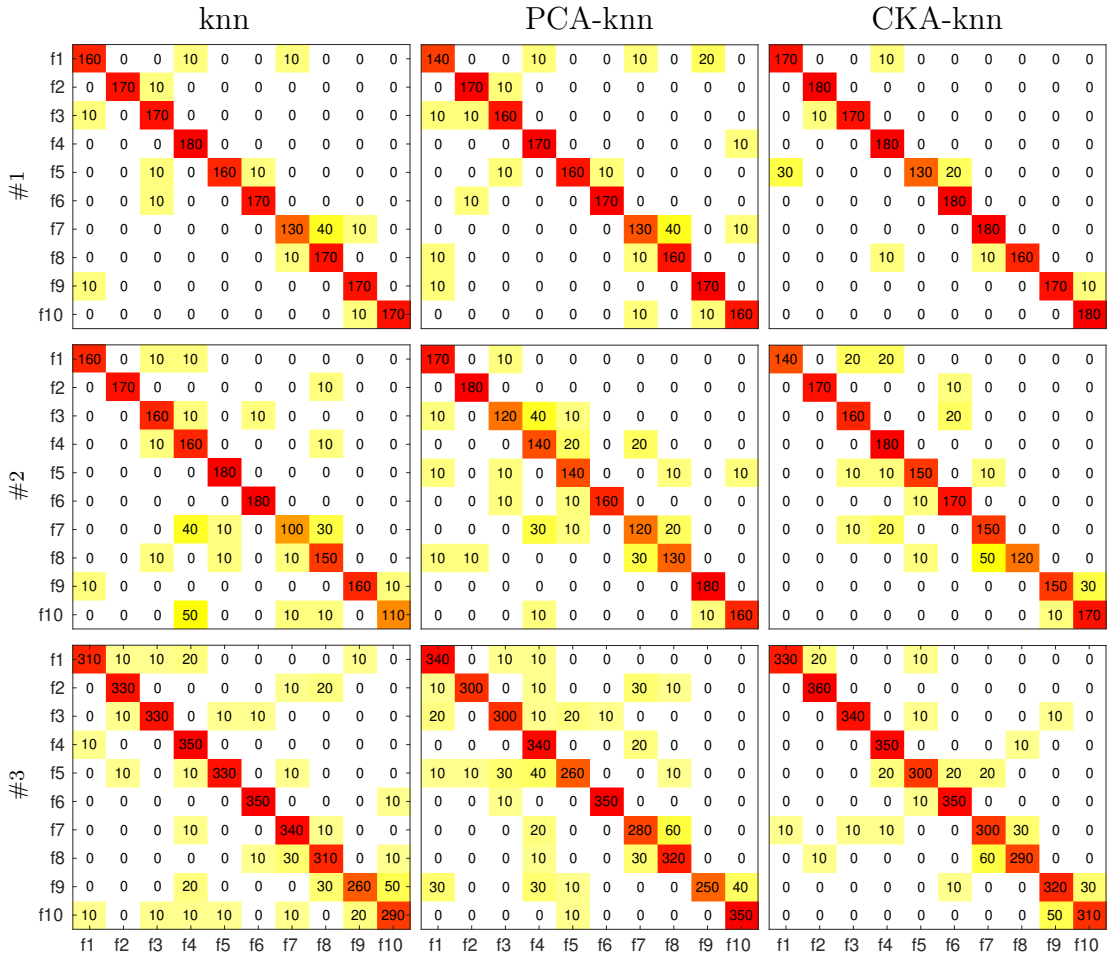


Figure 2-4: Median confusion matrices of each methods.

role in the convergence of the algorithm. However, in real-world applications, labels are not always available, and PCA showed that it could provide excellent results by obtaining an accuracy above 80 %, taking into account that it is a multi-class problem. Regarding features used, our proposal allows determining the distribution of relevance for each of them, this being an important point to determine the influence of the features in time-domain, in frequency-domain, and in the time-frequency domain. In general, for PCA, both features the time-domain and the frequency-domain are highlighted, but the most relevant feature for all experiments corresponds to the kurtosis calculated on the frequency-marginal, this may be due to the frequency marginals they are strong concentrations of energy (Figs. **2-2** and **2-3**). On the other hand, CKA showed that the relevant features are in time-domain and in frequency-domain, which may be a consequence of the algorithm having the reference of the labels; also, the difference between the characteristics is marked.

For the classification phase, we used a simple classifier as is k nearest neighbor, because we wanted to concentrate the effort on the methods of extracting features. Additionally, we

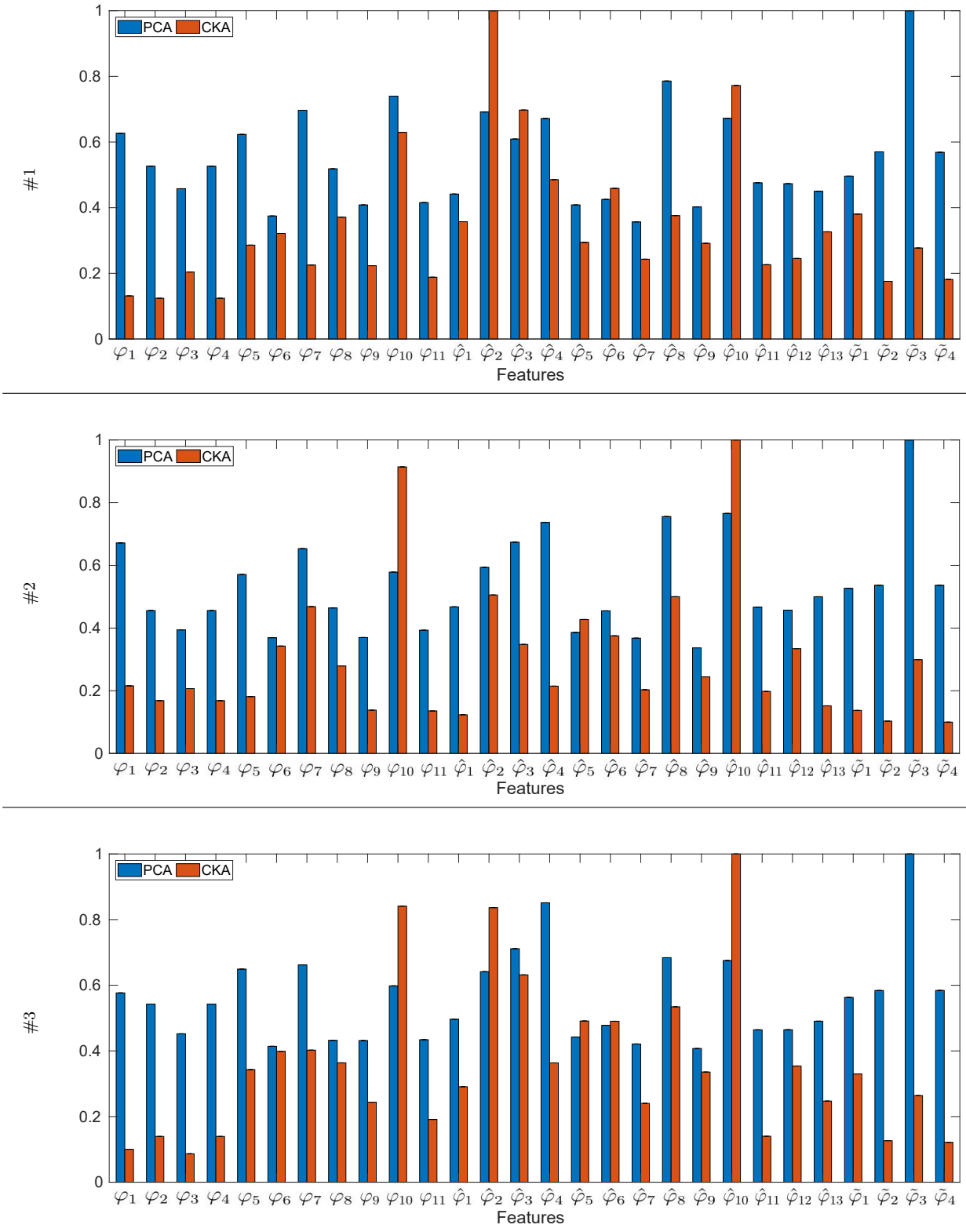


Figure 2-5: Features extraction for each of experiments.

took into account that the training and evaluation data were different for each fold and thus avoid any bias in the classifier. Finally, we found that for some types of faults, the accuracy was low (Fig. **2-4**). In general terms, the performance achieved in this work is remarkable. In this chapter, we test statistical features on different domains. Implementing feature extraction methods is important. This methods retain important features for fault detection. In addition, we show that the characteristics relevant to classification are those calculated in the domain of time and frequency. However, we note that some features are vulnerable to noise making the performance of the classification low. Finally, the classification results are acceptable because they allow to differentiate the incipient failures with great precision.

3 Fault identification by the comparison of blind source separation methods

3.1. Introduction

Data-driven fault identification methods have become important in recent years. The growth of complex industrial systems and the development of sensing technologies allow taking advantage of the captured signals. The main idea of data-driven methods is to use data processing algorithms to extract useful information and latent knowledge. In recent decades, there are methodologies for extracting and structuring information from high-dimensional, such as data-driven multivariate statistical process monitoring methods, which have generated greater interest in industrial research and attention (Žvokelj et al., 2016). The vibration signals collect different physical parameters of the mechanical systems. Hence, Blind source separation (BSS) allows separating fault patterns of components polluted by environment noise or other mechanical systems.

As is explain in Hyvärinen et al. (2004), several methods are very closely related to BSS. We used Independent Component Analysis (ICA) given that does not need the prior knowledge of the problem, it is enough to assume statistical independence between the sources. Thus, we can identify fault characteristics by source separation, where sources represent a force that influences a machine component known as a fault in the mechanical systems (Yang and Nagarajaiah, 2014). Various ICA-based methods have been developed. Hyvärinen and Oja (2000) proposed the FastICA algorithm, it is based on a fixed-point iteration scheme for finding a maximum of the non-Gaussianity. In Žvokelj et al. (2016) has utilized Ensemble Empirical Mode Decomposition and ICA for bearing fault detection. Besides, FastICA had been used as fault feature extraction (Wu and Xiong, 2019). These methods, require a stage of characterization increasing computational cost. Other authors proposed the SOBI algorithm based on the joint approximate diagonalization of multiple time-delayed correlation matrices (Belouchrani et al., 1997; Choi et al., 2002), and Cardoso (1999) developed the JADE algorithm, which is specifically a statistic based technique. JADE uses the joint diagonalization of a maximal set guarantees source identifiability. In Miao et al. (2020), proposed a fault separation method which it is based on the median filter and JADE but the filter can remove components from the fault and generate an error in reconstructs the characteristic signal. In this context, the chapter presents a formal comparison of three methods of the Independent Component Analysis: FastICA, JADE, SOBI, to determinate faults in rotative

machines and hence push for further development in this field.
The rest of the chapter is organized as follows: In Section 3.2 introduces the mathematical background of the Independent Component and spatial distribution information of the source. Section 3.3 demonstrates the present technique with experiments using synthetic signals and real data. The results of the experiments are discussed in Section 3.4.

3.2. Methods of independent component analysis (ICA)

To address the work done from a theoretical perspective, we must start from the fact that the final objective is to obtain a statistically independent representation of the vector of measures $\boldsymbol{x}(t) = [x_1(t), x_2(t), ..., x_M(t)]^\top$, which is defined $\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t)$, where $\boldsymbol{s}(t) = [s_1(t), s_2(t), ..., s_N(t)]^\top$ is the vector of unknown sources, which are considered statistically independent and $\boldsymbol{A} \in \mathbb{R}^{M\times N}$ is the mixing matrix. As a restriction we have that of number of accelerometers $m{\in}M$, is greater than or equal to the number of sources $n{\in}N$. However, the task of sources separation consist in recovery the vector $\boldsymbol{s}(t)$ of the observation $\boldsymbol{x}(t)$ without known $\boldsymbol{A}$. Essentially, the task consists in finding a linear projection that will transform the observations $\boldsymbol{x}(t)$ into statistically independent signals $\boldsymbol{y}(t)$, that according it must be a proportional version of the sources $\boldsymbol{s}(t)$. Therefore, the linear projections mentioned before it is can to represent basically like a matrix $\boldsymbol{W} \in \mathbb{R}^{N\times M}$ (seen Eq. (3-1)) that it is denominated as *demixing matrix*, such that, $\boldsymbol{W} = \boldsymbol{A}^{-1}$ and noting that for an orthogonal matrix $\boldsymbol{W}^{-1} = \boldsymbol{W}^\top$.

$$\boldsymbol{y}(t) = \boldsymbol{W}\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{A}\boldsymbol{s}(t) \tag{3-1}$$

We can not determine the order of the independent components, for this reason, we can freely change the order of the terms (Hyvärinen and Oja, 2000). On account of this, a permutation matrix $\boldsymbol{P}$ and its inverse are introduced as $\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{P}\boldsymbol{s}(t)$. So, we have the original independent variables $s_n(t)$ on the elements of $\boldsymbol{P}\boldsymbol{s}(t)$, but in other order. Additionally, the new unknown mixing matrix $\boldsymbol{A}\boldsymbol{P}^{-1}$ it is solved by the ICA algorithms.

3.2.1. FastICA

In FastICA, the vector $\boldsymbol{x}(t)$ is first whitened for obtaining the vector $\boldsymbol{z}(t) = \boldsymbol{V}\boldsymbol{x}(t)$ for which $\mathbb{E}\left\{\boldsymbol{z}\boldsymbol{z}^\top\right\} = \boldsymbol{I}$. Thus, in the model $\boldsymbol{z}(t) = \boldsymbol{V}\boldsymbol{x}(t) = \boldsymbol{V}\boldsymbol{A}\boldsymbol{s}(t)$ both $\boldsymbol{z}(t)$ and $\boldsymbol{s}(t)$ are white; $\boldsymbol{z}(t)$ because of explicit white, and $\boldsymbol{s}(t)$ by the assumptions of zero mean, unit variance, and independence of the sources. If $\boldsymbol{A}$ were known, the sources could be directly solved from $\boldsymbol{s}(t) = \boldsymbol{V}\boldsymbol{A}^\top\boldsymbol{z}(t)$. In the FastICA algorithm, the solution is sought as:

$$\boldsymbol{y}(t) = \boldsymbol{W}\boldsymbol{z}(t) \tag{3-2}$$

The separation matrix $\boldsymbol{W}$ is found by numerical algorithms. By writing the rows of $\boldsymbol{W} = \boldsymbol{w}_n^\top$. The FastICA algorithm (Hyvarinen, 1999) is an iterative method to find the local maxim of a cost function:

$$\mathcal{J}_G = \sum_{n=1}^{N} \mathbb{E}\left\{G(\boldsymbol{w}_n^\top \boldsymbol{z}(t))\right\} \tag{3-3}$$

with G a nonlinear function which is usually assumed even and symmetrical and $\mathbb{E}\{\cdot\}$ stands for the expectation operator. The cost function to be maximized can be negative mutual information, likelihood, some approximation of non-Gaussianity such as higher order cumulants, or some extension of these. A widely used cost function is the fourth-order cumulant or kurtosis, defined for any random variable $\boldsymbol{y}$ as:

$$kurt(\boldsymbol{y}) = \mathbb{E}\left\{\boldsymbol{y}^4\right\} - 3\left(\mathbb{E}\left\{\boldsymbol{y}^2\right\}\right)^2 \tag{3-4}$$

The argument $\boldsymbol{y}_n = \boldsymbol{w}_n^\top \boldsymbol{z}$ is restricted to have unit variance, and thus its kurtosis is $\mathbb{E}\{\boldsymbol{y}^4\} - 3$. In maximization, the second term can be dropped, and the criterion becomes :

$$\mathcal{J}_G^{kurt} = \sum_{n=1}^{N} \mathbb{E}\left\{(\boldsymbol{w}_n^\top \boldsymbol{z})^4\right\} \tag{3-5}$$

For one of the rows $\boldsymbol{w}_n^\top$, the FastICA algorithm for kurtosis maximization makes the basic updating step:

$$\boldsymbol{w}_n^+ = \mathbb{E}\left\{\boldsymbol{z}(\boldsymbol{w}_n^\top \boldsymbol{z})\right\} - 3\boldsymbol{w}_i \tag{3-6}$$

followed by normalization of vector $\boldsymbol{w}_n^+$ to unit norm.

3.2.2. Joint approximate diagonalization of Eigenmatrices (JADE)

JADE is an algorithm based on the joint diagonalization of cumulative matrices, under the assumption that sources have non-Gaussian distributions, using second-order statistics to decorrelate data and fourth-order cumulant for the joint diagonalization the matrix of mixes (Cardoso and Souloumiac, 1993). The objective is to maximize the non-Gaussianity of the complete set of observed signals $\boldsymbol{X}$, for obtain the mixing matrix $\boldsymbol{A}$ and the estimation of source signals $\boldsymbol{Y}$, as seen in Eq. (3-1). Then, we whiten the signals for obtaining the vector $\boldsymbol{z}(t)$. Now, we calculate the fourth-order cumulants as follows:

$$\mathcal{C}_z = cum(\boldsymbol{z}_i, \boldsymbol{z}_j, \boldsymbol{z}_k, \boldsymbol{z}_l) | 1 \leq i, j, k, l \leq d \tag{3-7}$$

where d is the dimension of vector $\boldsymbol{z}$ and cum is the computation of the cummulants, which can be defined as:

$$\begin{aligned} cum(\boldsymbol{z}_i, \boldsymbol{z}_j, \boldsymbol{z}_k, \boldsymbol{z}_l) = {} & \mathbb{E}\{\boldsymbol{z}_i, \boldsymbol{z}_j, \boldsymbol{z}_k, \boldsymbol{z}_l\} - \mathbb{E}\{\boldsymbol{z}_i, \boldsymbol{z}_j\}\mathbb{E}\{\boldsymbol{z}_k, \boldsymbol{z}_l\} \\ & - \mathbb{E}\{\boldsymbol{z}_i, \boldsymbol{z}_k\}\mathbb{E}\{\boldsymbol{z}_j, \boldsymbol{z}_l\} - \mathbb{E}\{\boldsymbol{z}_i, \boldsymbol{z}_l\}\mathbb{E}\{\boldsymbol{z}_j, \boldsymbol{z}_k\} \end{aligned} \tag{3-8a}$$

Then, we estimate a maximal set of cumulant matrices as:

$$\mathcal{C}_z(\boldsymbol{E}_i) = \lambda_i \boldsymbol{E}_i \tag{3-9}$$

where λ_i is the eigenvalue and $\boldsymbol{E}_i$ is the eigen-matrix. We need that the cumulant matrices be as diagonal as possible, so, we introduce the unitary matrix $\hat{\boldsymbol{U}}$:

$$\hat{\boldsymbol{U}} = \arg\min \sum_i off(\boldsymbol{U}^{\#} \mathcal{C}_z(\boldsymbol{E}_i)\boldsymbol{U}) \tag{3-10}$$

where # denotes the pseudo-inverse, off is the square of the non-diagonal elements and $\boldsymbol{U}$ is the rotation matrix. Finally, matrix $\boldsymbol{A}$ will be computed as $\hat{\boldsymbol{U}}$

$$\boldsymbol{A} = \hat{\boldsymbol{U}}\boldsymbol{W}^{\#} \tag{3-11}$$

3.2.3. Second-order blind identification (SOBI)

Choi et al. (2002) proposed the SOBI algorithm based on the joint approximation diagonalization of multiple time-delayed correlation matrix. SOBI takes advantage of the temporal structure in the observed data, assuming that are stationary sources of second-order and not correlated mutually. Thus, we compute a set of time-delayed correlation matrices beginning with $\boldsymbol{z}(t)$, so, $\boldsymbol{R}_z(\tau_i)$, $\forall i = 1, \ldots, K$. The SOBI finds an unitary transformation $\boldsymbol{D}$ such that

$$\boldsymbol{D}^{\top}\boldsymbol{R}_z(\tau_i)\boldsymbol{D} = \boldsymbol{\Lambda}_i, \forall i = 1, \ldots, K \tag{3-12}$$

where $\boldsymbol{\Lambda}_i$ is a set of diagonal matrices. Given that, SOBI employ several correlation matrices, it reduce the probability that $\boldsymbol{R}_z(\tau_i)$ has the same diagonal elements by some bad choice of time-lag τ_i. Furthermore, the exploitation of several correlation matrices increases the statistical efficiency. As a result, matrix $\boldsymbol{A}$ can be factored as

$$\boldsymbol{A} = \boldsymbol{W}^{\#}\boldsymbol{D} \tag{3-13}$$

where # denotes the Moore-Penrose pseudoinverse. For detailed description, see Belouchrani et al. (1997).

3.3. Experimental Set-Up

The methodological development of this approach comprises the following stage: *i)* Estimation of ICs on data using FastICA, JADE and SOBI, *ii)* Analysis of envelope of the signals of vibrations captured by the accelerometer followed by the calculation of faults frequency. The proposed approach (seen Fig. **3-1**) is evaluated on synthetics data and on test rig over a fixed machine.

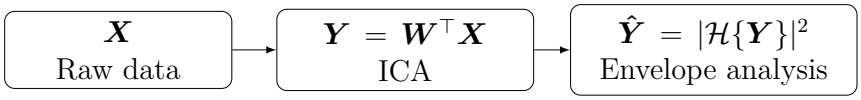


Figure 3-1: Proposed approach for differents methods ICA.

3.3.1. Synthetic data

We consider a normal pair of gears, such that all teeth of the gear are identical and are equally spaced, under a constant load and speed (Fan and Zuo, 2006). The behavior of the vibration signal is presented as follows:

$$\psi(t) = \sum_{l=0}^{L} \Psi_l cos(2\pi l f_c t + \phi_l) \tag{3-14}$$

where $f_c = Z f_r$ given that, Z is the number of teeth and f_r mean the frequency of the rotating shaft. L is the number of harmonics, Ψ_l and ϕ_l are, the amplitude and the phase angle of the l-th harmonic of mesh frequency, respectively. A faulty gear tooth generates a change in vibration signature and it can yield amplitude and phase modulation. Thus, the gear fault signal can be represented by Eq. (3-15):

$$g(t) = \sum_{l=0}^{L} \Psi_l (1 + b_l(t)) cos(2\pi l f_c t + \phi_l + c_l(t)) \tag{3-15}$$

where $b_l(t)$ and $c_l(t)$, respectively, are the amplitude and phase modulation functions. In general, the types of faults on gear tooth are not identical and the faults yield a periodic excitation (Kim and Lee, 2018). Hence, we define the modulation function as follows:

$$b_l(t) = \sum_{l'=1}^{L'} B_{ll'} cos(2\pi l' f_c t + \beta_{ll'}) \tag{3-16}$$

$$c_l(t) = \sum_{l'=1}^{L'} C_{ll'} cos(2\pi l' f_c t + \gamma_{ll'}) \tag{3-17}$$

where L' is the number of sidebands around tooth-meshing harmonics, $B_{ll'}$, $C_{ll'}$ are amplitudes and $\beta_{ll'}$, $\gamma_{ll'}$ are the phase at the l'-th sidebands of amplitude and phase-modulating signals, respectively, around the l-th meshing harmonic. By introducing Eqs. (3-16) and (3-17) into Eq. (3-15), appear additional frequency components that vary with time. Finally, the model for a gearbox with a faulty gear tooth is given in equation Eq. (3-18)

$$\begin{aligned} g(t) = & \sum_{l=0}^{L} \Psi_l \left(1 + \sum_{l'=1}^{L'} B_{ll'} cos(2\pi l' f_c(t) t + \beta_{ll'})\right) \\ & \times cos\left(2\pi l f_c(t) t + \phi_l + \sum_{l'=1}^{L'} C_{ll'} cos(2\pi l' f_c(t) t + \gamma_{ll'})\right) \end{aligned} \tag{3-18}$$

In order to simulate a frequency variation phenomenon respect to time, the signal $cos(2\pi\times1560t)+cos(2\pi\times1600t)$ is modulated by frequencies of 40 and 80 Hz, where the second frequency is established for short intervals of time along the signal. Finally, a random signal with a normal distribution is added. This simulated signal is denoted as $S1$ and it is showed in Fig. **3-2**.

$$g(t) = (1+sin(2\pi 40t))(cos(2\pi 1560t)+sin(2\pi 40t)) + \begin{cases} (1+sin(2\pi 80t))(cos(2\pi 1600t)+sin(2\pi 80t)), & t = [0, 0.099], \\ cos(2\pi 1600t), & t = [0.1, 0.3], \\ (1+sin(2\pi 80t))(cos(2\pi 1600t)+sin(2\pi 80t)), & t = [0.301, 0.5], \end{cases} \tag{3-19}$$

Besides, a synthetic signal is generated with periodic components of low frequency and transitory components of high frequency for simulating multiple fail signals (Wang and Tse, 2012), this behavior is described by the following formula:

$$q_1(t) = \exp(-\lambda(tF_s/f_b)/F_s)\times sin(2\pi f_0 t/F_s) \tag{3-20}$$

$$q_2(t) = sin(2\pi f_1 t/F_s) + 0.8sin(2\pi f_2 t/F_s) \tag{3-21}$$

The fault signal is present in Eq. (3-20) where $\lambda = 900$; $f_b = 100\,Hz$ is the fault characteristic frequency and the resonant frequency is define by $f_0 = 3700\,Hz$ labeled as $S2$. The low frequencies are shown case-by-case due to that signal is similar to the carrier wave that is describe in Eq. (3-21), which is denoted as $S3$ where $f_1 = 60\,Hz$ and $f_2 = 90\,Hz$ are the components of low frequency, and the sample frequency $F_s = 8192\,Hz$. At last, we use 4090 samples to simulate signal and we add a random signal with a normal distribution. From synthetic signals, we generate the matrix sources such that $\boldsymbol{S} = [g(t); q_1(t); q_2(t)]$ we use the notation of Matlab for concatenate. We produce a random matrix mixing $\boldsymbol{A}$ to linearly combine the source and to simulate matrix measures $\boldsymbol{X} = \boldsymbol{AS}$ as shown in Fig. **3-2** as $Mixture$.

Comparing the original spectrums showed in Fig. **3-2**, against the spectrums of the reconstructed signals shown in Fig. **3-3**, we can see that the overall performance of the methods is significant, given that the three methods separate mixed signals without mistake. Moreover, the spatial distribution information of sources was not applied to this experiment because the data have no spatial information.

3.3.2. Experiment in bearing fault detection

As shown in Fig. **3-4**, experimental test rig includes a $2\,HP$ electromotor `Siemens` with $1800\,rpm$ maximum speed. The motor is connected to shaft by a rigid coupling and has two supports, each one holding a ball bearing `SKF-6005NR` and two wheels. Drilling wells are designed to create either static or dynamic unbalance problems. To measure machine mechanical vibration, accelerometers are also included, which are located perpendicularly to

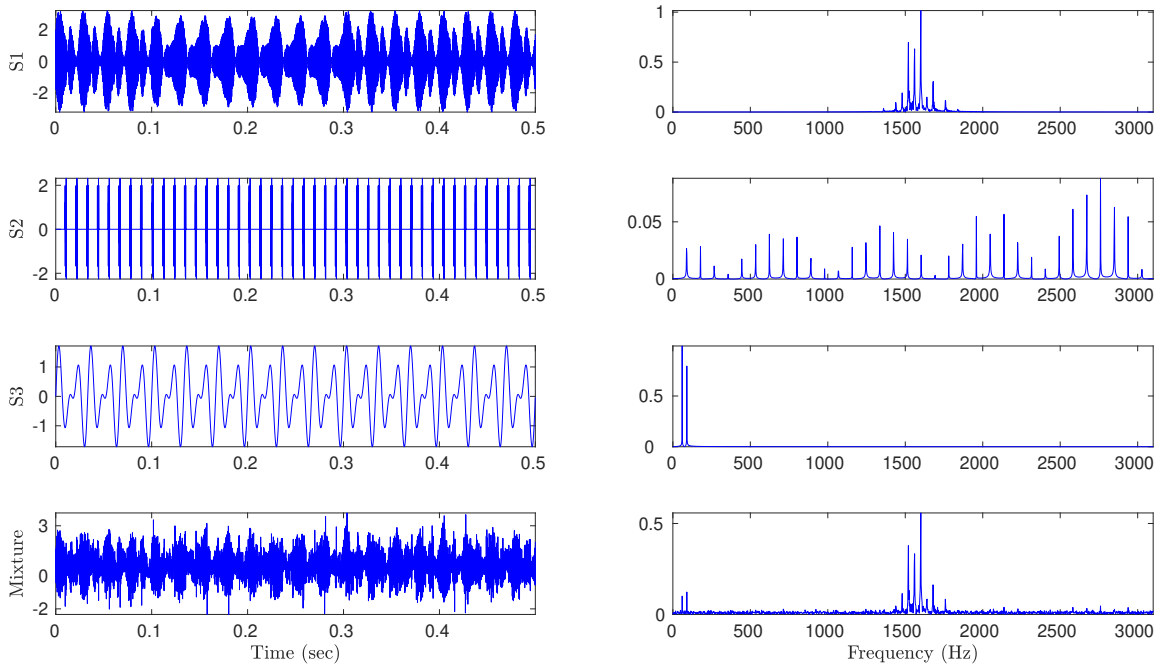


Figure 3-2: Simulate data synthetics

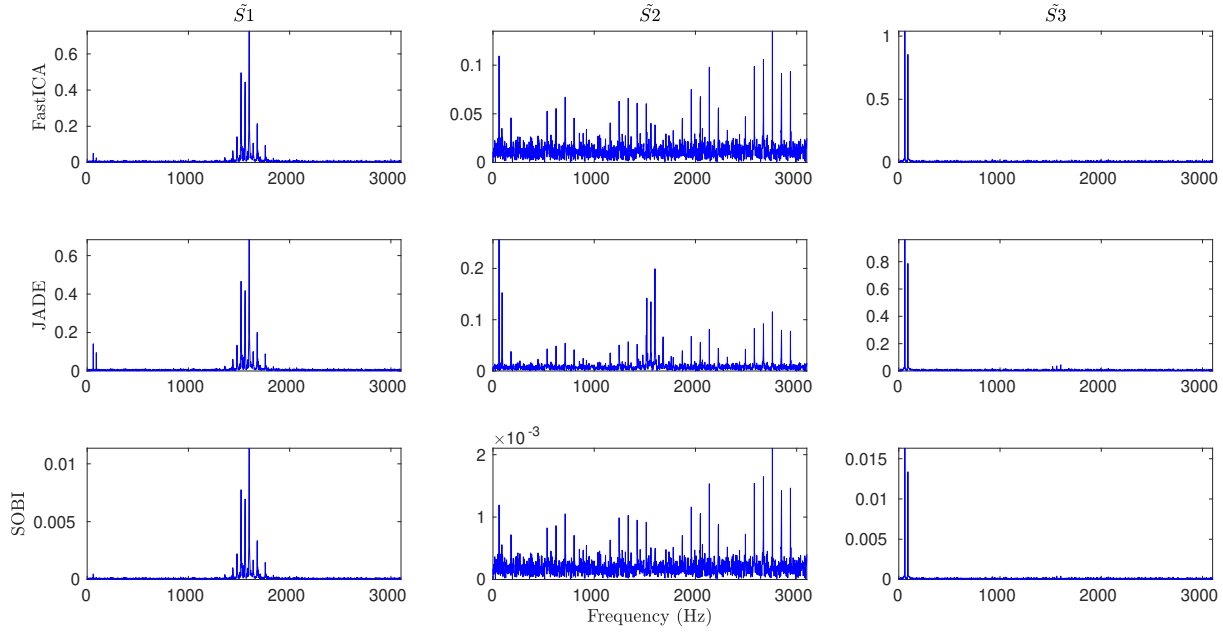


Figure 3-3: Spectrum of ICs data synthetic

the shaft horizontal plane (labeled as *accelerometer location*). In this experiment, just the `ACC102` accelerometer placed near the machine is employed, which has a measurement range of $0-10\,kHz$ and $100\,mV/g$ of sensibility. The `National Instruments USB-6009` data acquisition card acquires vibration recordings at $20\,kHz$ sampling frequency.

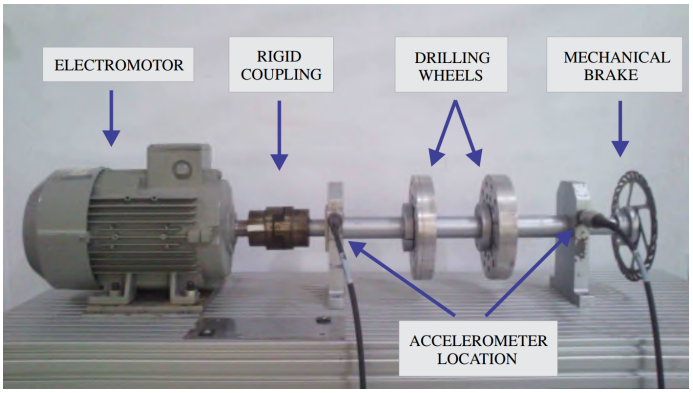


Figure 3-4: Experiment set-up for test rig.

It is selected several registers to apply the methodology described in Section 3.3. The feature principal that it is sought in the registers was select those in which the fault inner race and fault balls were found present in the axes horizontal and vertical of both accelerometers, we seek to verify the performance of ICA for BSS both in fault inner race as in fault ball. In particular, the bearing fault frequency may be approximated by the geometrical properties of the rolling element bearings and the shaft rotational speed f_r of the system (Smith and Randall, 2015b). Thus, the bearing fault frequencies are as follows:

Table 3-1: Theoretical bearing fault frequencies, where z is the number of rolling elements, θ is the contact angle of the load from radial plane. and $d^{'} = d/D$, being D and d the bearing pitch and ball diameters, respectively.

Ball pass frequency, outer race:	$BPFO = zf_r(1 - d^{'}\cos\theta)/2$
Ball pass frequency, inner race:	$BPFI = zf_r(1 + d^{'}\cos\theta)/2$
Ball spin frequency:	$BSF = Df_r(1 - (d^{'}\cos\theta)^2)/2d$
Fundamental train frequency:	$FTF = f_r(1 - d^{'}\cos\theta)/2$

where z is the number of rolling elements, D and d are respectively the bearing pitch and ball diameters, θ is the contact angle of the load from the radial plane. Fault frequencies above are based on kinematic relationships assuming that there is not frequency slip, yet it is always present having a variation from the calculated frequency of up to $1-2\,\%$ (Smith and Randall, 2015b).

The main spectral characteristics from outer race, inner race, and rolling elements, which may be identified in the envelope spectrum according to Taylor (1994) are: *i)* $BPFI$ and harmonics, sidebands spaced at f_r; And *ii)* BSF and harmonics (even harmonics often dominant), sidebands spaced at FTF, harmonics of FTF.

Experiment in ball pass frequency, inner race

In Fig. **3-5** is shown register 19, we can appreciate the signals in the time domain for each of mixture signals collected for accelerometers denote with $S1_h$ and $S1_v$ for first support, in the same the way $S2_h$ and $S2_v$ for the second support, where it is not visible a patron of fault for the signals $S1$ and $S2$ in both axes.

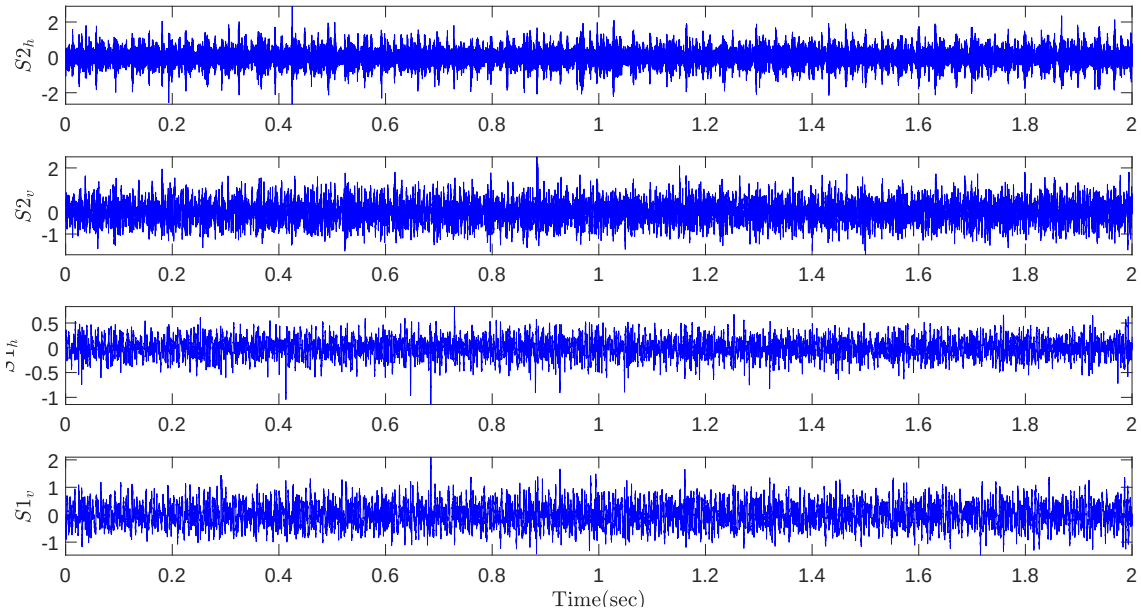


Figure 3-5: Signals collected of accelerometers fault inner race.

The step following is to carry out the envelope analysis and Fourier transform, additionally, we establish the fault frequency inner race, in order to easily appreciate the fault collected by all the accelerometers as illustrated in Fig. **3-6**. Thus, it is very difficult determined in which of the supports is found the fault given that we have a sign of fault but we can not ensure the support it belongs to fault with absolute precision.

We evaluated three algorithms of BSS: FastICA, JADE, SOBI the results are shown in Fig. **3-7** respectively. The signals of Fig. **3-5** make up the input mixture matrix $\boldsymbol{X}$ for ICA. It should be borne in mind that the first dominant peak for $S2_h$, $S2_v$, $S1_v$ corresponds to the rotational frequency of the shaft and its second harmonic corresponds for the first dominant peak in $S1_h$.

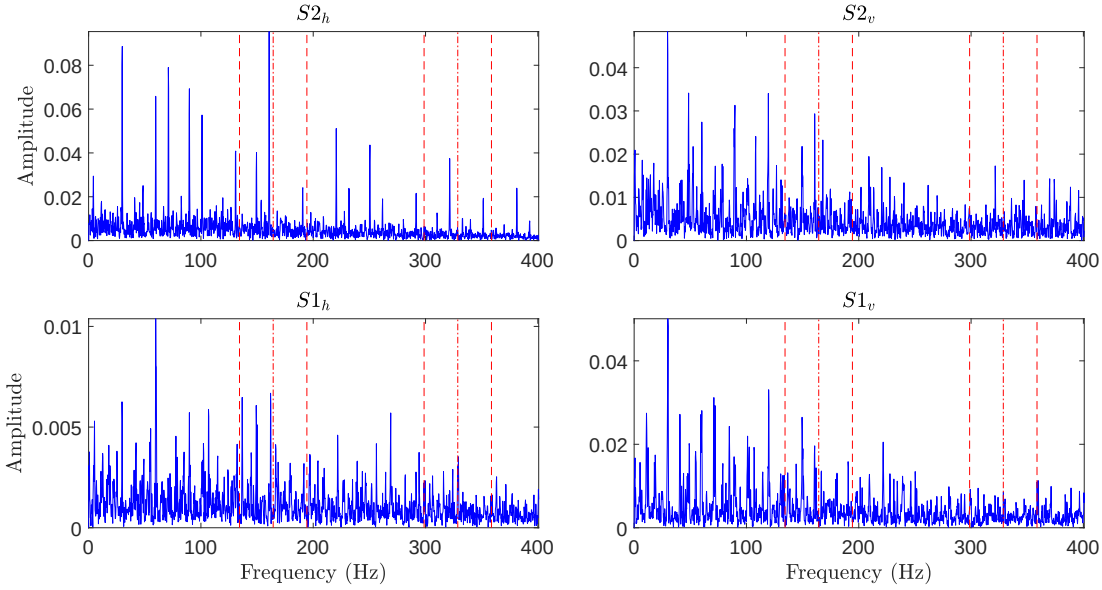


Figure 3-6: Spectrum of envelope of accelerometers fault inner race. Harmonic cursors multiples of $BPFI$ ($164.3\,Hz$) and sidebands cursors multiples of $BPFI \pm fr$ ($164.3 \pm 30\,Hz$) are displayed in red colour.

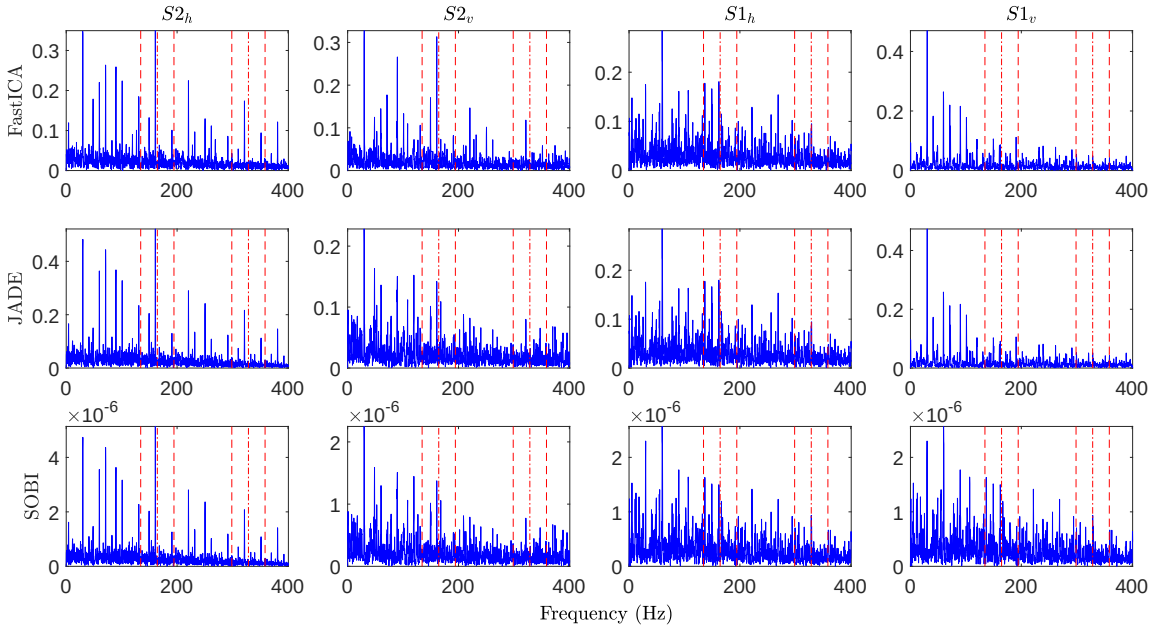


Figure 3-7: Spectrum of ICs fault inner race for FastICA, JADE, SOBI. Harmonic cursors multiples of $BPFI$ ($164.3\,Hz$) and sidebands cursors multiples of $BPFI \pm fr$ ($164.3 \pm 30\,Hz$) are displayed in red colour.

As evidenced in Fig. **3-7**, after applying the proposed methodology is achieved to establish the origin of the fault in $S2$ in both axes, however, of the three algorithms raised FastICA and JADE are those that present a solution relevant to BSS because they are able to turn off the component of the failure for the axes $S1_v$, in contrast to what happens in SOBI that keeps the components of the fault turned on for $S1_v$. Over and above, FastICA achieves the best performance by virtue of that highlights the fault in the accelerometer $S2$, both $S2_h$ and $S2_v$ axes, contrary to the methods of JADE and SOBI, which are able to separate the component of fault in $S2_h$ but fail to highlight the relevant information of the fault for the axis $S2_v$. The spatial distribution information works correctly for the association of the ICs with the location of the sources because the measurements of the data were carried out with a prior knowledge of the bearing that originates the fault, which in this case is recorded by accelerometer $S2$ as presented in Fig. **3-7**.

Experiment in ball spin frequency

Following the methodology presented in Section 3.3.2, the experiment was performed for the ball fault, selecting the registry 59 as evidenced in Fig. **3-8** in the time domain and Fig. **3-9** in the frequency domain, where it is not visible a patron of fault for the signals $S1$ and $S2$ in both axes. Besides, it should be taken into account that for this experiment the fault is in $S1_h$ and $S1_v$.

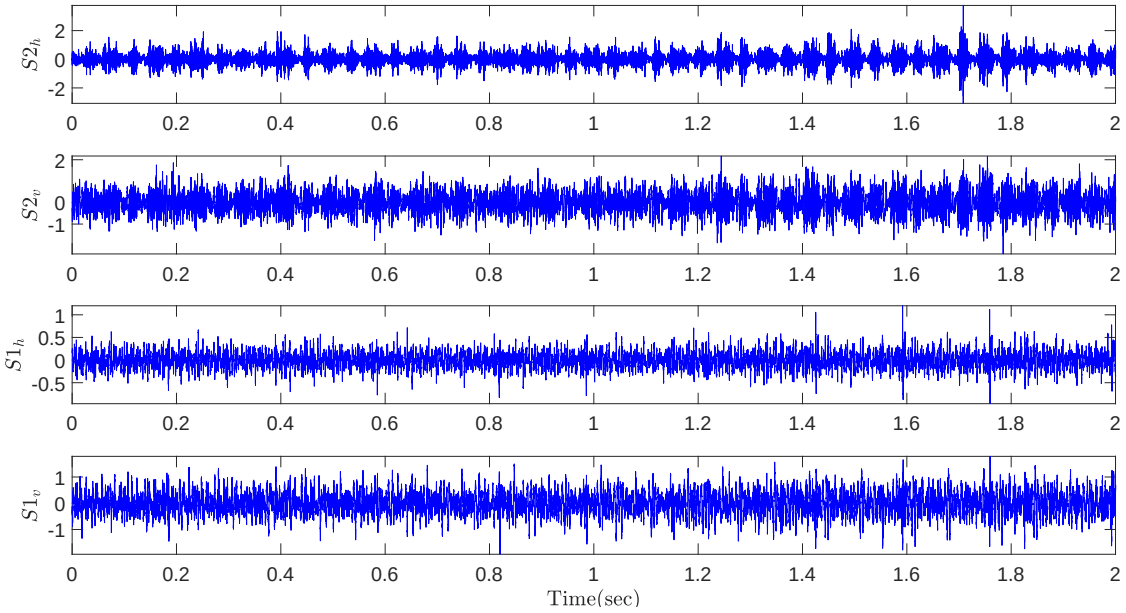


Figure 3-8: Signals collected of accelerometers fault ball.

The signals of Fig. **3-8** make up the input mixture matrix $\boldsymbol{X}$ for the methods of ICA, we establish the fault frequency ball and shaft spin frequency given that in this type of fault is

common for peaks to be presented in harmonic cursors multiples as illustrated in Fig. **3-9**. Thus, it is very difficult determined in which of the supports is the fault.

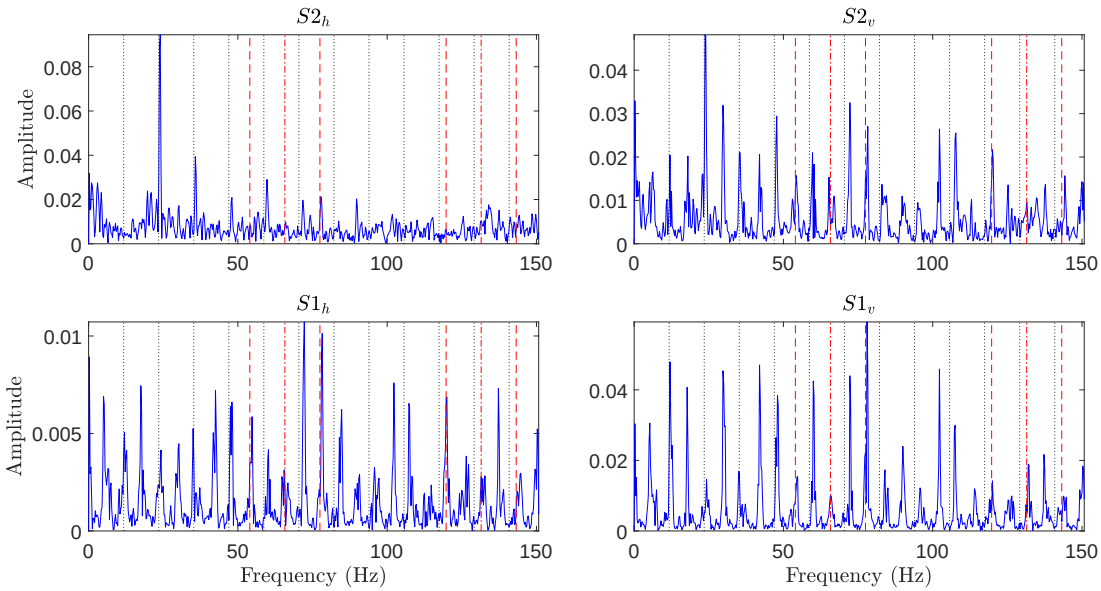


Figure 3-9: Spectrum of envelope of accelerometers fault ball. Harmonic cursors multiples of BSF (65.7 Hz) and sidebands cursors multiples of $BSF \pm FTF$ (65.7$\pm$11.7 Hz) are displayed in red colour. Harmonic cursors multiples of FTF (11.7 Hz) are displayed in black colour.

We evaluated the algorithms of BSS: FastICA, JADE, SOBI and the results are shown in Fig. **3-10** respectively. We find the same behavior of the Section 3.3.2, where the ICs calculated by JADE and SOBI highlight information important of the fault for $S1$ in both axes, but only attenuate the peaks of the multiple harmonics of $BSF \pm FTF$ and FTF for $S2_h$ and not for the axis $S2_v$, complicating the association of fault. Contrarily, FastICA presents the best performance because it highlights the behavior of fault in $S1$ separating both the first and second harmonic of fault (lines displayed in red colour) and shows multiple harmonics of FTF that indicate the presence of fault for that accelerometer, additionally, in $S2_h$ is shown only harmonics of FTF and there is no evidence of the fault, meanwhile, $S2_v$ extinguishes all relevant peaks which leads us to conclude in a correct way the origin of fault. In regard to the spatial distribution information allows us to correctly associate the calculated ICs to the accelerometer $S1$ which is expected for this experiment. It should be noted that for the support $S2$ in the three algorithms of ICA, presents a peak in the frequency of $2FTF$ which corresponds to the frequency of rotation of the shaft and is not representative for the identification of fault.

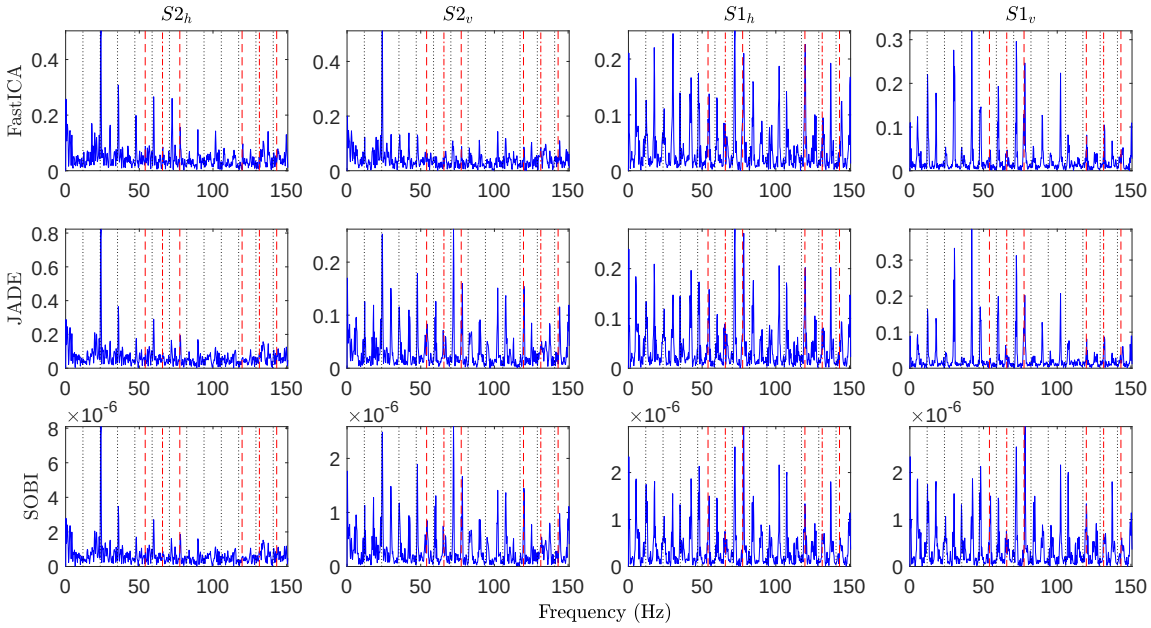


Figure 3-10: Spectrum of ICs fautls ball for FastICA, JADE, SOBI. Harmonic cursors multiples of BSF $(65.7\,Hz)$ and sidebands cursors multiples of $BSF \pm FTF$ $(65.7 \pm 11.7\,Hz)$ are displayed in red colour. Harmonic cursors multiples of FTF $(11.7\,Hz)$ are displayed in black colour.

3.4. Discussion

It was observed that for synthetic data the methods of BSS: FastICA, JADE, and SOBI tested in this chapter, achieve the estimation of ICs for the proposed dynamics without any error taking into account the information of the linear mixtures of the sources without any other a priori information, in this way we validate the correct operation of the BSS methods mentioned above. On the other hand, for our database, the best performance achieved is FastICA followed by JADE, where the method of SOBI does not achieve a good performance because the method takes advantage of the temporal structure of the observed data, i. e. is based on the estimation of the components from the correlation, and as seen in Figs. **3-5** and **3-8**, the data captured by the accelerometers do not offer any discriminating information about the faults, therefore, second-order statistics are not enough to properly estimate ICs. For its part, JADE uses fourth-order statistics and minimizes the crossing of cumulants to achieve independence among the estimated components (Choi et al., 2002). FastICA estimates a set of components that are maximally independent using a non-Gaussianity measurement and therefore provides a global separation measure, while JADE uses the measure of diagonalization for component separation that did not provide good results.

In this chapter, we test the most common methods of blind source separation in the state of art. We develop a formal comparison of three methods of the Independent Component

Analysis for fault identification in rotative machines. The results show that they are a useful tool for revealing fault patterns. However, note that each of the methods employed provides different features of failure.

4 Enhanced fault localization in rolling element bearings using multiple constrained ICA

4.1. Introduction

Rolling element bearing (REB) is one of the most critical mechanisms in rotating machinery, and its failure implies in many cases substantial economic losses and catastrophic incidents in industry. For that reason, the permanent monitoring of REB is crucial for early fault detection (Randall and Antoni, 2011). In recent decades, the vibration signal processing arises as a useful technique to diagnose a bearing fault, where the tracking of its degradation process could provide valuable information reducing a sudden damage (Kumar and Kumar, 2018; Cardona-Morales and Castellanos-Dominguez, 2018). For the sake of simplification, the maintenance engineers prefer to utilize the envelope analysis, also called *classical analysis*, since, by visual inspection of the envelope spectrum, it is possible relating specific frequencies with the corresponding bearing failures (Taylor, 1994). Nonetheless, a false diagnostic is feasible in machines with several REBs because it is possible to wrongly sense a broken bearing from the other bearing location (Kass et al., 2019). Therefore, a source fault localization strategy is needed to increase the condition monitoring accuracy in complex machines and systems.

Independent component analysis (ICA) was developed to deal with blind source separation (BSS) tasks, in particular, instantaneous mixtures of several statistical independent components (Hyvärinen and Oja, 2000). Nevertheless, ICA has two ambiguities (scaling and permutation) that limit its usage, it means that the IC amplitudes change on each algorithm iteration, and the appropriate IC order cannot be determined. In that regards, some authors had been suggested to modify the BSS paradigm into extracting a single signal or component that encloses the bearing fault dynamic, also called *blind signal extraction* (Antoni, 2005; Cardona-Morales et al., 2018; Smith and Randall, 2015a). Among the multiple methods adopted in the state-of-the-art related with BSE, constraint independent component analysis (cICA) or ICA with reference (Lu and Rajapakse, 2006; Zhang, 2008) has been designed to extract a single component (from the mixture) with a characteristic pattern. In REB fault detection was utilized cICA in Wang et al. (2011); Jing et al. (2014), however, the method was addressed to obtain a specific bearing fault, and in Yang et al. (2018), cICA

was employed over non-stationary vibration signals where the machine speed is fluctuating. Nonetheless, the bearing fault localization is still an open issue that increases its relevance nowadays, since the technology development allows think into sensor networks that provide a simultaneously monitoring.
Nonetheless, the ICA model has two ambiguities (i.e. scaling and permutation) that limit its application. The scaling implies a mismatch between the IC and source amplitudes, yet it is insignificant since in bearing fault identification is more relevant the waveform and its spectral information. In change, the permutation ambiguity does not allow to establish the correct IC order, and it affects the failure localization. This drawback could be alleviate using the constrained ICA inasmuch it retrieves a single IC that comprises the hidden pattern.
In this chapter, we propose a REB fault diagnostic method that allows identifying and locating the fault. Our approach is based on the extraction of a characteristic envelope that describes the faulty process because we use as a mixture the envelope signal of several measurements. To this end, we modify the cost function of cICA by the Kullback-Leibler divergence, introducing multiple constraints associated with any bearing fault. Finally, we identify and locate the faulty bearing using the Jaccard similarity computed between the obtained IC and the envelope signals of each sensor. The mcICA algorithm is tested both synthetic and real-world datasets, and compared against the classical analysis. The last case is characterized by two run-to-failure tests where it is important to determine the bearing fault and where it occurs.
The agenda is as follows: Firstly, Section 4.2 describes the mathematical background of cICA and it is present the measures of similarity used for bearing fault localization. The developed experiments and achieved results are described in Section 4.3, respectively. Finally, the discussion are provided in Section 4.4.

4.2. Methods

4.2.1. Constrained Independent component analysis (cICA)

Independent Component Analysis (ICA) aims at separating a multivariate vibration signal $\hat{\boldsymbol{x}}(t) = \{x_m(t) : m \in M\}$, where $\hat{\boldsymbol{x}}(t) \in \mathcal{L}^2(t)$, going along with the signal $\boldsymbol{x}(t)$, lasting $T \in \mathbb{R}$ and measured from M accelerometers, into N additive hidden patterns (or sources) that are assumed to be non-Gaussian and independent from each other. Based on their underlying higher-order statistical structures, the unknown components are linearly combined, that is $\hat{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{s}(t)$, through the full rank matrix $\boldsymbol{A} \in \mathbb{R}^{M \times N}$ that is unknown also. Therefore, the ICA approaches estimate a de-mixing vector $\boldsymbol{w} \in \boldsymbol{W}^{N \times M}$, such that, $\boldsymbol{W} = \boldsymbol{A}^{-1}$ and noting that for an orthogonal matrix $\boldsymbol{W}^{-1} = \boldsymbol{W}^{\top}$, thus, $\boldsymbol{y} = \boldsymbol{w}^{\top}\boldsymbol{A}\boldsymbol{s}$, using a two-step optimizing procedure: A linear whitening transformation is applied to the measured time series to get a

set of uncorrelated random signals set, each one having unit variance. Next, the optimization objective function of ICA is defined, according to a selected high-order statistical maximum criterion.

In fault vibration analysis, however, only an small source subset is of interest, from which some properties or a priori available information can be incorporated into the cost function, using constraints regarding the closeness $g(\boldsymbol{y})$ and scale $h(\boldsymbol{y})$ and resulting in the constrained ICA algorithm (cICA):

$$\mathcal{J}(\boldsymbol{y}) \approx \rho\Big(\mathbb{E}\left\{G(\boldsymbol{y})\right\} - \mathbb{E}\left\{G(\boldsymbol{\nu})\right\}\Big)^2, \tag{4-1a}$$

$$\text{s.t.: } g(\boldsymbol{w}) = \epsilon(\boldsymbol{y},\boldsymbol{r}) - \xi;\ h(\boldsymbol{w}) = \mathbb{E}\left\{\boldsymbol{y}^2\right\} - 1 = 0, \mathbb{E}\left\{\boldsymbol{r}^2\right\} - 1 = 0 \tag{4-1b}$$

where $\mathbb{E}\left\{\cdot\right\}$ stands for the expectation operator, $\rho \in \mathbb{R}^+$ is a positive-definite constant, $\boldsymbol{\nu} \in \mathbb{R}^{M\times 1}$ is a Gaussian variable having zero mean and unit variance, $G(\cdot)$ is any non-quadratic function, ϵ is the similarity measure between the independent component $\boldsymbol{y}$ and the reference signal $\boldsymbol{r} \in \mathbb{R}^{M\times 1}$, and $\xi \in \mathbb{R}^+$ is a certain similarity threshold.

Nevertheless, cICA performs the cost function minimization in Eq. (4-1a) to reduce to the smallest possible amount the difference between acquired data and the reference signal for computing the optimum de-mixing vector, $\boldsymbol{w}^*$. Instead, we propose to minimize the distance D_{YR} between the spectral power estimates of sources $\boldsymbol{y}$ and multiple references $\{\boldsymbol{r}_p : p \in P\}$, as follows:

$$\boldsymbol{w}^* = \min_{\forall p} \arg\min_{\boldsymbol{w}} D_{YR}(Y(k)||R_p(k)). \tag{4-2}$$

where $Y(k)$ and $R(k)$ are the spectral density of the time series $Y(k) = \mathcal{F}\{\boldsymbol{y}(t)\} = \mathcal{F}\{\boldsymbol{w}^\top \hat{\boldsymbol{x}}(t)\}$ and $R_p(k) = \mathcal{F}\{\boldsymbol{r}_p(t)\}$. Notation $\mathcal{F}\{\cdot\}$ stands for the Fourier Transform.

On the other hand, the widely-known mean squared error (MSE) is frequently applied as the distance D_{YR}, when assuming the stationarity of acquired data. Since the presence of faults induces non-stationarities, we introduce the use of Kullback-Leibler divergence to measure the discrepancy between a wide range of fault-derived distributions. Therefore, relying on Eq. (4-2), we devise a multiple-contrained ICA (termed *mcICA*) as follows:

$$\max_{n,p} \quad \rho\left(\mathbb{E}\left\{G(\boldsymbol{w}^\top \hat{\boldsymbol{x}})\right\} - \mathbb{E}\left\{G(\boldsymbol{\nu})\right\}\right)^2 - \lambda D_{YR}(\mathcal{F}\{\boldsymbol{w}^\top \hat{\boldsymbol{x}}\}||R_p) \tag{4-3a}$$

$$\text{s.t.:} \quad \mathbb{E}\left\{(\boldsymbol{w}^\top \hat{\boldsymbol{x}})^2\right\} = 1, \qquad \mathbb{E}\left\{r_p^2\right\} = 1 \quad \forall p \tag{4-3b}$$

Under the equality restrictions in Eq. (4-3b), the contrained optimization problem of Eq. (4-3a) is solved by the corresponding augmented Lagrangian function through a Newton-like learning algorithm as detailed in Huang and Mi (2007).

4.2.2. Bearing fault identification and localization

We propose a methodology for the bearing fault identification and localization that described in Fig. 4-1.

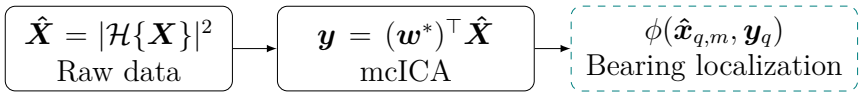


Figure 4-1: Proposed approach for fault localization using multiple constrained ICA.

Regarding the bearing fault identification, we fix a threshold to $thres = \mu + 6\sigma$, where the mean and the standard deviation of the machine vibration level are computed using the RMS or kurtosis values of the first two hundred recordings (Abboud et al., 2019). Thereby, a bearing fault exists when the vibration level indicator exceeds the threshold. Also, an indicator associated with each p-th reference signal is obtained which allows us to determine the type of fault. Therefore, a fault localization stage is required since it is unknown where the failure occurs.

With the purpose to perform the bearing fault localization, we compare the measures of similarity exhibited in Table 4-1. We estimate Cross-correlation Power Spectral Density (CPSD) that is an extension of the correlation coefficient. CPSD searchs powerful components locally concentrated in the spectral density of signal (Sierra-Alonso et al., 2014). For its parts, Jaccard is a measure of similarity between finite sets (Giusti and Batista, 2013). Further, we use Mean Squared Error (MSE) and Cross-correlation.

Table 4-1: Measures of similarity

CPSD	$\frac{1}{2LFTF}\sum_{l=1}^{L} Y(k)\|_\alpha$
Jaccard	$\frac{\sum_{t=1}^{T}(\boldsymbol{y}(t)-\boldsymbol{x}(t))^2}{\sum_{t=1}^{T}(\boldsymbol{y}^2(t)+\hat{\boldsymbol{x}}^2(t)-\boldsymbol{y}(t)\hat{\boldsymbol{x}}(t))^2}$
MSE	$\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{y}(t)-\hat{\boldsymbol{x}}(t))^2$
Cross-correlation	$\mathbb{E}\left\{\boldsymbol{y}(t)\hat{\boldsymbol{x}}^\top(t)\right\}$

where $Y = \mathcal{F}\{\mathbb{E}\left\{\boldsymbol{y}(t)\hat{\boldsymbol{x}}^\top(t)\right\}\}$, L is de number of harmonics, $\alpha \in f_i \pm FTF$, f_i is $BPFO$ or $BPFI$. For all measures, we calculate the similarity between $\boldsymbol{y}_q(t)$ and $\hat{\boldsymbol{x}}_{q,m}(t)$ where $q \in Q$ represent the register that exceeds the threshold, Q is the number total of the registers and m is each of the accelerometers. Finally, we estimate the percent of accuracy acc_j for the j-th measures of similarity as seen in Algorithm 1.

Algorithm 1 Weighted accuracy of bearing fault localization.

1. To find the register q that exceeds the threshold γ.

   for q **to** Q:

2. To compute the signal reconstructed by *mcICA* from the record (fault identification):

$$\boldsymbol{y}_q = (\boldsymbol{w}^*)^\top \hat{\boldsymbol{X}}_{\boldsymbol{q}}$$

3. To calculate proximity measures for each of the accelerometers m, where $m = [1, \ldots, M]$ (fault localization):

$$\phi_m(\hat{\boldsymbol{x}}_{q,m}, \boldsymbol{y}_q))$$

4. To assign a vote for bearing with the maximum index of d_m:

$$l_{q,m} = \max(\phi_m(\hat{\boldsymbol{x}}_{q,m}, \boldsymbol{y}_q)) \ \ where \ \ l_{q,m} = [1, \ldots, M]$$

   end for

5. To estimate the performance of the proposed method:

$$acc_m = \mathbb{E}\{l_{q,m}\} \ \ \forall m$$

4.3. Experimental Set-Up

For evaluation purposes, the *mcICA* algorithm comprises the following stages (see Fig. **4-1**): *i)* computing the envelope of measured signal vibrations using the Hilbert transform; *ii)* Data-driven *mcICA* decomposition, using a probabilistic distance-based cost function; And *iii)* identification and localization of faults, measuring the similarity between the bearing signals and their estimated latent components.

The proposed *mcICA* algorithm is validated on two different datasets. First, a set of synthetic signals to demonstrate the blind separability, and second, a run-to-failure test of rolling element bearings acquired from a test rig to early localization of a bearing fault.

4.3.1. Synthetic separable data

Several simulated faulty signals are generated by mixing three vibration sources, adding noise with zero-mean and unit-variance as follows:

$$\begin{cases} \boldsymbol{s}_1 = & \sin(2\pi f_1 t + \cos(2\pi f_2 t)) \\ \boldsymbol{s}_2 = & \cos(2\pi f_3 t) \\ \boldsymbol{s}_3 = & \eta(t),\, \mathcal{N}_\eta(0,1) \\ \boldsymbol{s}_4 = & \sum_{i=1}^{M}(A\cos(2\pi f_r t + \alpha_A) + C) \\ & \exp(-B(t - iT_b - \tau))\cos(2\pi f_n(t - iT_b - \tau) + \alpha_w) \end{cases} \tag{4-4}$$

where $\boldsymbol{s}_1$ is a frequency-modulated signal, $\boldsymbol{s}_2$ is a plain cosine, $\boldsymbol{s}_3$ is a random Gaussian noise, and $\boldsymbol{s}_4$ is a simulated rolling element bearing signal with inner race fault that emulates a surface deterioration such as spalling fatigue and abrasive wear, rolling element random slip, tolerance, and amplitude modulation (Randall et al., 2001). In this case, the following parameter values are fixed: $f_1 = 610\,Hz$, $f_2 = 540\,Hz$, $f_3 = 280\,Hz$, $f_r = 33.33\,Hz$ is the rotary frequency, which is generally equated to the rotating speed of the shaft; $f_b = 296\,Hz$ is the fault frequency associated to a bearing effect; $f_n = 2000\,Hz$ is the natural frequency associated to system or bearing; $A = 3$ is the resonance intensity; $B = 800$ is the system resonance damping coefficient; T_b is the attenuated damping oscillation with mean impulse period $T_b = 1/f_b$; $\tau = 0.01$ is an small fluctuation around T_b due to the presence of slip. For the sake of simplicity, the phase parameters are a zero-valued constant: $C = 0$, $\alpha_A = 0$, $\alpha_w = 0$.

The validating signal set is generated using 20480 samples at a sampling frequency $20\,kHz$, including the multiple channels to simulate the tested faults and holding non-stationary behavior as shown by spectrograms in Fig. **4-2**.

During testing of synthetic data, the multiple references and simulated signal are fed into the mcICA algorithm. Since the reference signal is a crucial point to perform an accurate IC reconstruction, a square pulse train is employed, having the characteristic fault frequency as suggested in Wang et al. (2011); Lu and Rajapakse (2006). Using the Jaccard dissimilarity, the extracted IC component is $\boldsymbol{s_4}$ that matches the simulated bearing source with inner race fault as shown in Fig. **4-3**, reaching a high value of similarity (correlation index close to 0.999).

4.3.2. Bearing fault localization on real-world data

The used bearing data were obtained from Prognostics Center Excellence through the prognostic data repository provided by Intelligent Maintenance System (*IMS*), which holds a test-to-failure experiment performed on a bearing test rig, including four Rexnord ZA-2115 double-row bearings on one shaft driven by an AC motor and coupled by rub belts (Qiu

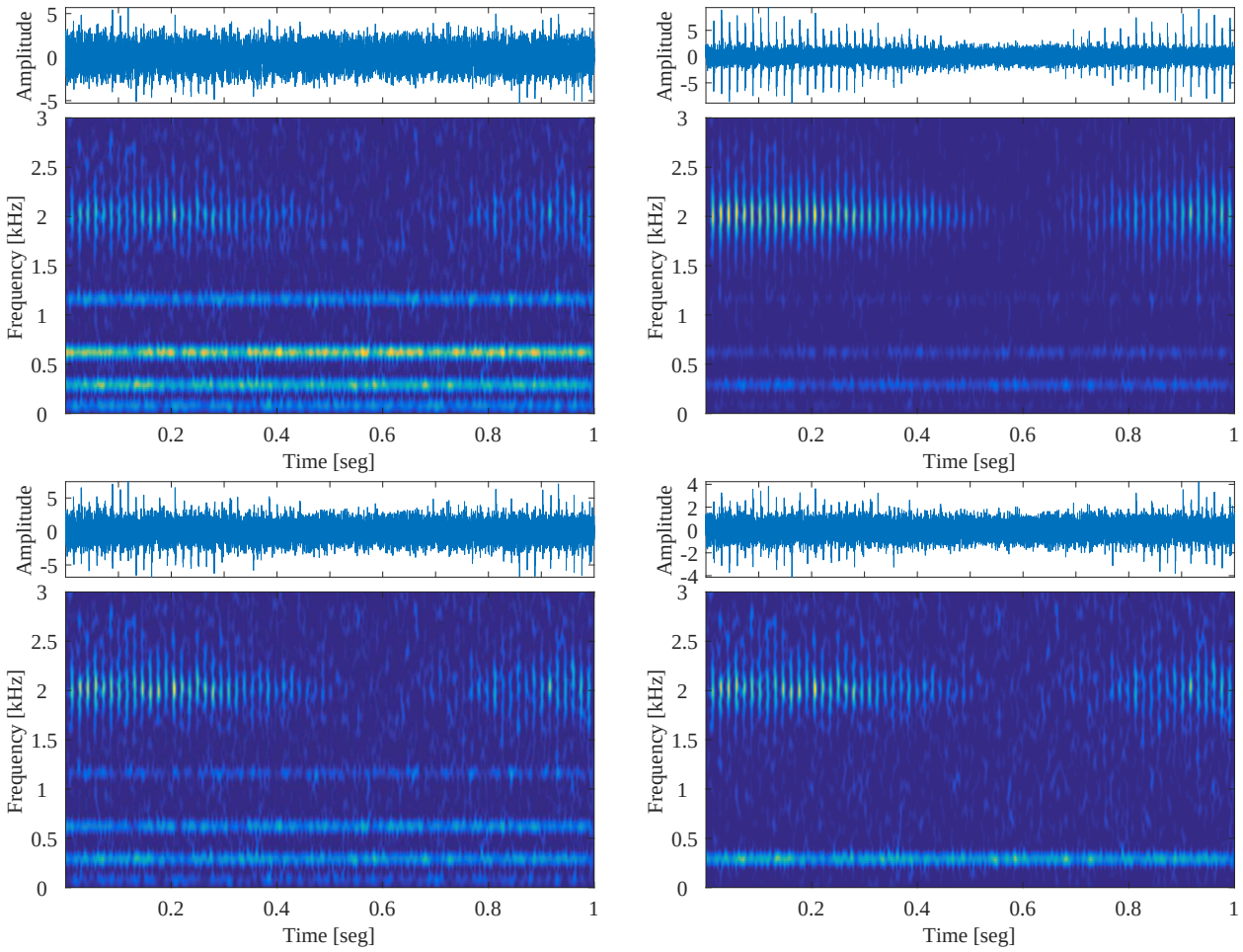


Figure 4-2: Mixing source signals generated by the synthetic data. It is shown the signal in time domain and its time-frequency representation (spectrogram with hamming window, 50 % overlap, and 2048 bins)

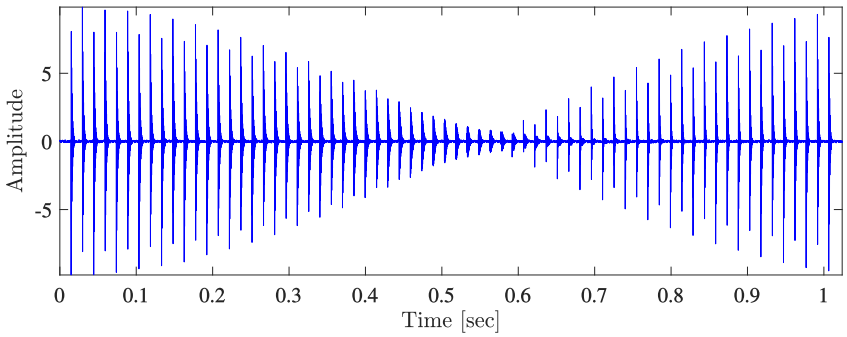


Figure 4-3: IC reconstruction obtained using our mcICA algorithm. Jaccard index is 1 and correlation index is 0.9993

et al., 2006). During the experiment, the rotation speed remained constant at 2000 r/min, and each bearing had 16 rollers per row, a diameter of 2.815 in, roller diameter of 0.331 in, and a tapered contact angle of 15.17°. A PCB 353B33 High Sensitivity Quartz ICP Accelerometer was installed on each bearing housing. The data sampling rate was adjusted to 20 kHz, and collecting one vibration record every 10 min.

In this work, the test #2 and #3 are used for validation, holding an outer race defect in the bearings labeled as #1 and #3, respectively. The bearing fault frequencies are approximated by the geometrical properties of the rolling element bearings and shaft rotational speed f_r as shown in Table **3-1**. Therefore, we build a set of four square pulse trains based on the expected fault frequencies as follows: $BPFO = 236.4\,Hz$, $BPFI = 296.9\,Hz$, $BSF = 139.9\,Hz$, and $FTF = 14.7\,Hz$. It is worth noting that the frequencies mentioned above may present a frequency slip up to $1 - 2\,\%$ (Smith and Randall, 2015a).

To characterize the machine health state on real-world data, we feed the signal envelope of each bearing and the multiple references into the mcICA algorithm, extracting the IC reconstruction that holds the dominant fault pattern. Then, we assess the machine vibration level using the kurtosis value over the obtained IC spectrum, namely, on the frequency bands associated with the bearing fault frequencies described in Table **3-1** . For the sake of comparison, we also carry out the classical envelope analysis, for which the RMS value is the indicator of vibration level.

In order to localize the damaged bearing, different similarity measures are estimated because it is possible that several rolling element bearings evince as the lowest Jaccard dissimilarity value computed between the IC and signal envelope for each bearing. Therefore, the accuracy provided by the mcICA approach is estimated from the recordings that exceed the threshold, such as show in Algorithm 1.

4.3.3. Case 1: Slow-growing of bearing fault

In the test #2, Fig. **4-4** shows that the classical analysis, as well as mcICA approach, detect the fault at recording 539 exceeding the threshold. Yet, the former analysis demands the knowledge about the bearing that will fail, and this situation is far from being normal. On the opposite, the proposed algorithm neglects this requirement since it runs over all bearings simultaneously. Note that the detected active fault is an outer race defect as reported in the database description.

For visual inspection, Fig. **4-5** displays the squared envelope spectrum computed for all bearings at the record 539, clearly showing that the characteristic bearing fault frequency ($BPFO$) is present in the bearing #1. Nonetheless, the bearing #2 also has some fault signs like the sidebands of the BPFO harmonics, and hence, a classical analysis could generate a false alarm concerning the faulty bearing. In contrast, the reconstructed IC obtained using our approach allows identifying the outer race defect selecting the proper damaged bearing.

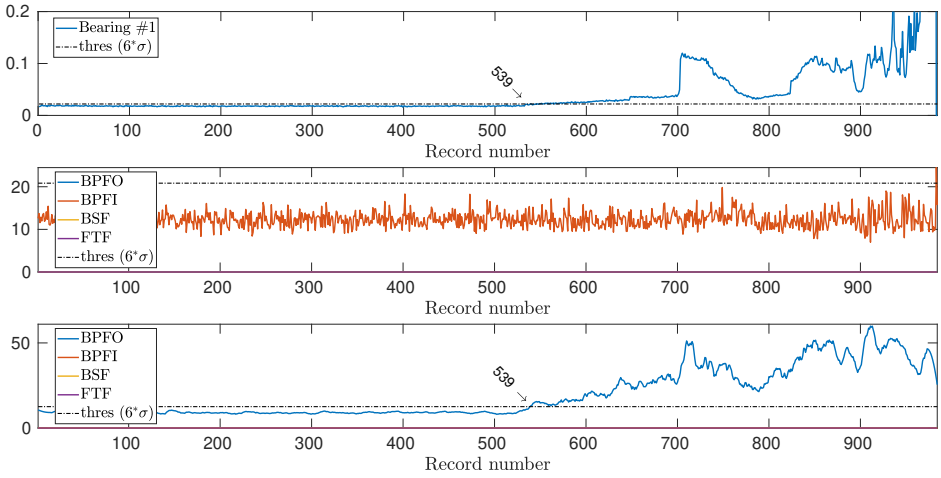


Figure 4-4: Bearing fault trending of the test #2, using classical analysis (top), cICA (middle) and mcICA (bottom).

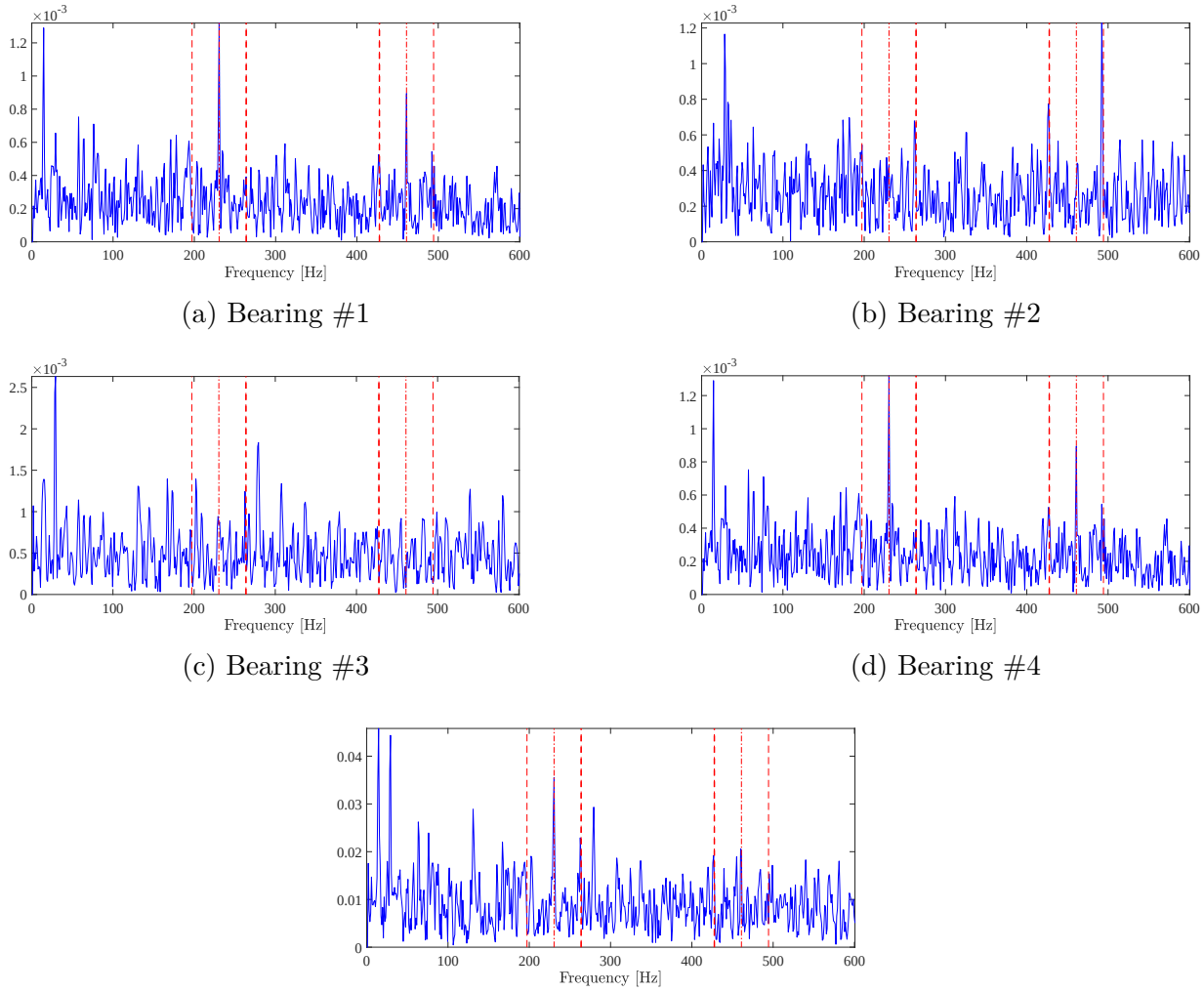
Table 4-2: Weighted accuracy of bearing fault localization for test #2.

Measures (%)	Bearing #1	Bearing #2	Bearing #3	Bearing #4
CPSD	**95.74**	0.45	2.24	1.57
Jaccard	**69.51**	3.81	26.23	0.45
MSE	**66.59**	4.71	28.25	0.45
Cross-correlation	32.51	20.63	**37.44**	9.42

4.3.4. Case 2: Fast-growing of bearing fault

In the experiment in test #3, as it is evinced in Fig. 4-6, the $BPFI$, BSF, and FTF fault holds inactive whereas outer race fault is activated. Particularly, the RMS analysis reaches the threshold at the record 5977, and it remains stable until the record 6071, where it is present a growing degradation process. In change, our approach exceeds the threshold at the record 6068, showing an accelerate increment in the damaged bearing.

Additionally, we perform the analysis of the envelope spectrum of the bearing #3 for the records 5977, 6068 and 6071 using both the classical analysis and our mcICA approach, since those records exceed the threshold (see Fig. 4-7). It is possible to observe in Fig. 4-7a that the record 5977 does not exhibit a clear outer race fault pattern, which could be considered a false alarm. Nevertheless, at the record 6068 reported by our algorithm, the IC presents significant sidebands of $BPFO$, indicating an early faulty process (see Fig. 4-7b). Finally,



(a) Bearing #1

(b) Bearing #2

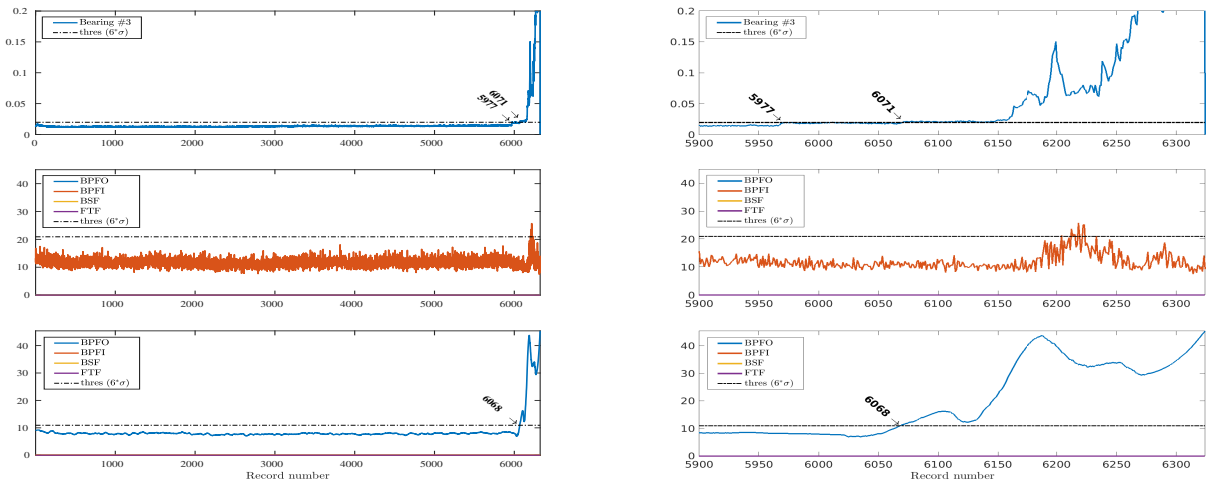
(c) Bearing #3

(d) Bearing #4

(e) IC reconstructed

Figure 4-5: Squared envelope spectrum for all bearings and IC by the record 539 of the test #2. Adjusted harmonic cursors multiples of $BPFO$ (230.5 Hz) and sideband cursors multiples of $BPFO \pm fr$ (230.5 ± 33.3 Hz) are displayed in red color.

the record 6071 showed in Fig. 4-7c exhibits a dominant spectral component at $BPFO$ for both approaches.



(a) Trending of the test #3

(b) Zoom trending of the test #3

Figure 4-6: Bearing fault trending of the test #3, using classical analysis (top), cICA (middle) and mcICA (bottom).

Table 4-3: Weighted accuracy of bearing fault localization for test #3.

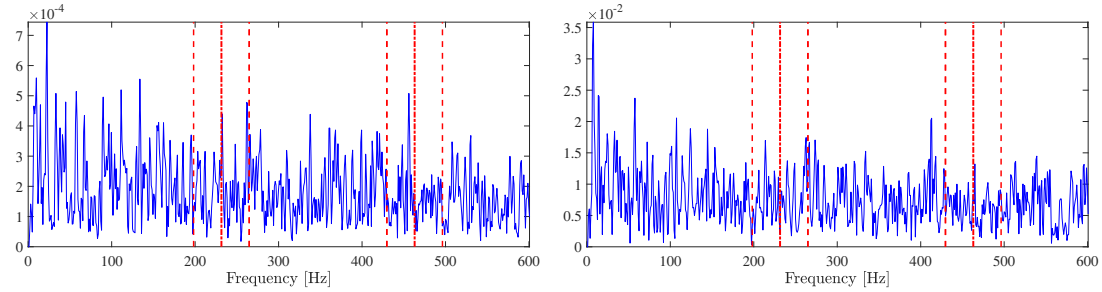
Measures (%)	Bearing #1	Bearing #2	Bearing #3	Bearing #4
CPSD	1.56	0.78	26.85	**72.82**
Jaccard	7.00	14.40	**78.21**	0.39
MSE	7.78	15.18	**74.71**	2.33
Cross-correlation	21.40	23.74	26.85	**28.02**

Fig. 4-8 illustrates the achieved localization performance in the case of 6180, at which the fault reaches steady behavior. Due to the fault is present at all sensors, its localization maybe not simple using the proximity measurements computed from the signal envelopes.

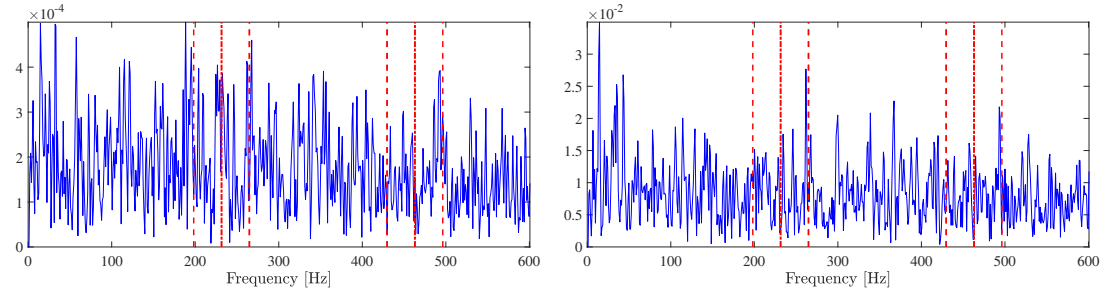
4.4. Discussion

The introduced mcICA algorithm applied to real-world data allows highlighting some aspects related to its performance to detect a machine degradation process, to identify the type of bearing fault and to spatially locate the damaged bearing.

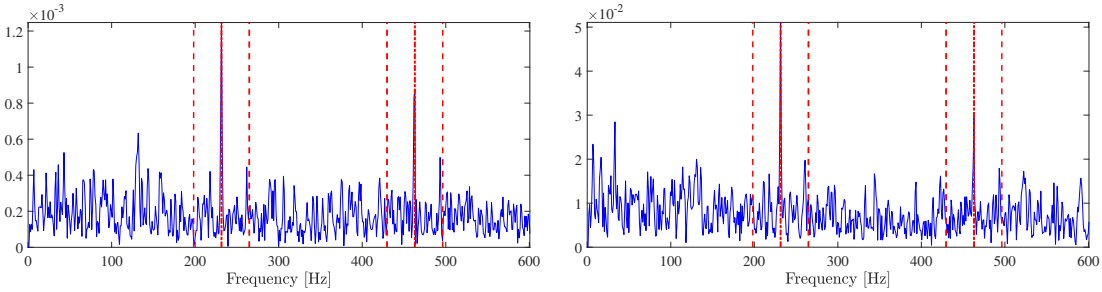
Regarding the degradation process, in the test #2 showed in Fig. 4-4, our approach holds a trending that continually increases until it reaches a maximum at the record 941, where the vibration level decreases due to that the fault surface has extended during the degradation



(a) Record 5977



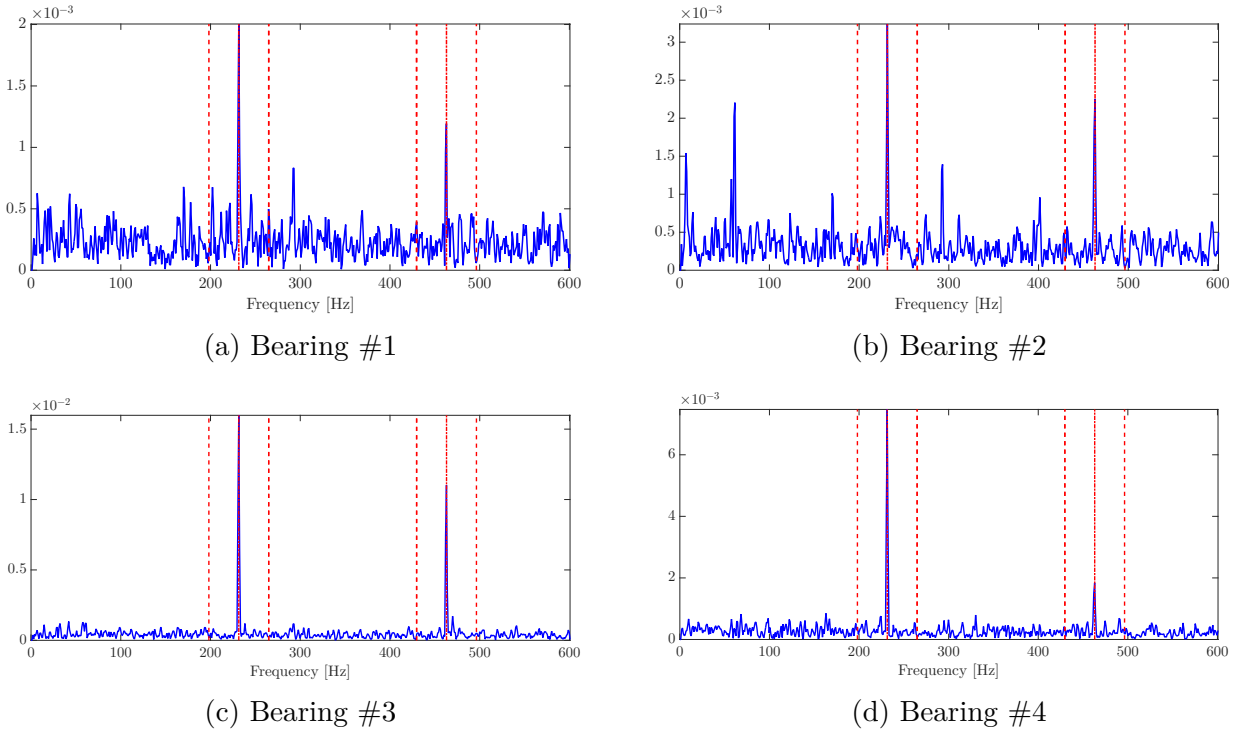
(b) Record 6068



(c) Record 6071

Figure 4-7: Squared envelope spectrum for bearing #3 both RMS (left) and IC (right) analysis in the test #3. Adjusted harmonic cursors multiples of $BPFO$ (231.5 Hz) and sideband cursors multiples of $BPFO \pm fr$ (231.5 ± 33.3 Hz) are displayed in red color.

process. Hence, it smooths the impulsive excitation (Abboud et al., 2019). In change, the classical analysis evinces a slow-growing process that rapidly grows up until the record #700, which could be catastrophic for the machine because it does not provide enough time for the maintenance. Besides, when the failure moves forward, i.e., around the record 790, the low vibration level hides the actual degradation stage. A similar dynamic is presented in the test #3 (Fig. **4-6**), where our approach provides a growing trend, while RMS detection holds constant for several records on the threshold value, displaying a clear degradation process 172 records from the first fault detection flag. Therefore, the obtained outcome of the record 5977 could be considered as an outlier since it does not present a clear fault pattern.



(a) Bearing #1 (b) Bearing #2

(c) Bearing #3 (d) Bearing #4

Figure 4-8: Squared envelope spectrum for all bearings by the record 6180 of the test #3. Adjusted harmonic cursors multiples of $BPFO$ (231.5 Hz) and sideband cursors multiples of $BPFO \pm fr$ (231.5 ± 33.3 Hz) are displayed in red color.

Concerning the fault identification task, in the test #2, an outer race bearing fault is found using both the classical and *mcICA* approaches, while cICA does not correctly identify the failure. Nonetheless, the traditional analysis needs carrying out an envelope analysis and visual inspection to distinguish the type of fault, yet our approach matches the proper reference to automatically identifying it. Therefore, it allows optimizing the diagnostic time in the maintenance task. Additionally, it is worth noting that the proposed *mcICA* algorithm has a low time consumption since it utilizes optimized procedures like Fast-ICA and KL-divergence instead of MSE cost function. This change allows comparing the power spectral densities of the reference signals and the IC reconstruction, highlighting the characteristic fault frequencies and achieving a better algorithm performance. In contrast, conventional cICA presents several limitations such as *i)* taking into account only one reference signal that makes the selection more complex and *ii)* requiring a priori knowledge about the source to be retrieved.

With respect to fault localization, the employed measures show acceptable behavior in both tests. In the test #2, it should be noted that the success rate of the CPSD measure is the best (95.74 %), determining the bearing location where the damage takes place. However, that measure does not reach the same performance in the test #3, since the bearing degradation arises rapidly, and it is observed at nearby sensors, as shown in Fig. 4-8. Therefore, CPSD

performance is affected due to the high similarity between the power spectral densities. Nevertheless, the measure determines that the fault is in bearing #3 or #4 with success rate of 26.85 % and 72.82 %, respectively, which discards the bearings #1 with 1.56 % and #2 with 0.78 %. In particular, JI is best for test #3 since, as a measure of sets, it takes advantage of the join and intersection of the data to associate the reconstructed IC with the bearing #3. Finally, CC fails to determine in any case the origin of the fault due to the vibration signals are highly correlated in time-domain, as evidenced in the closeness of the percentages obtained (Tables **4-2** and **4-3**).

In this chapter, we propose an algorithm for fault localization in rolling element bearings. We modified the cost function of the traditional cICA, in order to take advantage of the characteristics of faults in the frequency space. Besides, we introduce multiple references that help with the fault identification. Thus, the reconstructed IC contains clear fault patterns, facilitating the association with the vibration signal coming from the bearings. Therefore, it is necessary to look for a measure of similarity that allows the location of the fault. Finally, the proposed measures show a high performance in the location of the fault.

5 Conclusion and Future Work

5.1. Conclusion

Firstly, feature extraction based on unsupervised and supervised learning methods to diagnose gearbox faults is presented. This model employed different features from the time-domain, the frequency-domain, and the time-frequency domain to extract the gear health conditions. The classifier based on k nearest neighbor classification algorithm to construct a reliable and accurate fault diagnosis method. The data were collected from the gears under different loads and motor speeds to generate typical cases of gear faults. The analysis results on the gear data show that the proposed methodology enables the classification of gear damage.

Secondly, it was proposed an application of ICA methods for multi-fault rotating systems getting essential results in the separation of faults signals acquired by accelerometers, providing relevant information on the discrimination of faults in different supports. The three algorithms showed a performance acceptable for the task of BSS. However, the algorithm FastICA showed a top performance about the others, the support $S2$ indicate very clearly the presence of faults in both axis, horizontal and vertical for fault inner race and the support $S1$ for fault ball, as well as we note that the algorithm powers the information of the fault being a fundamental part because it allows us to find the support to which the fault corresponds. Moreover, for fault ball pass frequency, out race, charges no importance to apply the methodology proposal because it was only recorded by the support $S1$, which means that no mixture of the fault signals was presented; therefore we have no problem of BSS to analyze with the algorithms implemented in this document.

Thirdly, a fault localization method in machines or systems with rolling element bearings is presented. The methodology is based on an improved version of the constrained ICA algorithm, where multiple constraints are fed into the optimization problem. The change of the cost function by a divergence measure allows inferring that our approach could provide relevant information under non-stationary operating conditions. Utilizing a dissimilarity measure was feasibly detecting the faulty bearing among the multiple measurements, which holds open the blind source separation issue for industrial applications.

5.2. Future Work

The proposed methodologies perform well for the databases under consideration. However, validation of these with other types of signals would allow determining the capacity of the methodology to extrapolate its operation to other rotary machine configurations.

The fault localization method provides us detail information spacial about the machine. This permits us to have a complete panorama of the phenomena that can affect the correct functioning of the elements present in the bearings, also provides us with a tool to diagnose the health of the machine. Therefore, we propose to test our mcICA approach in the monitoring of different industrial applications since it is promising in fault detection and localization.

Finally, it is proposed as future work to test different classifiers, since there are more advanced classification techniques (such as neural networks among others), which can take better advantage of the extraction of features performed previously. Besides, it is proposed to make a feature selection based on the relevances obtained in the development of this work.

Bibliografía

Abboud, D., M. Elbadaoui, W. A. Smith, and R. B. Randall
2019. Advanced bearing diagnostics: A comparative study of two powerful approaches. *Mechanical Systems and Signal Processing*.

Alvarez-Meza, A. M., A. Orozco-Gutierrez, and G. Castellanos-Dominguez
2017. Kernel-based relevance analysis with enhanced interpretability for detection of brain activity patterns. *Frontiers in neuroscience*, 11:550.

Antoni, J.
2005. Blind separation of vibration components: Principles and demonstrations. *Mechanical Systems and Signal Processing*, 19(6):1166 – 1180. Special Issue: Blind Source Separation.

Atoui, I., H. Meradi, R. Boulkroune, and R. Saidi
2013. Fault detection and diagnosis in rotating machinery by vibration monitoring using fft and wavelet techniques. In *2013 8th international workshop on systems, signal processing and their applications (WoSSPA)*, Pp. 401–406. IEEE.

Belouchrani, A., K. Abed-Meraim, J.-F. Cardoso, and E. Moulines
1997. A blind source separation technique using second-order statistics. *IEEE Transactions on signal processing*, 45(2):434–444.

Caesarendra, W. and T. Tjahjowidodo
2017. A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing. *Machines*, 5(4):21.

Cardona-Morales, O.
2011. Análisis tiempo-frecuencia de señales de vibraciones mecánicas para la detección de fallos en máquinas rotativas. *masters, Universidad Nacional de Colombia-Sede Manizales*.

Cardona-Morales, O. and G. Castellanos-Dominguez
2018. Fault diagnostic of machines under variable speed operating conditions using order tracking and novelty detection. In *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*, A. Timofiejczuk, F. Chaari, R. Zimroz, W. Bartelmus, and M. Haddar, eds., Pp. 179–189, Cham. Springer International Publishing.

Cardona-Morales, O., E. F. Sierra-Alonso, and G. Castellanos-Dominguez
2018. Bearing fault identification based on blind extraction of cyclostationary signals using order tracking. In *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*, A. Timofiejczuk, F. Chaari, R. Zimroz, W. Bartelmus, and M. Haddar, eds., Pp. 191–201, Cham. Springer International Publishing.

Cardoso, J.-F.
1999. High-order contrasts for independent component analysis. *Neural computation*, 11(1):157–192.

Cardoso, J.-F. and A. Souloumiac
1993. Blind beamforming for non-gaussian signals. In *IEE proceedings F (radar and signal processing)*, volume 140, Pp. 362–370. IET.

Cerrada, M., J. Aguilar, J. Altamiranda, and R.-V. Sánchez
2019. A hybrid heuristic algorithm for evolving models in simultaneous scenarios of classification and clustering. *Knowledge and Information Systems*, 61(2):755–798.

Cerrada, M., R.-V. Sánchez, F. Pacheco, D. Cabrera, G. Zurita, and C. Li
2016. Hierarchical feature selection based on relative dependency for gear fault diagnosis. *Applied Intelligence*, 44(3):687–703.

Chen, J., Z. Li, J. Pan, G. Chen, Y. Zi, J. Yuan, B. Chen, and Z. He
2016. Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. *Mechanical systems and signal processing*, 70:1–35.

Choi, S., A. Cichocki, and A. Beloucharni
2002. Second order nonstationary source separation. *J. VLSI Signal Process. Syst.*, 32(1-2):93–104.

Fan, X. and M. J. Zuo
2006. Gearbox fault detection using hilbert and wavelet packet transform. *Mechanical Systems and Signal Processing*, 20(4):966 – 982.

Giusti, R. and G. E. Batista
2013. An empirical comparison of dissimilarity measures for time series classification. In *2013 Brazilian Conference on Intelligent Systems*, Pp. 82–88. IEEE.

Haidong, S., J. Hongkai, L. Xingqiu, and W. Shuaipeng
2018. Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine. *Knowledge-Based Systems*, 140:1–14.

Huang, D.-S. and J.-X. Mi
2007. A new constrained independent component analysis method. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 18(5):1532–1535.

Hyvarinen, A.
1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634.

Hyvärinen, A., J. Karhunen, and E. Oja
2004. *Independent component analysis*, volume 46. John Wiley & Sons.

Hyvärinen, A. and E. Oja
2000. Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430.

Igba, J., K. Alemzadeh, C. Durugbo, and E. T. Eiriksson
2016. Analysing rms and peak values of vibration signals for condition monitoring of wind turbine gearboxes. *Renewable Energy*, 91:90–106.

Isermann, R.
2005. Model-based fault-detection and diagnosis–status and applications. *Annual Reviews in control*, 29(1):71–85.

Isermann, R.
2006. *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Science & Business Media.

Jardine, A. K., D. Lin, and D. Banjevic
2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510.

Jiang, G., H. He, J. Yan, and P. Xie
2018. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox. *IEEE Transactions on Industrial Electronics*, 66(4):3196–3207.

Jing, L., M. Zhao, P. Li, and X. Xu
2017. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement*, 111:1–10.

Jing, S. X., S. T. Guo, J. F. Leng, and X. Y. Zhao
2014. Application of fault diagnosis method based on cica to gearbox. In *Mechatronics and Mechanical Engineering I*, volume 664 of *Applied Mechanics and Materials*, Pp. 148–152. Trans Tech Publications.

Kass, S., A. Raad, and J. Antoni
2019. Self-running fault diagnosis method for rolling element bearing. In *Mechanism, Machine, Robotics and Mechatronics Sciences*, R. Rizk and M. Awad, eds., Pp. 127–140, Cham. Springer International Publishing.

Kim, J.-S. and S.-K. Lee
2018. Identification of tooth fault in a gearbox based on cyclostationarity and empirical mode decomposition. *Structural Health Monitoring*, 17(3):494–513.

Kopsaftopoulos, F. and S. Fassois
2013. A functional model based statistical time series method for vibration based damage detection, localization, and magnitude estimation. *Mechanical Systems and Signal Processing*, 39(1-2):143–161.

Kumar, A. and R. Kumar
2018. Role of signal processing, modeling and decision making in the diagnosis of rolling element bearing defect: A review. *Journal of Nondestructive Evaluation*, 38(1):5.

Lei, Y., Z. He, and Y. Zi
2008. A new approach to intelligent fault diagnosis of rotating machinery. *Expert Systems with Applications*.

Lei, Y., M. J. Zuo, Z. He, and Y. Zi
2010. A multidimensional hybrid intelligent method for gear fault diagnosis. *Expert Systems with Applications*, 37(2):1419–1430.

Leng, J., S. Jing, and C. Luo
2018. Fault feature extraction of single-channel signal from gearbox based on eemd and cica. *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, 11(1):69–75.

Li, C., R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez
2015. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing*, 168:119–127.

Lu, W. and J. C. Rajapakse
2006. ICA with Reference. *Neurocomputing*, 69(16-18):2244–2257.

Medina, R., X. Alvarez, D. Jadán, M. Cerrada, R.-V. Sánchez, and J. C. Macancela
2017. Poincaré plot features from vibration signal for gearbox fault diagnosis. In *2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)*, Pp. 1–6. IEEE.

Miao, F., R. Zhao, X. Wang, and L. Jia
2020. A new fault feature extraction method for rotating machinery based on multiple sensors. *Sensors*, 20(6):1713.

Popescu, T. D.
2010. Blind separation of vibration signals and source change detection–application to machine monitoring. *Applied Mathematical Modelling*, 34(11):3408–3421.

Qiu, H., J. Lee, J. Lin, and G. Yu
2006. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of sound and vibration*, 289(4-5):1066–1090.

Randall, R. B. and J. Antoni
2011. Rolling element bearing diagnosticsâ€"a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485 – 520.

Randall, R. B., J. Antoni, and S. Chobsaard
2001. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15(5):945–962.

Shen, Z., Z. He, X. Chen, C. Sun, and Z. Liu
2012. A monotonic degradation assessment index of rolling bearings using fuzzy support vector data description and running time. *Sensors*, 12(8):10109–10135.

Sierra-Alonso, E. F., O. Cardona-Morales, C. D. Acosta-Medina, and G. Castellanos-Dominguez
2014. Spectral correlation measure for selecting intrinsic mode functions. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, E. Bayro-Corrochano and E. Hancock, eds., Pp. 231–238, Cham. Springer International Publishing.

Singh, D. S. and Q. Zhao
2016. Pseudo-fault signal assisted emd for fault detection and isolation in rotating machines. *Mechanical Systems and Signal Processing*, 81:202–218.

Smith, W. A. and R. B. Randall
2015a. Rolling element bearing diagnostics using the case western reserve university data: A benchmark study. *Mechanical Systems and Signal Processing*, 64-65:100 – 131.

Smith, W. A. and R. B. Randall
2015b. Rolling element bearing diagnostics using the case western reserve university data: A benchmark study. *Mechanical Systems and Signal Processing*, 64-65:100 – 131.

Tao, Y., X. Wang, R.-V. Sánchez, S. Yang, and Y. Bai
2019. Spur gear fault diagnosis using a multilayer gated recurrent unit approach with vibration signal. *IEEE Access*, 7:56880–56889.

Taylor, J.
1994. *The vibration analysis handbook*. Vibration Consultants.

Van Der Maaten, L., E. Postma, and J. Van den Herik
2009. Dimensionality reduction: a comparative. *J Mach Learn Res*, 10(66-71):13.

Wang, D. and P. W. Tse
2012. A new blind fault component separation algorithm for a single-channel mechanical signal mixture. *Journal of Sound and Vibration*, 331(22):4956 – 4970.

Wang, Z., J. Chen, G. Dong, and Y. Zhou
2011. Constrained independent component analysis and its application to machine fault diagnosis. *Mechanical Systems and Signal Processing*.

Wu, T. and X. Xiong
2019. Fault feature extraction method of rolling bearings based on citd and fastica. In *2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS)*, Pp. 849–854. IEEE.

Yan, X. and M. Jia
2018. A novel optimized svm classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing. *Neurocomputing*, 313:47–64.

Yang, T., Y. Guo, X. Wu, J. Na, and R.-F. Fung
2018. Fault feature extraction based on combination of envelope order tracking and cica for rolling element bearings. *Mechanical Systems and Signal Processing*, 113:131 – 144. SI: IMETI-MechElectro.

Yang, Y. and S. Nagarajaiah
2014. Blind identification of damage in time-varying systems using independent component analysis with wavelet transform. *mechanical systems and signal processing*, 47(1-2):3–20.

Yellin, D. and E. Weinstein
1994. Criteria for multichannel signal separation. *IEEE transactions on signal processing*, 42(8):2158–2168.

Zhang, Z. L.
2008. Morphologically constrained ICA for extracting weak temporally correlated signals. *Neurocomputing*, 71(7-9):1669–1679.

Žvokelj, M., S. Zupan, and I. Prebil
2016. Eemd-based multiscale ica method for slewing bearing fault detection and diagnosis. *Journal of Sound and Vibration*, 370:394–423.