

## Article

# Evaluation of Entropy Analysis as a Fault-Related Feature for Detecting Faults in Induction Motors and Their Kinematic Chain

Arturo Y. Jaen-Cuellar <sup>1</sup>, Juan J. Saucedo-Dorantes <sup>1</sup>, David A. Elvira-Ortiz <sup>1</sup>  
and Rene de J. Romero-Troncoso <sup>2,\*</sup>

<sup>1</sup> Academic Center of Advanced and Sustainable Technology (CATAS), Tequisquiapan Campus, Autonomous University of Queretaro, Carretera No. 120, San Juan del Río-Xilitla, Km 19+500, Tequisquiapan 76750, Queretaro, Mexico; ayjaen@hspdigital.org (A.Y.J.-C.); jsaucedo@hspdigital.org (J.J.S.-D.); delvira@hspdigital.org (D.A.E.-O.)

<sup>2</sup> HSPdigital CA-Mecatronica Engineering Faculty, Autonomous University of Queretaro, San Juan del Río 76806, Queretaro, Mexico

\* Correspondence: troncoso@hspdigital.org

**Abstract:** The induction motors found in industrial and commercial applications are responsible for most of the energy consumption in the world. These machines are widely used because of their advantages like high efficiency, robustness, and practicality; nevertheless, the occurrence of unexpected faults may affect their proper operation leading to unnecessary breakdowns with economic repercussions. For that reason, the development of methodologies that ensure their proper operation is very important, and in this sense, this paper presents an evaluation of signal entropy as an alternative fault-related feature for detecting faults in induction motors and their kinematic chain. The novelty and contribution lie in calculating a set of entropy-related features from vibration and stator current signals measured from an induction motor operating under different fault conditions. The aim of this work is to identify changes and trends in entropy-related features produced by faulty conditions such as broken rotor bars, damage in bearings, misalignment, unbalance, as well as different severities of uniform wear in gearboxes. The estimated entropy-related features are compared to other classical features in order to determine the sensitivity and potentiality of entropy in providing valuable information that could be useful in future work for developing a complete methodology for identifying and classifying faults. The performed analysis is applied to real experimental data acquired from a laboratory test bench and the obtained results depict that entropy-related features can provide significant information related to particular faults in induction motors and their kinematic chain.

**Keywords:** entropy analysis; fault detection; fault indicator; induction motors



**Citation:** Jaen-Cuellar, A.Y.; Saucedo-Dorantes, J.J.; Elvira-Ortiz, D.A.; Romero-Troncoso, R.d.J. Evaluation of Entropy Analysis as a Fault-Related Feature for Detecting Faults in Induction Motors and Their Kinematic Chain. *Electronics* **2024**, *13*, 1524. <https://doi.org/10.3390/electronics13081524>

Academic Editor: Davide Astolfi

Received: 12 March 2024

Revised: 11 April 2024

Accepted: 15 April 2024

Published: 17 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The industry has been constantly subjected to changes due to the emerging technology up to its relatively new version known as Industry 4.0, and despite the technological improvements with views towards efficient energy consumption, it still strongly depends on the use of electrical and mechanical elements such as gearboxes (GBs), induction motors (IMs), and their associated kinematic chains to propel most of their processes [1,2]. Some of the most common applications of GBs and IMs are [3–5] conveyors, pumps, cranes, fans, lifts, shredders, blowers, benders, and mixers, among others. The main characteristics of GB and IM are their general robustness, variety of prices and models, power–speed adjustable relationship, reliability, high-efficiency, easy maintenance, wide applicability, and hard environmental resistance [6,7]. Certainly, the use of IMs allows for the maximization of the productivity and reliability of the processes, but, in consequence, it must be kept in optimal conditions to ensure its continuous working without undesirable and unexpected

production breakdowns [8]. In fact, IMs remain the most used electromechanical equipment and they may be found in approximately 80% of all industrial applications [9,10]. Accordingly, its wide use leads to the production of between 40% and 60% of the world's energy consumption according to the International Energy Agency and other research works [4,5,11,12]. For that reason, it is mandatory to ensure the correct operation of the IMs and their associated kinematic chain through adequately scheduled preventive maintenance; otherwise, it could result in severe economic losses, equipment damage, or human injuries [10]. In summary, the analysis and study of these machines are currently active since there are some challenges that have not been previously addressed in reported works, i.e., the detection and identification of faults regardless of the operating condition. For that reason, the development of new approaches using novel and non-conventional features is still required, to complement, or not, the existing methodologies for increasing the performance in the diagnostic.

Regarding the reported methodologies in the literature for detecting and classifying faults on IMs, there are two general strategies that can be identified: (1) classical methodologies and (2) Data-Driven-based Approaches (DDA). Most of the classical methodologies apply conventional techniques over measured signals in the time, frequency, and time–frequency domains to visualize abnormal behaviors related to the occurrence of faults [13,14]; the most used techniques are the Fast Fourier Transform (FFT), the Discrete Wavelet Transform (DWT), and the Multiple Signal Classification (MUSIC), among others. Although these techniques have demonstrated their effectiveness, some challenges and limitations have not been completely addressed. Hence, limitations such as the strong dependency on the spectral content and the overlapping of some fault-related frequencies can increase the degree of difficulty in differentiating between different faulty conditions. Accordingly, some challenges that have to be faced are the resolution requirements of the spectral-based methods, the adequate selection of a mother function in DWT-based methods, the selection of adequate decomposition levels in empirical mode decomposition, or the expert knowledge required in advance due to the technique complexity, to mention a few [15]. On the other hand, the data-driven-based approaches are strategies for decision making based on the analysis and interpretation of data, and they take advantage of their powerful capabilities for processing high amounts of data and grouping them as clusters, such as the Principal Component Analysis (PCA), the Linear Discriminant Analysis (LDA), k-Nearest Neighbors (kNN), and Dictionary Learning (DicL), among others [4,16,17]. Nevertheless, the common drawbacks of these techniques comprise the need to use historical or high volumes of collected data, where a low volume of data could yield misleading labeling, the application of the techniques being limited to the data used for training, the propensity to overfitting, data can be biased, in many cases there exists a requirement of additional data pre-processing prior to apply the clustering mechanisms, and so on [18]. Some schemes under the DDA are the Shallow Machine Learning (SML) approaches, which combine, in general, the feature extraction, such as statistical features, frequency features, etc., with some classifiers, like Fuzzy Logic (FL), Artificial Neural Networks (ANN), Support Vectors Machine (SVM), and so on, for categorizing the faults [19,20]. The general weaknesses of these techniques are the processing of a high number of features, the estimation of redundant or correlated information in the feature calculation stage, the technique overfitting, the adequate selection of the mechanisms for feature extraction, the classifier configuration, and the optimal values of the techniques hyperparameters, to mention a few [21]. Additionally, a subset of the SML approaches is the Deep Learning (DL) techniques, which are schemes that take advantage of new reinforced structures based on neural networks such as Sparse Auto-encoder (SAE), Convolutional Neural Networks (CNN), Self-organizing Maps (SOM), Generative Adversarial Networks (GAN), Expandable Continuous Learning (ECL), and Transfer Learning (TL), among others [22–26]. These structures are capable of learning features, patterns, profiles, and clues with the aim of directly classifying the faults from the signals in the time, frequency, and time–frequency domains, or from features computed from any of these domains. Furthermore, recent research has led to some improvements;

for example, GAN, ECL, and TL have integrated intelligent fault diagnosis by adding data augmentation properties, extending the learning capability through elastic expansion mechanisms that progressively introduce new network branches, or combining the transfer learning to the DL techniques [24–26]. These intelligent approaches overcome limitations of DL such as the incapability of processing a low volume of data samples, fusing diverse information, augmenting the representation learning capability, anticipating the influence of variability in working conditions, and speeding up the training processes. However, the main disadvantages of the DL techniques are the expertise and knowledge required to define the structure of the network, the selection of their optimal hyperparameters to guarantee the best performance for the faults diagnostic, misleading diagnosis for multiple or combined faults with similar response behavior, non-linearities, non-stationary behavior, data quantity/quality/complexity, and noise addition [27]. In this sense, recent works have also directed towards the use of non-conventional feature extraction, for example, metrics related to the behavior of the signal energy, such as entropy.

Several works have been developed considering non-conventional indicators in applications related to the detection of abnormal conditions in industrial equipment, such as the entropy calculation [28–31]. For example, in [32], twelve entropy-based features were extracted from vibration signals in time, frequency, and time–frequency domains for detecting problems in bearings, and a Z-score was used as a threshold of the features for showing the fault. Under the same framework, a theoretical review was developed in [31], where the entropy measurements related to the statistical mechanics, the information theory, and the dynamical systems theory were analyzed to give some insights about its potential for detecting bearing faults. The research in [33] presents a methodology for diagnosing faults in bearings by decomposing the vibration signal through the Wavelet Packet Transform (WPT) and, at the same time, it computes the Teager energy entropy ratio of the resonant frequency sub-band. Likewise, in other work, an approach for monitoring the conditions of rolling bearings in wind turbines is developed [34]. Firstly, the vibration signal is decomposed through the Variational Mode Decomposition (VMD). Next, the statistical features are extracted, and a permutation entropy analysis is performed to analyze the time series complexity, then feature selection is carried out and, finally, a multi-class-SVM performs the classification task. In another example, six entropy-based features were extracted from vibration signals to automatically diagnose the bevel gear wear [35]. In that work, an analytic wavelet transform is applied to decompose the signal into sub-bands, and then the entropy features are extracted and fed to a Least-Squares Support Vector Machine (LS-SVM) classifier. Additionally, the Kruskal–Wallis test is carried out to generate statistically meaningful results. The work reported in [30] makes use of several entropy-based methods for fault classification in bearings, and they are compared to provide assessments and recommendations for future developments in new classification methodologies. On the other hand, the research described in [36] proposes an intelligent fault diagnosis through the Hierarchical Diversity Entropy (HDE) as a feature extraction tool and the Random Forest (RF) classifier for detecting early bearing faults from signals with high noise levels and weak fault symptoms. In another case, a methodology that proposes to combine a Multiscale Weighted Permutation Entropy (MWPE) with a Twin Extreme Learning Machine (TELM) for extracting enhanced features with non-stationary and non-linear characteristics is developed in [37]. In another field, the work in [38] presents a hybrid methodology between VMD, permutation entropy, and k-means clustering for selecting a low-resistance grounding fault line. Here, the VMD serves to reduce the influence of harmonic content and noise to facilitate feature extraction. From the previously analyzed works, it is important to mention that entropy features have some pros and cons that are worth mentioning; for example, as cons, the direct feature extraction from time domain signals could be ineffective depending on the noise level, and also the extraction of entropy features could be enhanced through a multiscale or multiscale weighted approach [37,38]. In contrast, the pros are that several novel methods to enhance classical fault detection can be developed using entropy features; for example, they can address systems with non-linearities and

non-stationary characteristics, even with non-linear time series complexity, but also they provide consistency, strong robustness, and high calculation efficiency [28,29,31,36–38]. Regarding the field of rotative machinery, most of the reported works have focused their efforts on studying the faults in bearing elements due to the frequency relation to the entropy calculation, but other type of faults related to the kinematic chain of the motor could be addressed, such as coupling misalignment, system unbalance, or gears wear in gearboxes. In summary, any non-conventional features, like those based in the entropy that helps to provide clues of the different conditions of a system, are still considered a field of interest and their validation through correlation tests could be a relevant aspect for the detection of faults in IMs and the associated kinematic chain.

Therefore, the contribution of this work lies in evaluating the use of entropy features to determine its discriminant properties for condition monitoring of faults that may unexpectedly occur in gearboxes, induction motors, and their related kinematic chain. To perform the entropy signal analysis, this work proposes the calculation of six entropy parameters such as spectral, permutation, sample, approximated, fuzzy, and Rényi entropy; these features are estimated from different available physical magnitudes, like vibration signals and stator current signatures, which are acquired during different experimental tests. Four different severities of uniform wear are tested in the gearbox and in the induction motor conditions like bearing defect, broken rotor bar, unbalance, and misalignment are evaluated, and each of the aforementioned conditions is iteratively tested in the gearbox and the induction motor under different operating conditions. The entropy features are extracted, and further processing is performed to assess the operating condition of the IM and the related kinematic chain. In a sense of a complement tool, the entropy features are analyzed to demonstrate which of them provide meaningful and discriminant information that could improve the efficiency of classical approaches; meaning that when entropy indicators are used, they allow a more efficient discrimination of failure conditions. Additionally, the validation of the features is essential, and verification mechanisms are also important; for this reason, the Kruskal–Wallis test and the Fisher discriminant score could in a reliable way enhance the proposed analysis. The obtained results show that entropy features have advantages for the detection of faults, compared to other methods, which can be used to improve the detection efficiency.

## 2. Theoretical Background

Entropy is a physical magnitude that has been widely studied in physics and thermodynamics. This magnitude is commonly associated with the degree of disorder in a system; however, a most formal definition derives from the second law of thermodynamics that establishes that every system evolves to its most probable configuration, as this is the one with the most microstates, and it coincides with the one with the greatest entropy. In this sense, entropy can be explained as a measure of the probability that a system reaches a specific final state based on the number of different configurations that lead to such a final state [35]. In other words, the more different configurations that lead to the same final state, the higher the probability of reaching that state and the greater the entropy is. This notion of entropy has been employed in the field of data analytics and signal processing to determine the possible causes of a signal obtaining a specific trend, allowing us to explain how two different phenomena are related to each other. The most used expression for quantifying the entropy of a signal or dataset was introduced by Claude E. Shannon in 1948 [31,39], and it is shown in Equation (1).

$$H(X) = - \sum_{i=1}^n p(x_i) \log(p(x_i)) \quad (1)$$

where  $H(X)$  is the Shannon entropy of the signal  $X$ ;  $x_i$  is the  $i$ -th element of the signal  $X$ ; and  $p(x_i)$  is the probability that the value  $x_i$  appears in the signal.

Although the Shannon entropy is good for detecting irregularities in a signal, some variants have been developed to increase the accuracy and performance of this feature. For instance, when it is necessary to estimate the complexity of the data, the spectral entropy (*SpecEn*) is a better feature. The spectral entropy needs to compute the spectrum of the signal  $X$  using the Fourier transform and calculate the power of each frequency component ( $p_f$ ); thus, the spectral entropy is finally obtained using Equation (2).

$$SpecEn(X) = \sum_f p_f \log\left(\frac{1}{p_f}\right) \tag{2}$$

Another variant of the Shannon entropy that works better in the quantification of uncertainty and randomness in time series is the Rényi entropy (*ReEn*). This is a generalization of the Shannon entropy that can be computed using Equation (3).

$$ReEn(X) = \frac{\alpha}{1 - \alpha} \sum_{i=1}^n \log_2(p(x_i)^\alpha) \tag{3}$$

where  $p(x_i)$  is the probability of the time series  $X = \{x_1, x_2, \dots, x_n\}$ , and  $\alpha \neq 1$  represents the order of the feature. Regarding the parameter  $\alpha$ , it is common to use values higher than 2 to obtain smoother results. Moreover, if  $\alpha = 1$ , then the Rényi entropy becomes the Shannon entropy.

In some applications, it is also required to know the irregularity and unpredictability of the time series, and the approximate entropy (*ApEn*) is the preferred tool for such a task. This feature can assess the probability that a new pattern appears in the signal behavior, and it can be expressed as shown in Equation (4).

$$ApEn(X) = \phi^m(r) - \phi^{m+1}(r) \tag{4}$$

where  $\phi^m(r)$  and  $\phi^{m+1}(r)$  are the mean values of the logarithm pattern and they are calculated according to Equation (5).

$$\phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln \left[ \frac{1}{N - m + 1} \text{num}\{d[x(i), x(j)] < r\} \right] \tag{5}$$

where  $m$  is the length of the pattern,  $r = 0.2SD$  (with  $SD$  as the standard deviation of the signal) is the tolerance of the signal,  $N$  is the length of the signal, and  $\text{num}\{d[x(i), x(j)] < r\}$  is the distance between the terms  $x(i)$  and  $x(j)$ . To achieve better results in the literature, it is recommended to use a value of  $m = 2$  for the pattern length [29]. The *ApEn* presents some immunity against interference and noise, but it is highly dependent on the signal length, and small signals usually deliver unexpected results. Additionally, the estimation regarding the complexity of the signal is not very accurate. Therefore, to overcome the issues related to the *ApEn*, it was developed the sample entropy (*SampEn*) that can be calculated using Equation (6).

$$SampEn = \ln\left(\frac{B^{m+1}(r)}{B^m(r)}\right) \tag{6}$$

where  $B^m(r)$  and  $B^{m+1}(r)$  are the mean values of the pattern count, and they can be obtained through Equation (7).

$$B^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \left[ \frac{1}{N - m} \text{num}\{d[x(i), x(j)] < r\} \right] \tag{7}$$

where  $r$  is the tolerance of the signal,  $m$  is the embedding dimension,  $N$  is the signal length, and  $\text{num}\{d[x(i), x(j)] < r\}$  is the distance between  $x(i)$  and  $x(j)$  with  $i = 1, 2, \dots, N - m + 1$ , and  $i \neq j$ .

Also, the fuzzy entropy (*FuzEn*) is an important feature that allows us to determine the similarity between two signals [28]. In this sense, the similarity can be expressed by Equation (8).

$$S_{ij}^m = e^{-\ln(2)\left(\frac{d_{ij}^m}{r}\right)^n} \quad (8)$$

where  $m$  is the embedding dimension,  $d_{ij}^m$  is the distance between the samples  $x_i^m$  and  $x_j^m$ , and  $r$  is the similarity tolerance. Also, the operator  $\varphi^m(N, r)$  must be defined as in Equation (9).

$$\varphi^m(N, r) = \frac{1}{N - m} \sum_{i=1}^{N-m} \frac{1}{N - m + 1} \sum_{j=1, j \neq i}^{N-m} S_{ij}^m \quad (9)$$

With  $N$  being the length of the signal, therefore, the *FuzEn* can be calculated as Equation (10) depicts.

$$\text{FuzEn}(m, r, N) = \ln \varphi^m(N, r) - \ln \varphi^{m+1}(N, r) \quad (10)$$

On its part, the permutation entropy (*PermEn*) also allows us to estimate the complexity of a signal, with the difference that it uses a phase space reconstruction to compute the complexity [30]. This is a particularly good approach for dealing with signals that introduce non-linearities. The permutation entropy is calculated using Equation (11).

$$\text{PermEn} = -\sum p(\pi) \log_2 p(\pi) \quad (11)$$

where  $p(\pi)$  is the relative frequency for each permutation.

Together, all these entropy variations can provide information regarding the anomalies and disturbances in a signal and how these variations are related to a specific operating condition in a system such as an induction motor.

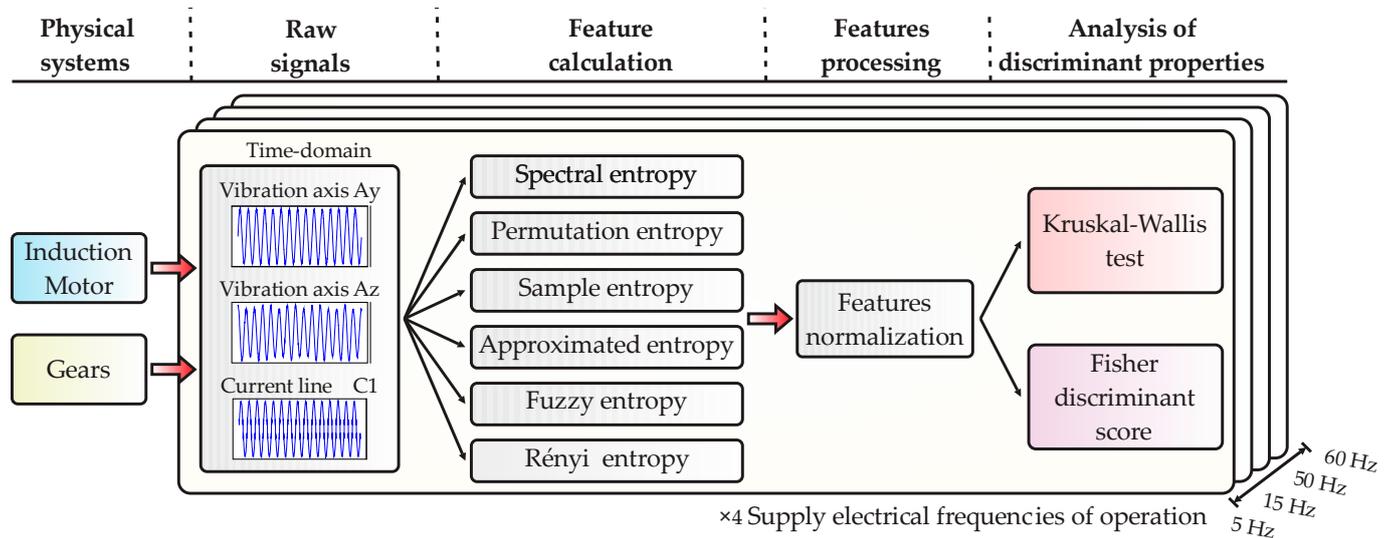
### 3. Methodology

As mentioned above, it is still a field of interest in the development of new diagnosis methodologies capable of distinguishing between different fault conditions that a machine could present. In this sense, Figure 1 presents the general flow chart of the approach proposed in this work, which is focused on an analysis of six different entropy features that are estimated from raw signals coming from data banks of an induction motor with fault conditions, with the main purpose of finding the discriminant properties for distinguishing the conditions of the motor (faults). Thus, the proposed approach is developed through five main stages: (i) physical systems, (ii) raw signals, (iii) feature calculation, (iv) feature processing, and (v) analysis of discriminant properties.

#### 3.1. Physical Systems

The entropy signal analysis developed in this work is considered to be a physical system that is an experimental test bench based on a kinematic chain composed of a three-phase IM, a 4:1 ratio gearbox (GB), and a DC generator acting as mechanical load (ML). In this kinematic chain, the three-phase motor speed is controlled by a variable frequency drive (VFD), and the experimental trials were run at four different supply electrical frequencies of operation: 5 Hz, 15 Hz, 50 Hz, and 60 Hz. Next, in the kinematic chain, the motor is mechanically coupled to the gearbox, which, in turn, is coupled to the DC generator. It must be highlighted that the DC generator represents a low mechanical load, entailing around 20% of the nominal load. The configuration of the low-level load is important because it can produce a downloading effect, which causes misleading fault detection, which is a challenge for the diagnostic in classic methodologies [40,41]. Thus, the entropy analysis is explored as an alternative in fault analysis under unfavorable conditions of low

load. On the other hand, the gearbox has a gear with uniform wear in all teeth, where four conditions are considered: wear of 0% (healthy gear and no wear in the teeth), wear of 25%, wear of 50%, and wear of 75%. Also, five different conditions are generated in the induction machine: healthy condition (motor without any fault), motor with bearing defect, a broken rotor bar, system unbalance, and coupling misalignment. It is important to mention that the corresponding experimental trials of every fault condition were individually carried out. The specific characteristics of the kinematic chain are explained in detail in the Section 4.



**Figure 1.** General flow chart of the proposed methodology for analyzing the discriminant properties of six different entropy features extracted from raw signals of induction motors with fault conditions.

### 3.2. Raw Signals

The proposed analysis considers the processing of real signals that are acquired from the experimental test bench described in the previous subsection. The experimentation is carried out individually for each system condition; as a result, the corresponding measurements consider a specific fault condition stored in a data bank for further offline processing. Now, in the kinematic chain, two sensors are considered to carry out the signal measurements, that is, a triaxial vibration sensor and a current sensor based on the Hall effect. Yet, although vibrations in the kinematic chain are measured in the  $x$ ,  $y$ , and  $z$  directions (axial, radial, and tangential axes), only the vibrations measured in the perpendicular plane (vibrations in the  $y$ -axis and the  $z$ -axis) of the kinematic chain rotating axis are considered in this work; such consideration is because it has been proven that this plane may experience an increase in vibrations when unexpected fault conditions appear in rotating machines [42,43]. On the other hand, only one stator current is considered to be analyzed because all the studied conditions are affected in a similar way to all the phases, hence, the stator current signature measured in the first line is taken into account. In the proposed analysis, the raw signals of the data bank are directly processed on its time-domain representation without the need for a space transform technique; the idea is to find out the usability of new features. The specific details about the sensor characteristics used in this work and the sampling frequencies considered for the data acquisition system are detailed in the Section 4.

### 3.3. Feature Calculation

Once the signals are measured and stored, the next step is to calculate a set of features inherent in the signals. In this case, six different features based on the entropy metric are addressed as counterparts of the classic methodologies that typically use statistical features. Thus, as observed in Figure 1, the entropy features obtained using expressions (1) to (11) are the following: *SpecEn*, *PermEn*, *SampEn*, *ApEn*, *FuzEn*, and *ReEn*. The feature calculation is carried out by segmenting each signal in equal parts of one second and then the proposed

entropy features are estimated from each segmented part. The main advantages of using entropy features rely on the fact that they describe energy behavior properties of a measured signal, instead of pure data distribution as is commonly performed by classical statistical features. Moreover, entropy features also allow us to address systems with non-linearities and non-stationary characteristics, even with non-linear time series complexity, but they also provide consistency, strong robustness, and high calculation efficiency, as stated in the Introduction.

### 3.4. Feature Processing

The feature processing considered for the entropy features calculated from the raw signals is the data normalization. The normalization that is considered is based on the well-known approach that leads to the transformation of a signal to zero mean and unit variance, and the main purpose of normalizing the features is due to the need to visualize the capabilities of entropy features in discriminating any condition in the data, like patterns, profiles, and tendencies, among others. Additionally, it is important to consider the conventional advantages of data normalization such as minimizing and correcting duplicated data, avoiding data anomalies, optimizing the storage space, avoiding the creation of undesired relationships and dependencies between data, reducing the time and complexity in data revision, facilitating the data access and interpretation, and preventing the undesired data erasing.

### 3.5. Analysis of Discriminant Properties

In the final stage of this proposal, the analysis of discriminant properties is carried out. Such analysis is performed with the aim of determining the most significant fault-related properties that are inherent to the calculation of entropy features. In this regard, the discriminant properties of the entropy features are then analyzed through the Kruskal–Wallis (KW) test and by calculating their Fisher discriminant score (FDS), which are a non-parametric and parametric technique, respectively, which can be used to identify useless sample features [44,45]. Thereby, the normalized entropy features are first analyzed using the non-parametric KW test, and this test leads to the determination of whether a group of data comes from the same population, since the probability distribution is not assumed. Certainly, under a null hypothesis, the implementation of the KW test identifies whether these data come from the same distribution and in condition monitoring strategies allows us to analyze if there are statistically relevant differences between two or more categories. Hence, for  $C$  classes with  $n$  samples (per class), the test is carried out as follows [35,46]:

- Step I: all samples ( $N$ ) for all  $C$  classes are sorted and then a rank is assigned in ascending order by ignoring the class to which the samples belong;
- Step II: find the sum of all  $R_i$  ranks for each individual class with  $n_i$  samples;
- Step III: compute the KW statistic ( $S_{KW}$ ) by applying Equation (12);
- Step IV: the significance of the resulting  $S_{KW}$  values for all  $C$  classes are assessed through the Chi-square test ( $\chi^2$ ); thus, the  $S_{KW}$  values are statistically significant if  $S_{KW}$  is equal or larger than  $\chi^2$ .

Once the KW test is performed, the  $p$ -value can be also estimated as a quantitative metric to determine if any of the differences between the medians are statistically significant; consequently, it has been determined that higher  $p$ -values are associated with statistically significant differences.

$$S_{KW} = \frac{12}{N(N+1)} \sum_{i=1}^C \frac{R_i^2}{n_i} - 3(N-1) \quad (12)$$

On the other hand, the FDS is a statistical technique carried out to determine the separability between individuals (samples) that belong to different populations (classes); FDS is used in condition monitoring strategies to analyze and choose features that provide significant discriminant information [47]. Ideally, the FDS is theoretically applied to two-class

problems but it has been generalized to multi-class problems; thereby, for  $C$  classes with  $n$  samples per class, the analysis through FDS considers the computation of the between-class ( $\overline{S}_b$ ) and within-class ( $\overline{S}_w$ ) scatter matrices following Equations (13) and (14) [48], where  $\mu_{f_i}$  and  $\mu$  are the mean of the  $k$ th class and the overall mean of all samples for all classes, respectively. On the other hand,  $x_{f_i}^k$  represents each  $i$ th sample of the  $k$ th class. Accordingly, the estimation of  $\overline{S}_b$  and  $\overline{S}_w$  is aiming to quantify the separability between classes and to measure the compactness of data points in the same class, respectively. Hence, the FDS is achieved as a ratio ( $FDS_r$ ) between  $\overline{S}_b$  and  $\overline{S}_w$ , which measures the linear separation between two different classes, as specified in (15) [49,50]. Therefore, values of  $FS_r$  higher than 1 are related to features that provide relevant discriminant information to separate the classes under analysis, whereas values of  $FS_r$  smaller than 1 can lead to overlapping problems when two classes are intended to be separated.

$$\overline{S}_b = \sum_{k=1}^C n_k (\mu_{f_i} - \mu) (\mu_{f_i} - \mu)^T \quad (13)$$

$$\overline{S}_w = \sum_{k=1}^C \sum_{i=1}^{n_k} (x_{f_i}^k - \mu_{f_i}) (x_{f_i}^k - \mu_{f_i})^T \quad (14)$$

$$FS_r = \frac{\overline{S}_b}{\overline{S}_w} \quad (15)$$

Finally, as it is observed in the block diagram of Figure 1, the proposed approach is repeated for each one of the operating frequencies that represent the supply electrical frequency of operation of the kinematic chain (5 Hz, 15 Hz, 50 Hz, and 60 Hz). Recently, the proposal of new methodologies applied to the diagnosis of faults may deal with different data distributions that are usually generated by different operating conditions like loads and speeds.

#### 4. Experimental Setup

The analysis of signal entropy performed in this work is applied to a real dataset that is acquired from an experimental laboratory test bench, which is composed of three main elements: an IM, a GB, and a ML. The test bench is shown in Figure 2, the IM is a 2.49 kW, 220 V-three-phase motor with one pair of poles (model WEG 00236ET3E145T-W22), and it is mechanically linked to a 4:1 ratio GB (BALDOR GCF4X01AA) through a rigid coupling; the GB is, in turn, coupled to a DC generator (BALDOR CDP3604), which is used as the ML. Also, a VFD (model WEG-CFW08) is used to feed and control the rotating speed of the IM output shaft.

The dataset comprised vibration signals and stator current signatures; in this regard, the vibrations of the whole kinematic chain are acquired using a triaxial accelerometer LIS3L02AS4 that is located on the top of the GB; meanwhile, the stator currents are measured through a set of three Hall-effect sensors from Tamura Corporation (model L08P050D15) that are placed in the power supply lines between the VDF and the IM. All signals are acquired by a self-designed Data Acquisition System (DAS), which is based on Field Programmable Gate Array (FPGA) technology, the proprietary DAS has a 12-bit analog-to-digital serial-output converter (model ADS 7841) to read the available measurements. Thus, the vibration signals and stator currents are acquired with sampling frequencies equal to 3 k samples/s and 4 k samples/s, respectively.

Regarding the evaluated conditions, four conditions are tested in the GB; meanwhile, five conditions are tested in the IM. Table 1 summarizes the details of each studied condition. Each one of the conditions under study is iteratively tested in the GB and the IM by replacing the healthy elements with the damaged ones; additionally, different supply electrical frequencies of operation are used in the VFD to analyze the behavior of the whole kinematic chain at different rotating speeds, and the supply frequencies are 5 Hz, 15 Hz, 50 Hz, and 60 Hz. The acquisition of vibrations and stator currents is carried out during

the 100 s in the steady-state regime of the IM; thus, a total of 300 k and 400 k samples are acquired and stored in a personal computer for further processing, respectively.

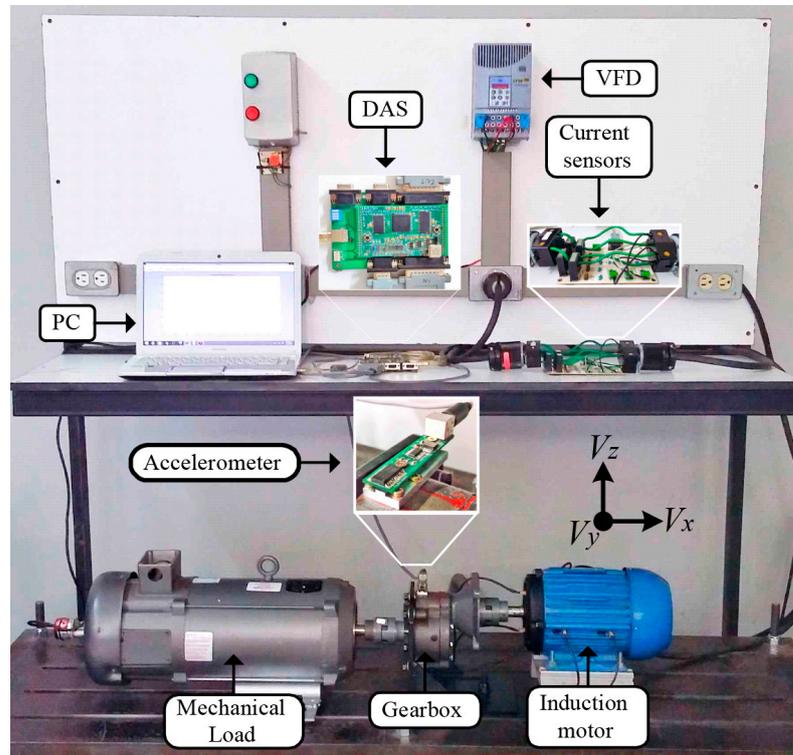


Figure 2. Experimental test bench used for the evaluation of entropy signals as a fault-related feature.

Table 1. Description of the different studied conditions tested in the GB and the IM.

Element under Evaluation	Studied Condition	Assigned Label	Description	Picture
Gearbox (GB)	GB with 0% of wear	GB0	The 4:1 GB is healthy, and the driven (72 teeth) and driver (18 teeth) gears are in perfect condition.	
	GB with 25% of wear	GB25	Uniform wear is artificially induced to the driven gear until all teeth are worn within 25% of the original condition.	
	GB with 50% of wear	GB50	Uniform wear is artificially induced to the driven gear until all teeth are worn within 50% of the original condition.	
	GB with 75% of wear	GB75	Uniform wear is artificially induced to the driven gear until all teeth are worn within 75% of the original condition.	

Table 1. Cont.

Element under Evaluation	Studied Condition	Assigned Label	Description	Picture
Induction motor (IM)	Healthy IM	HLT	The IM is healthy, and all the elements of the IM are in perfect condition.	
	Bearing defect	BD	The outer bearing race in a 6205 bearing is artificially generated by drilling a through-hole with a 1.191 mm diameter.	
	Broken rotor bar	BRB	A hole with 14 mm of depth is drilled in an IM rotor to completely break a rotor bar and to produce artificial damage.	
	Unbalance	UNB	An external mass (bolt) is attached to the rigid coupling that links the IM to the GB to produce an artificial unbalanced condition.	
	Misalignment	MIS	The free end of the IM is displaced to generate an angular misalignment of about 6 degrees.	

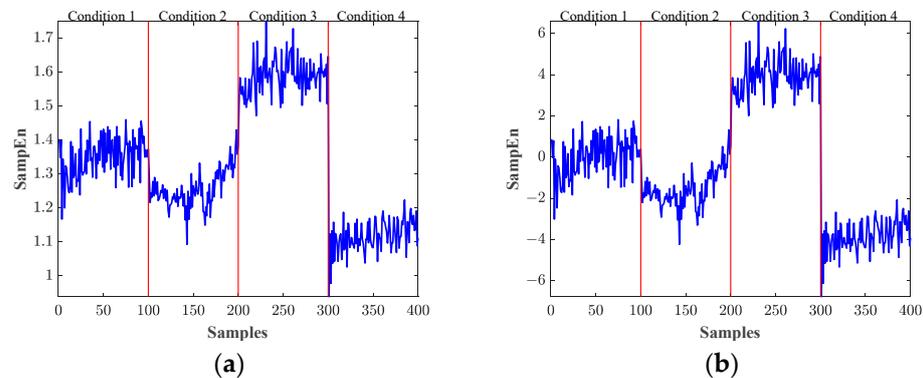
## 5. Results and Discussions

The goal of this work is to perform an analysis of signal entropy to identify changes and trends in entropy-related features produced by the occurrence of faults; in this sense, the proposed analysis is implemented under GNU Octave. Hence, the different faulty conditions, according to Table 1, are tested in a kinematic chain. Despite both elements, with GBs and IMs installed in the same kinematic chain, each one of the studied conditions is experimentally tested individually under the four different supply electrical frequencies of operation managed through the VFD.

### 5.1. Entropy Analysis for Detecting Faults in the GB

According to the explanation of the second stage of the methodology and the diagram observed in Figure 2, the two vibration signals subject to further processing are the  $V_y$  and the  $V_z$ , and the stator's current signature is  $C_1$ . Thus, the six entropy features previously mentioned are computed from each one of these signals according to the explanation of the third stage of the proposed methodology. Hence, as a result, three characteristic feature matrices are obtained, which are  $V_y$ ,  $V_z$ , and  $C_1$ ; each feature matrix has a dimension of  $6 \times 100$  where the columns represent the six entropy features and the rows belong to each one of the segmented parts of  $V_y$ ,  $V_z$ , and  $C_1$ , respectively. In addition, the characteristic features matrices are obtained to generate a consecutive set of samples (100 samples) for each one of the entropy features; accordingly, due to the four different supply electrical frequencies of operation used to carry out the experimentation for each condition twelve characteristic feature matrices are computed; for example, for the gearbox in healthy condition (GB0), the matrices obtained are  $V_y^{GB0@5Hz}$ ,  $V_z^{GB0@5Hz}$ ,  $C_1^{GB0@5Hz}$ ,  $V_y^{GB0@15Hz}$ ,  $V_z^{GB0@15Hz}$ ,  $C_1^{GB0@15Hz}$ ,

$V_y^{GB0@50Hz}$ ,  $V_z^{GB0@50Hz}$ ,  $C_1^{GB0@50Hz}$ ,  $V_y^{GB0@60Hz}$ ,  $V_z^{GB0@60Hz}$ , and  $C_1^{GB0@60Hz}$ . Once all the feature matrices are computed for all the considered conditions, the feature normalization is carried out by normalizing each entropy feature to zero mean and unit variance; this procedure plays a key role in condition monitoring strategies since it leads to improving the model performance and allows to ensure convergence. Figure 3a shows the qualitative visual representation of the *SampEn* feature without normalization that is estimated directly from the  $V_z$  signal for all conditions of uniform wear (GB0, GB25, GB50, and GB75). On the other hand, Figure 3b also shows the *SampEn* feature after applying the normalization procedure. As observed, the amplitude scale of the vertical axis in Figure 3b is significantly higher than the amplitude scale of the vertical axis in Figure 3a. This particular difference between amplitude scales in vertical axes may produce a significant increase in terms of accuracy and performance when normalized features are used to produce the fault assessment. The normalization procedure is applied to all feature matrices for all studied conditions and normalized feature matrices are then obtained ( $NV_y$ ,  $NV_z$ , and  $NC_1$ ); for example, for the GB0 condition, the normalized matrices obtained are  $NV_y^{GB0@5Hz}$ ,  $NV_z^{GB0@5Hz}$ , and  $NC_1^{GB0@5Hz}$ , when the kinematic chain operates at 5 Hz for the signals  $V_y$ ,  $V_z$ , and  $C_1$ , respectively.



**Figure 3.** Qualitative representation of the *SampEn* feature estimated from the  $V_z$  signal for all evaluated conditions in the gearbox (GB0, GB25, GB50, and GB75) when the kinematic chain operates at 60 Hz: (a) *SampEn* without normalization and (b) *SampEn* normalized to mean zero and unit variance.

Afterward, the normalized entropy features are then analyzed using two different approaches with the aim of determining their capabilities for being included in diagnosis strategies applied to the detection and identification of faults in kinematic chains. Thus, the KW test and the FDS are applied over the previously normalized features; hence, the KW test is first applied to each entropy feature for all studied conditions in the GB. Thereby, Equation (12) is used to carry out the KW test, for the conditions tested in the GB the number of classes is defined as  $C = 4$ , with a total number of samples  $N = 400$  for all classes with  $n_i = 100$  samples per class.

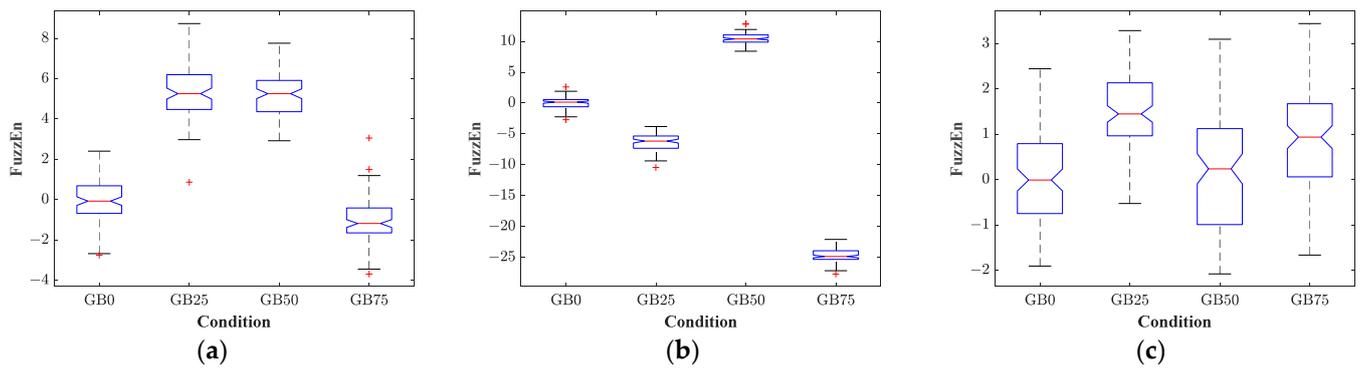
Specifically, for each available signal and each operating condition, the KW test is individually applied to each entropy feature of the normalized feature matrices  $NV_y$ ,  $NV_z$ , and  $NC_1$ , i.e., the KW test leads to achieving twelve numerical  $p$ -values for only one entropy feature for the signals  $V_y$ ,  $V_z$ , and  $C_1$  for all GB conditions. This test is carried out to determine whether a group of data points (samples) comes from the same population. In this regard, the quantitative metric known as  $p$ -value may depict the level of distinction between the assessed conditions (classes); in particular,  $p$ -values smaller than 0.05 mean that there exist statistically significant differences between analyzed classes. Consequently, Table 2 summarizes the  $p$ -values achieved during the KW test applied to the corresponding normalized entropy features estimated from signals  $V_y$ ,  $V_z$ , and  $C_1$  for the conditions GB0, GB25, GB50, and GB75 that are tested in the GB for all the supply electrical frequencies of operation 5 Hz, 15 Hz, 50 Hz, and 60 Hz. As appreciated in Table 2, most of the entropy features for the signals  $V_y$ ,  $V_z$ , and  $C_1$  produce  $p$ -values smaller than 0.05 demonstrating that

these features have significant information to differentiate between the analyzed conditions. Specifically, most of the  $p$ -values obtained during the test are represented on a scale that varies from  $1 \times 10^{-13}$  to  $1 \times 10^{-81}$ , which theoretically can be taken into account as zero.

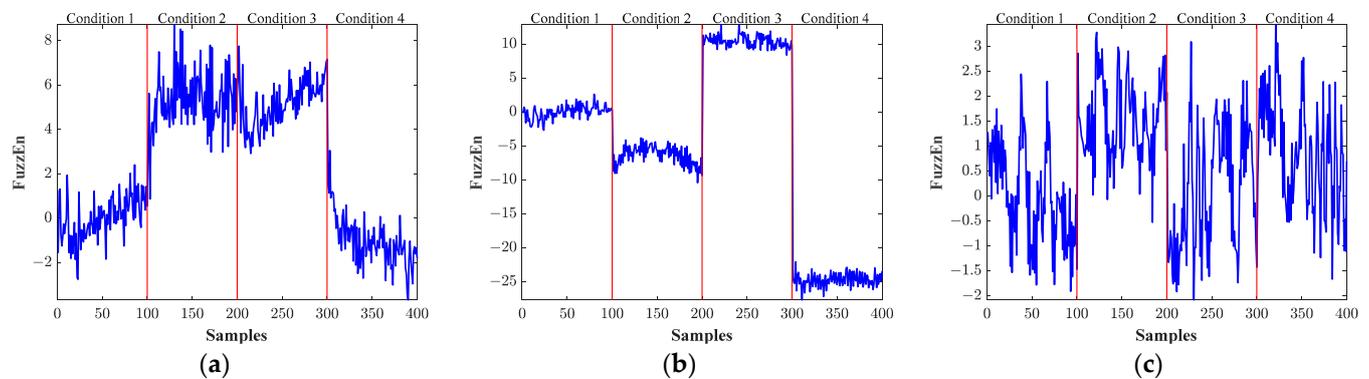
**Table 2.** Quantitative  $p$ -value achieved during the KW test applied to the entropy features estimated from signals  $V_y$ ,  $V_z$ , and  $C_1$  for the conditions GB0, GB25, GB50, and GB75 tested in the GB.

Entropy Features	Physical Magnitude	Supply Electrical Frequencies of Operation			
		5 Hz	15 Hz	50 Hz	60 Hz
<i>SpecEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>PermEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>SampEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>ApEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>FuzzEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>ReEn</i>	$V_y$	0.00536	0.0716	0.0462	0.00522
	$V_z$	0.441	0.00230	0.0269	0.0218
	$C$	0.318	0.448	0.217	0.664

The KW test also allows to obtain boxplots with the aim of comparing the data distribution of the classes under analysis; Figure 4a–c show the boxplots for the *FuzzEn* feature computed from the signals  $V_y$ ,  $V_z$ , and  $C_1$  for all conditions tested in the GB when the kinematic chain operates at 60 Hz, respectively. The corresponding  $p$ -values computed for these boxplots are  $7.10 \times 10^{-67}$ ,  $9.17 \times 10^{-81}$ , and  $1.92 \times 10^{-22}$ , respectively, and the labels in the horizontal axes directly indicate the conditions tested in the GB. On the other side, the horizontal red line inside the boxes represents the median of the tested conditions, whereas the red symbols '+' are samples detected as outliers. Accordingly, based on the KW test, it can be assumed that the *FuzzEn* feature provides statistically significant information to discriminate between assessed conditions in the GB, and this assertion is supported by the achieved  $p$ -values, which are smaller than 0.05. In fact, the *FuzzEn* estimated from the signal  $V_z$  is the most representative because the medians of all classes (horizontal red lines inside the boxes) are vertically separated from each other, avoiding the overlapping between the considered conditions. In its counterpart, the *FuzzEn* estimated from signals  $V_y$  and  $C_1$  also produce  $p$ -values smaller than 0.05, but their corresponding boxplots show a slight overlapping; that is, in Figure 4a, medians of conditions GB25 and GB50 appear horizontally aligned and in Figure 4c there is not a considerable separation between the medians of all conditions. Finally, the quantitative representation of the *FuzzEn* feature estimated from signals  $V_y$ ,  $V_z$ , and  $C_1$  when the kinematic chain operates at 60 Hz is shown in Figure 5a–c, respectively. These quantitative values belong to the normalized *FuzzEn* for the four conditions (100 samples per condition). As appreciated, there exists a clear difference between the values of *FuzzEn* in Figure 5b for each assessed condition, whereas the values of *FuzzEn* in Figure 5a only present different amplitudes for the GB0, GB25, and GB50 conditions, and the values of *FuzzEn* in Figure 5c do not show considerable changes between all conditions.



**Figure 4.** Boxplots achieved for the *FuzzEn* feature for the GB0, GB25, GB50, and GB75 conditions tested in the GB when the kinematic chain operates at 60 Hz; the boxplots for *FuzzEn* belong to the signals: (a)  $V_y$ , (b)  $V_z$ , and (c)  $C_1$ .



**Figure 5.** Quantitative representation of the estimated fuzzy entropy when evaluating the conditions of 0%, 25%, 50%, and 75% of uniform wear in the GB with an operating frequency of 60 Hz, the fuzzy entropy is estimated for (a)  $V_y$ , (b)  $V_z$ , and (c)  $C_1$ .

Following the proposed method, the entropy features are also analyzed through the FDS approach to determine their discriminant properties for distinguishing between classes. Hence, for each one of the three signals and each supply electrical frequency of operation in the VFD, the analysis based on FDS is individually carried out over each entropy feature as follows:

- (i) An entropy feature from the normalized feature matrices  $NV_y$ ,  $NV_z$ , and  $NC_1$  is selected, such as, for example, the *FuzzEn*.
- (ii) When analyzing the studied conditions in the GB, firstly, the samples of the *FuzzEn* feature from the GB0 condition are compared to the samples of the *FuzzEn* feature from the GB25 condition by using (13)–(15). As a result, a discriminant ratio ( $FS_r$ ) is computed, which measures the linear separation between the GB0 and GB25 conditions.
- (iii) The previous step is repeated for the same entropy feature (i.e., *FuzzEn*) until all faulty conditions (GB25, GB50, and GB75) are iteratively faced with the healthy condition (GB0).
- (iv) Another entropy feature is selected and steps (II) and (III) are performed; over the new entropy feature, the procedure is applied to all available entropy features.

Thereby, Table 3 summarizes the  $FS_r$  scores achieved by applying the FDS analysis to each normalized entropy feature estimated from the signals  $V_y$ ,  $V_z$  and  $C_1$  for the operating supply frequencies 5 Hz and 15 Hz, whereas Table 4 summarizes the  $FS_r$  scores for the operating supply frequencies 50 Hz and 60 Hz. Theoretically, values of  $FS_r$  higher than 1 are associated with features that provide significant discriminant information to separate the classes under analysis; that is, entropy features that produce  $FS_r > 1$  ensures the separability of classes. Thus, for the normalized values of *FuzzEn* features presented in

Figure 5a, the  $FS_r$  scores are 26.65, 41.89, and 0.06 when facing GB0 vs. GB25, GB0 vs. GB50, and GB0 vs. GB75, respectively; these  $FS_r$  scores mean that the  $FuzzEn$  estimated from  $V_y$  signal has the capability to distinguish the GB0 condition from GB25 and GB50 conditions, but not from the GB75 condition. On the other hand, for the normalized  $FuzzEn$  feature of Figure 5b, the computed  $FS_r$  scores are 46.41, 971.21, and 23,643.02; these values describe a high capability of the  $FuzzEn$  feature estimated from the  $V_z$  signal for separating any of the faulty conditions (GB25, GB50, and GB75) from the healthy one (GB0). Likewise, the  $FS_r$  scores for the normalized  $FuzzEn$  feature, estimated from the  $C_1$  signal, and shown in Figure 5c are 0.52, 0.00, and 0.03, respectively. Hence, these values depict a lack of meaningful information to separate the healthy condition from the faulty ones.

**Table 3.** Quantitative Fisher score values achieved for the evaluation of the gearbox conditions (5 Hz and 15 Hz).

		Supply Electrical Frequencies of Operation					
		5 Hz			15 Hz		
Entropy Features	Physical Magnitude	0% vs. 25%	0% vs. 50%	0% vs. 75%	0% vs. 25%	0% vs. 50%	0% vs. 75%
<i>SpecEn</i>	$V_y$	0.13	513.29	0.05	1.08	0.68	0.68
	$V_z$	0.00	20.07	0.00	13.05	349.14	0.12
	$C$	0.21	213.38	115.88	0.01	0.00	0.00
<i>PermEn</i>	$V_y$	0.00	0.00	0.18	0.03	0.38	0.02
	$V_z$	0.00	13.67	0.33	0.04	113.63	0.45
	$C$	0.01	0.22	0.00	0.13	0.04	0.01
<i>SampEn</i>	$V_y$	0.00	0.10	1.20	2.41	0.00	2.26
	$V_z$	0.00	0.08	0.00	0.03	88.19	0.31
	$C$	0.02	0.01	0.04	0.00	0.00	0.24
<i>ApEn</i>	$V_y$	0.61	13.71	0.23	0.01	0.20	16.00
	$V_z$	0.01	6.80	0.14	0.03	190.27	0.10
	$C$	0.02	0.01	0.08	0.00	0.00	0.17
<i>FuzzEn</i>	$V_y$	21.05	1480.97	31.14	1.50	20,259.24	381.99
	$V_z$	0.01	0.91	10.34	856.53	12,493.92	3160.64
	$C$	0.02	0.07	0.00	0.02	0.10	12.86
<i>ReEn</i>	$V_y$	0.00	0.00	0.00	0.00	0.00	0.00
	$V_z$	0.00	0.00	0.00	0.00	0.00	0.00
	$C$	0.00	0.00	0.00	0.00	0.00	0.00

**Table 4.** Quantitative Fisher score values achieved for the evaluation of the gearbox conditions (50 Hz and 60 Hz).

		Supply Electrical Frequencies of Operation					
		50 Hz			60 Hz		
Entropy Features	Physical Magnitude	0% vs. 25%	0% vs. 50%	0% vs. 75%	0% vs. 25%	0% vs. 50%	0% vs. 75%
<i>SpecEn</i>	$V_y$	113.05	73.85	12.61	0.07	5.82	4.66
	$V_z$	72.75	187.76	0.79	0.00	2.04	30.72
	$C$	20,062.81	5.38	0.22	3460.57	246.86	3.67
<i>PermEn</i>	$V_y$	173.48	6.91	481.62	6.42	0.61	0.71
	$V_z$	11.03	18.68	3.20	0.00	0.00	8.24
	$C$	1096.29	1.73	0.04	466.80	63.29	0.29
<i>SampEn</i>	$V_y$	0.53	0.00	1.78	0.13	0.43	0.03
	$V_z$	7.25	10.61	44.39	0.42	19.83	21.16
	$C$	42.42	0.00	0.00	21.13	0.57	0.00

Table 4. Cont.

		Supply Electrical Frequencies of Operation					
		50 Hz			60 Hz		
Entropy Features	Physical Magnitude	0% vs. 25%	0% vs. 50%	0% vs. 75%	0% vs. 25%	0% vs. 50%	0% vs. 75%
<i>ApEn</i>	$V_y$	22.30	2.44	4.83	30.91	29.19	1.87
	$V_z$	131.91	0.32	1844.60	1.17	116.48	384.40
	$C$	46.03	0.00	0.00	43.97	1.05	0.00
<i>FuzzEn</i>	$V_y$	0.63	1.68	1.07	26.65	41.89	0.06
	$V_z$	33,901.59	1138.96	66,594.98	46.41	971.21	23,643.02
	$C$	0.02	3.07	196.82	0.52	0.00	0.03
<i>ReEn</i>	$V_y$	0.00	0.00	0.00	0.00	0.00	0.00
	$V_z$	0.00	0.00	0.00	0.00	0.00	0.00
	$C$	0.00	0.00	0.00	0.00	0.00	0.00

In addition, to complement the proposed signal entropy analysis for the assessment of uniform wear in the GB, Figure 6a–c present the boxplots for the normalized *SpecEn* feature estimated from the signals  $V_y$ ,  $V_z$ , and  $C_1$ . The KW test applied to these normalized values leads to obtaining  $p$ -values equal to  $4.64 \times 10^{-71}$ ,  $2.52 \times 10^{-69}$ , and  $5.01 \times 10^{-80}$  for each corresponding signal when the kinematic chain operates at 60 Hz. Likewise, the most relevant  $FS_r$  scores are computed for the  $C_1$  signal with values about 3460.57, 246.86, and 3.67 when facing GB0 vs. GB25, GB0 vs. GB50, and GB0 vs. GB75, correspondingly. Also, another important  $FS_r$  score is calculated for the  $V_z$  signal when comparing GB0 vs. GB75 conditions, and such an  $FS_r$  score is equal to 30.72. Therefore, it should be highlighted that there exists a relationship between higher values in the  $FS_r$  score and the vertical separation of medians in the boxplots. On the other side, it is also probable that a specific feature computed from a particular signal does not provide sufficient information for distinguishing between different conditions, but another feature may provide enough information to carry out the class separation. In this regard, Figure 7a–c present the boxplots obtained for the normalized *ApEn* feature estimated from the signals  $V_y$ ,  $V_z$ , and  $C_1$  when the kinematic chain operates at 60 Hz, separately; their associated  $p$ -values are  $6.74 \times 10^{-79}$ ,  $1.74 \times 10^{-78}$ , and  $3.10 \times 10^{-66}$ . As well, the computed  $FS_r$  scores for the  $V_y$  signal are 30.91, 1.17, and 43.97, and, for the  $V_z$  signal they are 29.19, 116.48, and 1.05. In the case of the  $C_1$  signal, the corresponding scores are 1.87, 384.40, and 0.00 when the healthy condition is faced with each faulty condition. It is worth mentioning that the *ApEn* feature provides meaningful information for assessing different severities of uniform wear in the GB.

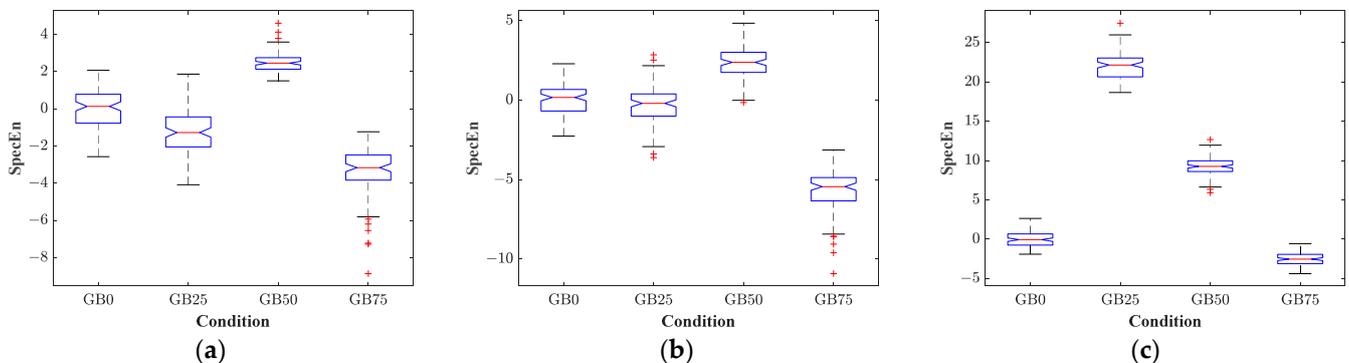
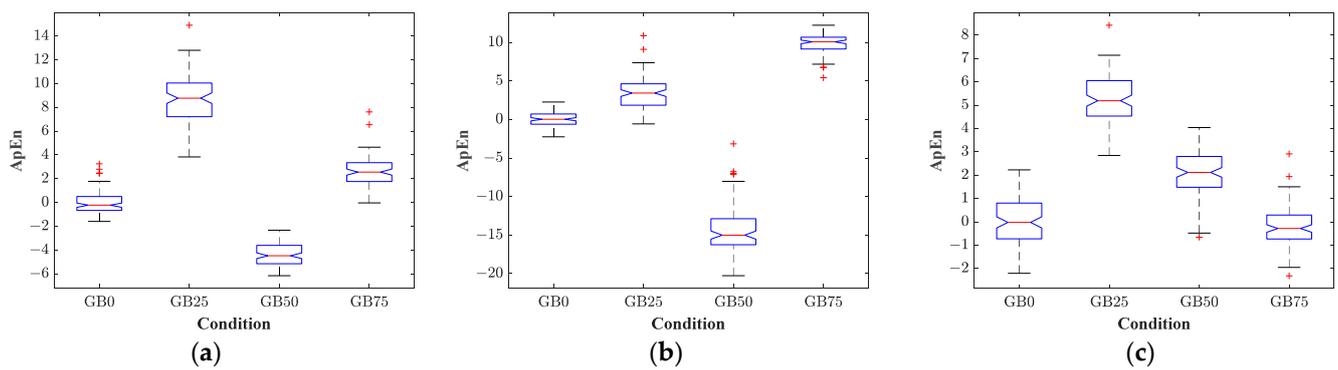


Figure 6. Boxplots of the estimated *SpecEn* feature when evaluating the GB0, GB25, GB50, and GB75 are tested in the GB with 60 Hz as operating frequency, the *SpecEn* is computed from signals: (a)  $V_y$ , (b)  $V_z$ , and (c)  $C_1$ .



**Figure 7.** Boxplots of the estimated *ApEn* feature when evaluating the GB0, GB25, GB50, and GB75 are tested in the GB and the kinematic chain operates at 60 Hz; the *ApEn* is computed from signals (a)  $V_y$ , (b)  $V_z$ , and (c)  $C_1$ .

Additionally, during the entropy signal analysis for detecting uniform wear in a GB, it can be noticed that the consideration of a single feature may not contain enough discriminant information to separate all the assessed conditions; in this regard, the combination of different entropy features may increase the capabilities of condition monitoring strategies to perform the fault assessment. Hence, to validate this assertion, the combination of three entropy features is tested and evaluated under the FDS analysis; that is, due to each tested condition characterized by 18 entropy features (6 entropy features per signal,  $V_y$ ,  $V_z$ , and  $C_1$  as summarized in the Table 5), there are  $C_3^{18} = 816$  possible combinations to be analyzed using Equations (13)–(15). Also, during this analysis, the GB0 condition is individually faced against the GB25, GB50, and GB75 conditions for each operating condition of the kinematic chain (5 Hz, 15 Hz, 50 Hz, and 60 Hz), the resulting  $FS_r$  scores are ranked in descending order and the first three ranked  $FS_r$  values are considered the most relevant. Table 6 summarizes the first three ranked  $FS_r$  values: as observed, any of the resulting combinations produce  $FS_r > 11$ , which ensures the separability of classes; the smallest  $FS_r$  scores are computed when GB0 is faced with GB25 when the kinematic chain operates at 5 Hz. Meanwhile, the highest  $FS_r$  scores are computed also when GB0 is faced with GB25 when the kinematic chain operates at 50 Hz.

**Table 5.** Identifiers assigned to each considered entropy feature for each signal under analysis  $V_y$ ,  $V_z$ , and  $C_1$ .

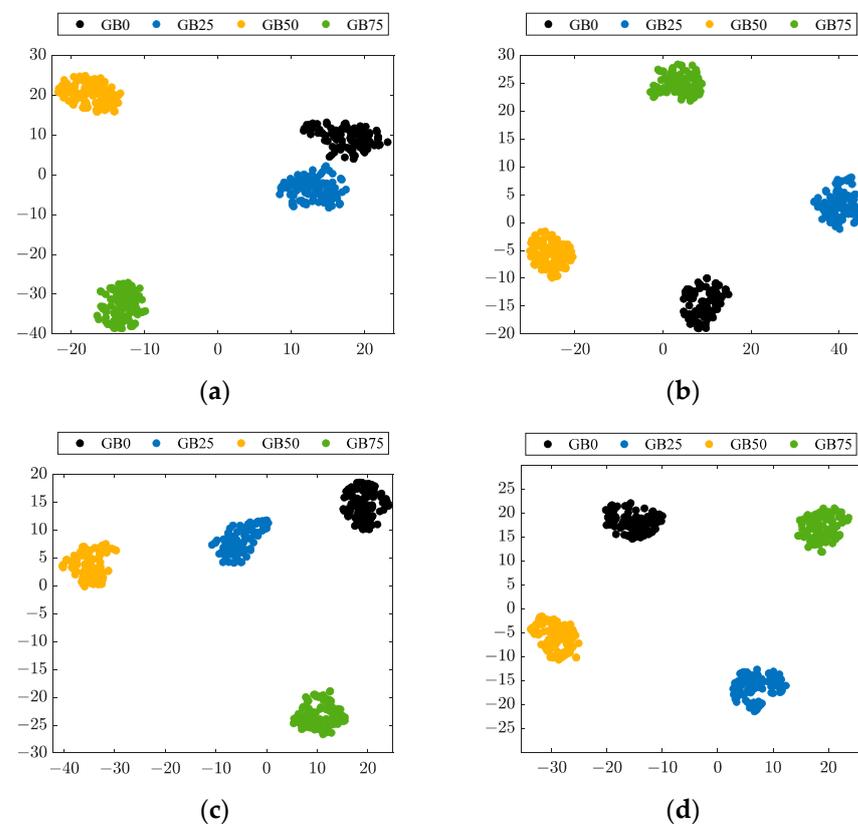
Entropy Feature	Signals under Analysis		
	$V_y$	$V_z$	$C_1$
<i>SpecEn</i>	1	7	13
<i>PermEn</i>	2	8	14
<i>SampEn</i>	3	9	15
<i>ApEn</i>	4	10	16
<i>FuzzEn</i>	5	11	17
<i>RenyiEn</i>	6	12	18

In order to validate that the combination of entropy features may improve the capability of condition monitoring strategies to perform the fault assessment, the *t*-distributed stochastic neighbor embedding (t-SNE) is used to obtain the visual representation of the best-ranked features (last row in Table 6) into a two-dimensional (2D) space. As a result, Figure 8a, Figure 8b, Figure 8c, and Figure 8d show the 2D visual representation of the data distribution for the best-ranked features resulting from evaluating all possible combinations ( $C_3^{18} = 816$ ) by the FDS when the kinematic chain operates at 5 Hz, 15 Hz, 50 Hz, and

60 Hz, respectively. It should be highlighted that all conditions of uniform wear (GB0, GB25, GB50, and GB75) tested in the GB appear separated from each other for any of the supply electrical frequencies of operation; thus, the use of entropy features estimated from vibration signals and stator current signatures can lead to the accurate assessment of faults in rotating machines like GBs.

**Table 6.** Quantitative Fisher score values obtained when combinations of three entropy features are considered during the evaluation of the gearbox conditions at all operating frequencies.

		Supply Electrical Frequencies of Operation								
		5 Hz		15 Hz		50 Hz		60 Hz		
		$FS_r$	Combined Features	$FS_r$	Combined Features	$FS_r$	Combined Features	$FS_r$	Combined Features	
Faced conditions	GB0 vs. GB25	13	5, 13, 17	363	5, 7, 11	39,594	11, 13, 14	3119	2, 13, 14	
		12	5, 13, 14	341	7, 8, 11	37,972	11, 13, 16	3115	12, 13, 14	
		11	5, 13, 16	340	4, 7, 11	37,176	11, 13, 15	3003	2, 12, 13	
	GB0 vs. GB50	1509	1, 5, 13	27,166	5, 11, 12	1037	1, 7, 11	777	11, 12, 13	
		1339	1, 5, 8	26,691	5, 6, 11	862	7, 11, 14	771	1, 11, 13	
		1339	1, 5, 17	26,618	4, 5, 11	862	1, 8, 11	769	2, 11, 13	
	GB0 vs. GB75	152	5, 11, 13	2975	4, 5, 11	22,278	1, 10, 11	11,637	9, 11, 12	
		142	5, 13, 17	2891	5, 11, 15	22,207	10, 11, 12	11,514	11, 12, 13	
		142	5, 13, 16	2882	5, 11, 16	20,742	1, 11, 12	10,910	11, 12, 15	
	Best-ranked features		4, 5, 6, 7, 8, 11, 12, 15, 16		1, 5, 8, 11, 13, 14, 16, 17		1, 7, 8, 10, 11, 12, 13, 4, 15, 16		1, 2, 9, 11, 12, 13, 14, 15	

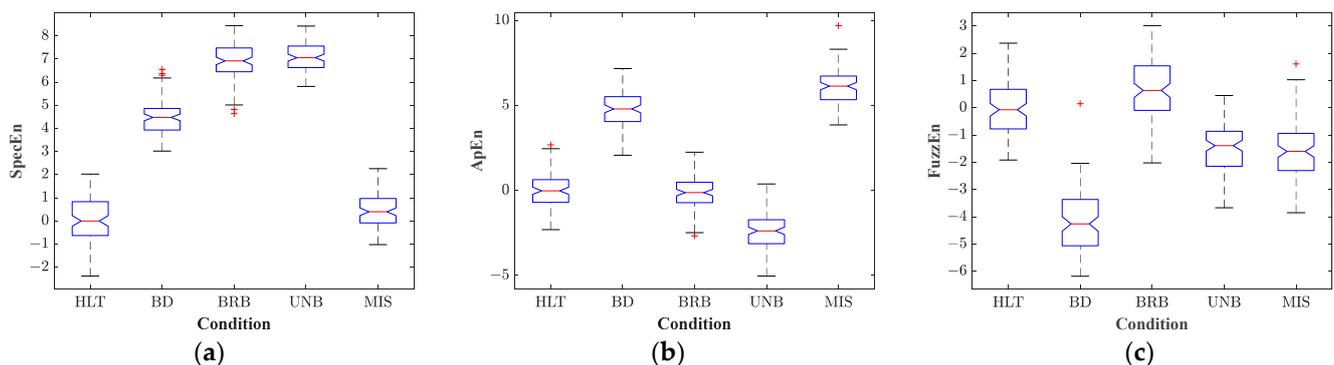


**Figure 8.** Resulting projection achieved by applying the t-distributed stochastic neighbor embedding (t-SNE) to the ranked entropy features when the GB0 condition is faced with GB25, GB50 and GB75 for the kinematic chain operating at (a) 5 Hz, (b) 15 Hz, (c) 50 Hz and (d) 60 Hz.

### 5.2. Entropy Analysis for Detecting Faults in the IM

Successively, as mentioned in the Section 3, the proposed entropy signal analysis is also applied to another dataset, which is also composed of vibration signals and stator current signatures, but, acquired during the evaluation of different faulty conditions such as HLT, BD, BRB, UNB, and MIS in an IM linked to a kinematic chain that operates at four different supply electrical frequencies of operation such as 5 Hz, 15 Hz, 50 Hz, and 60 Hz. The entropy analysis is applied to the same three signals  $V_y$ ,  $V_z$ , and  $C_1$ ; thus, these signals are equally segmented, and then the proposed set of entropy features from each segmented part is estimated. As a result, similarly to the analysis in the GB, three characteristic feature matrices are obtained,  $V_y$ ,  $V_z$ , and  $C_1$ , with a dimension of  $6 \times 100$ . Now, in this case, for the IM under the BRB condition, the following twelve matrices— $V_y^{\text{BRB}@5\text{Hz}}$ ,  $V_z^{\text{BRB}@5\text{Hz}}$ ,  $C_1^{\text{BRB}@5\text{Hz}}$ ,  $V_y^{\text{BRB}@15\text{Hz}}$ ,  $V_z^{\text{BRB}@15\text{Hz}}$ ,  $C_1^{\text{BRB}@15\text{Hz}}$ ,  $V_y^{\text{BRB}@50\text{Hz}}$ ,  $V_z^{\text{BRB}@50\text{Hz}}$ ,  $C_1^{\text{BRB}@50\text{Hz}}$ ,  $V_y^{\text{BRB}@60\text{Hz}}$ ,  $V_z^{\text{BRB}@60\text{Hz}}$ , and  $C_1^{\text{BRB}@60\text{Hz}}$ —are obtained for all frequencies. Afterward, the normalization procedure is carried out and the normalized feature matrices are generated,  $NV_y$ ,  $NV_z$  and  $NC_1$ , where entropy features have zero mean and unit variance, i.e., for the BRB condition, the matrices  $NV_y^{\text{BRB}@5\text{Hz}}$ ,  $NV_z^{\text{BRB}@5\text{Hz}}$ , and  $NC_1^{\text{BRB}@5\text{Hz}}$  are obtained when the kinematic chain operates at 5 Hz.

Once the normalization of all the feature matrices is performed, the KW test is carried out using Equation (12); thereby, for the conditions tested in the IM, the number of classes is defined as  $C = 5$ , with a total number of samples  $N = 500$ , for all classes with  $n_i = 100$  samples per class. Table 7 summarizes the resulting  $p$ -values; as observed, most of the  $p$ -values are smaller than 0.05, which describes statistically significant differences between the analyzed classes. In fact, most of the achieved  $p$ -values are represented on a scale that varies from  $1 \times 10^{-9}$  to  $1 \times 10^{-97}$ ; ideally, these values can be understood as zero. On the other hand, Figure 9a–c show the resulting boxplots achieved for the *SpecEn* feature estimated from  $V_y$ , the *ApEn* feature estimated from  $V_z$ , and the *FuzzEn* feature estimated from  $C_1$ , respectively, when the kinematic chain operates at 50 Hz. As appreciated, each one of the entropy features provides statistically significant information to separate specific conditions; that is, *SpecEn* can separate the HLT condition from BD, BRB, and UNB, but produce a slight overlapping with the MIS condition. Meanwhile, the *ApEn* has the capability of separating the HLT condition from BD, UNB, and MIS but not from the BRB condition. At last, the *FuzzEn* also separates the HLT condition from BD, UNB, and MIS; meanwhile, a slight overlapping is produced with the BRB condition.



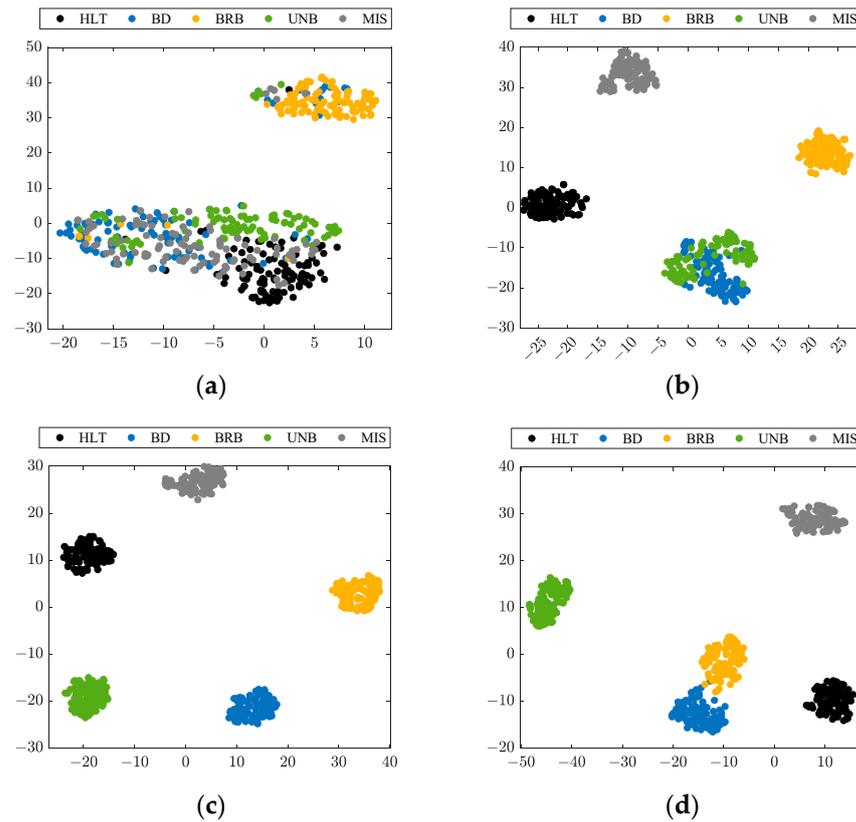
**Figure 9.** Boxplots obtained when evaluating the HLT, BD, BRB, UNB, and MIS conditions are evaluated in the IM and the kinematic chain operates at 50 Hz; the boxplots belong to (a) *SpecEn* feature estimated from  $V_y$ , (b) *ApEn* feature estimated from  $V_z$ , and (c) *FuzzEn* feature estimated from  $C_1$ .

**Table 7.** Quantitative  $p$ -value accomplished by applying the KW test to the entropy features estimated from signals  $V_y$ ,  $V_z$ , and  $C_1$  for the conditions BD, BRB, UNB, and MIS, evaluated in the IM for all operating frequencies.

Entropy Features	Physical Magnitude	Supply Electrical Frequencies of Operation			
		5 Hz	15 Hz	50 Hz	60 Hz
<i>SpecEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>PermEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>SampEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>ApEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>FuzzEn</i>	$V_y$	0	0	0	0
	$V_z$	0	0	0	0
	$C$	0	0	0	0
<i>ReEn</i>	$V_y$	0.762	0.452	0	0
	$V_z$	0.00588	0.0931	0.00877	0.0799
	$C$	0.199	0.395	0.227	0.891

Successively, the FDS analysis is performed by using Equations (13)–(15), with the aim of evaluating the discriminant properties of each normalized entropy feature and distinguishing between the conditions tested in the IM. Hence, Table 8 summarizes the  $FS_r$  scores achieved when the HLT condition is faced with each faulty condition and the supply electrical frequencies of operation used in the kinematic chain are 5 Hz and 15 Hz. Meanwhile, Table 9 summarizes the  $FS_r$  scores obtained when the kinematic chain operates at 50 Hz and 60 Hz. As appreciated in Table 8, only a few entropy features for some specific signals lead to the production of  $FS_r$  scores higher than 1 for the operating frequency of 5 Hz, whereas for 15 Hz, most of the  $FS_r$  scores are higher than 1; thereby, it could be possible to obtain misclassifications when  $FS_r$  scores smaller than 1 are produced. On the other hand, as depicted in Table 9, the values of  $FS_r$  scores are higher than 1 for almost all the entropy features for the three signals, and  $FS_r > 1$  may ensure the separability of assessed conditions. As previously stated, a specific entropy feature may not provide sufficient information to separate the healthy condition from the faulty ones. In this sense,  $C_3^{18} = 816$  possible combinations of three entropy features are also evaluated using Equations (13)–(15), accordingly; during the analysis, the HLT condition is faced with the BD, BRB, UNB, and MIS conditions for each operating supplying frequency of the kinematic chain (5 Hz, 15 Hz, 50 Hz, and 60 Hz). Table 10 summarizes the first three ranked  $FS_r$  values, as appreciated, and only the  $FS_r$  scores achieved when the HLT condition is faced with the MIS condition (operating at 5 Hz) produce  $FS_r < 1$ , whereas the remaining  $FS_r$  scores lead to  $FS_r > 1$ . Finally, the t-SNE is used to obtain the visual representation of the best-ranked features (last row in Table 10) in a 2D space; thus, Figure 10a–d show the data distributions for the best-ranked features when the kinematic chain operates at 5 Hz, 15 Hz, 50 Hz, and 60 Hz, respectively. It must be emphasized that all tested conditions in the IM (HLT, BD, BRB, UNB, and MIS) appear separated from each other for the operating frequencies of 50 Hz and 60 Hz (Figure 10c,d). Meanwhile, in Figure 10b, all the faulty conditions are separated from the HLT one when the kinematic chain operates at 15 Hz; nevertheless, an unexpected overlapping is produced between BD and UNB conditions. Unfortunately, in Figure 10a, the HLT, BD, UNB, and MIS conditions appear to overlap

between them when the kinematic chain operates at 5 Hz, and only the BRB condition tries to be separated from the other conditions. Therefore, it is worth mentioning that the use of entropy features estimated from vibration signals and stator current signatures can lead to the accurate diagnosis of faults like BD, BRB, UNB, and MIS in IMs; however, the assessment of faults in IMs through entropy features can be sensitive to the rotating speed, specifically, at high operating frequencies (50 Hz and 60 Hz), an accurate assessment of faults is performed, but some limitations are presented during the fault diagnosis at an extremely low operating frequency (5 Hz).



**Figure 10.** Resulting projection achieved through the t-SNE during the evaluation of the best-ranked entropy features when the HLT condition in the IMs is faced with BD, BRB, UNB, and MIS for the kinematic chain operating at (a) 5 Hz, (b) 15 Hz, (c) 50 Hz, and (d) 60 Hz.

**Table 8.** Quantitative Fisher score values achieved for the experimentation of different faulty conditions in the IM when the kinematic chain operates at 5 Hz and 15 Hz.

		Supply Electrical Frequencies of Operation							
		5 Hz				15 Hz			
Entropy Features	Physical Magnitude	HLT vs. BD	HLT vs. BRB	HLT vs. UNB	HLT vs. MIS	HLT vs. BD	HLT vs. BRB	HLT vs. UNB	HLT vs. MIS
<i>SpecEn</i>	$V_y$	0.04	0.03	0.00	0.00	0.09	23.79	2.62	0.49
	$V_z$	0.07	0.02	0.00	0.00	0.03	1.53	0.01	24.07
	$C$	0.39	0.12	0.47	0.28	0.00	0.01	0.00	2.05
<i>PermEn</i>	$V_y$	0.01	0.00	0.01	0.00	0.00	1.32	0.00	0.30
	$V_z$	0.00	0.01	0.00	0.00	0.01	2.64	0.00	0.36
	$C$	0.00	0.00	0.03	0.00	0.01	0.01	0.03	0.19



**Table 10.** Quantitative Fisher score values achieved when combinations of three entropy features are considered during the evaluation of different faulty conditions in the IMs for all operating frequencies.

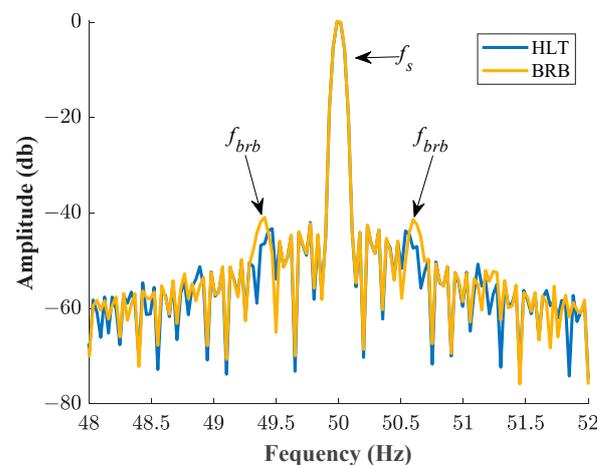
		Supply Electrical Frequencies of Operation							
		5 Hz		15 Hz		50 Hz		60 Hz	
		$FS_r$	Combined Features	$FS_r$	Combined Features	$FS_r$	Combined Features	$FS_r$	Combined Features
Faced conditions	HLT vs. BD	2.03	5, 11, 13	15	11, 12, 15	283	2, 5, 10	46	4, 5, 11
		1.39	5, 11, 17	15	9, 11, 12	280	2, 5, 17	45	1, 4, 5
		1.38	5, 7, 11	15	8, 11, 12	278	2, 5, 12	44	4, 5, 9
	HLT vs. BRB	20.26	5, 11, 13	496	5, 8, 11	221	1, 5, 7	110	4, 5, 11
		20.19	5, 10, 11	458	5, 11, 15	213	2, 5, 7	108	1, 4, 5
		19.05	5, 11, 17	446	5, 11, 16	208	1, 2, 7	105	4, 5, 9
	HLT vs. UNB	1.23	11, 13, 17	32	4, 8, 11	1392	1, 5, 11	110	1, 5, 12
		1.12	13, 15, 17	31	3, 8, 11	1186	2, 5, 11	108	1, 5, 7
		1.05	13, 16, 17	30	2, 8, 11	1086	5, 9, 11	105	5, 7, 12
	HLT vs. MIS	0.40	5, 11, 13	52	5, 7, 11	18,882	11, 12, 14	713	5, 13, 14
		0.30	3, 5, 13	51	7, 11, 13	18,677	11, 12, 16	698	5, 10, 13
		0.29	5, 13, 17	51	4, 7, 11	18,642	11, 12, 15	673	1, 5, 13
	Best-ranked features	2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 15, 16		3, 5, 7, 10, 11, 13, 15, 16, 17		1, 2, 5, 7, 9, 10, 11, 12, 14, 15, 16, 17		1, 4, 5, 7, 9, 10, 11, 12, 13, 14	

### 5.3. Comparison versus Classical Approaches

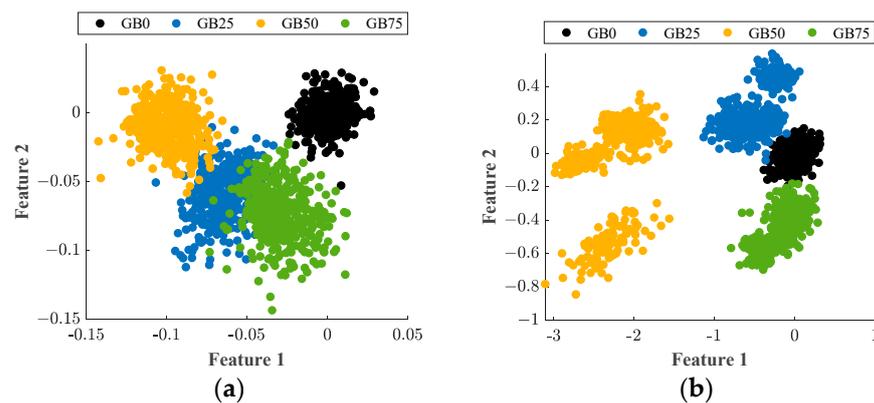
Finally, a comparison with classical approaches is performed to highlight that entropy features provide meaningful information to carry out the detection of faults in GBs and IMs; in this regard, the estimation of frequency spectra through the FFT has represented a suitable diagnosis tool for detecting the occurrence of faults in rotating machinery. For example, the detection of problems associated with rotor bars in IMs is commonly addressed due to the unbalance of the air-gap magnetic flux leading to the generation of a fault-related frequency component around the supply frequency ( $f_s$ ) in the frequency spectrum at  $f_{brb} = (1 \pm 2ks)f_s$ ,  $k + 1, 2, 3, \dots$ . Nevertheless, the detection of faults based on the analysis of frequency spectra requires prior knowledge and expertise since the calculation of fault-related frequency components depends on additional variables such as the per-unit slip ( $s = (n_s - n_r)/n_s$ ), which, in turn, is in terms of the synchronous ( $n_s$ ) and rotating rotor speed ( $n_r$ ). Figure 11 shows the obtained frequency spectra by applying the FFT technique to the stator current signatures acquired when the IM is tested under the HLT and BRB conditions at 50 Hz as the supply electrical frequency of operation, as observed, the spectrum associated with the BRB condition produces the  $f_{brb}$  components appearing at 49.4 Hz and 50.59 Hz for a slip of around 1.5%. Although faults like BRBs introduce modifications in specific fault-related frequency components, the accuracy of the diagnosis depends on the technician’s knowledge. Hence, the analysis of changes and trends through the estimation of entropy features represents a suitable option to achieve the occurrence of any fault in rotating machinery like IMs or GBs.

On the other hand, DDA can be successfully applied for detecting and identifying faults in rotating machines; however, the performance achieved by these approaches depends on two main stages where a set of fault-related features is estimated and then subjected to a reduction procedure. That is, the statistical features such as RMS value, Standard Deviation, Kurtosis, Skewness, Shape Factor, and Latitude factor are commonly computed from raw signals aiming to model trends and changes; meanwhile, dimensionality reduction techniques such as PCA and/or LDA are used to compress and transform an original dataset into a 2D feature space. Consequently, to validate that the entropy features studied in this work provide significant information for fault diagnosis methodologies, the aforementioned statistical features have been computed from the raw signals  $V_y$ ,  $V_z$  and  $C_1$  for all tested conditions (GB0, GB25, GB50, and GB75) in the GB under all operating

frequencies (5 Hz, 15 Hz, 50 Hz, and 60 Hz), and then are subjected to a reduction procedure using the LDA technique. Figure 12a shows the resulting 2D projection where it is possible to observe four main clusters representing the assessed conditions in the GB; unfortunately, the statistical features lead to overlapping between all faulty conditions. Additionally, to compare the performance between statistical features and entropy features, the entropy features studied in this work are also subjected to the reduction procedure by means of the LDA technique. Figure 12b presents the resulting 2D projection where an improvement between the separability of all considered conditions is obtained. It should be mentioned that for both sets of features (statistical and entropy), the reduction procedure has been applied to all samples of all studied conditions under all tested frequencies. Thereby, the obtained results support the use of entropy features as a part of condition monitoring methodologies for the detection and identification of faults that can occur in rotating elements like GBs and IMs.



**Figure 11.** Frequency spectra obtained by applying the FFT technique over the acquired stator current signature when the conditions of HLT and BRB are tested in the IM and the supply frequency is fixed at 50 Hz.



**Figure 12.** Resulting 2D projection obtained by applying the LDA technique to (a) statistical and (b) entropy features proposed in this work, both estimated from signals  $V_y$ ,  $V_z$ , and  $C_1$ . In the resulting projection, all evaluated frequencies have been considered (5 Hz, 15 Hz, 50 Hz, and 60 Hz).

Finally, Table 11 presents the most important highlights obtained during the comparison between the classical frequency analysis (performed by the FFT) and by comparing DDA in front of the entropy analysis performed in this work. In fact, it can be emphasized that the most critical issue is associated with prior knowledge and expertise for detecting those fault-related frequency components in a frequency spectrum; meanwhile, the use of entropy features and dimensionality reduction techniques, such as LDA, leads to improve

the diagnosis outcome. On the other hand, Table 12 summarizes the highlights of the comparison between the entropy features analysis versus classical statistical features, both used in a DDA using the LDA technique, and the estimation of statistical time domain features is a suitable option to characterize trends and changes in time domain signals with an appropriate trade-off between simplicity and quickly responses; however, the obtained results demonstrate that entropy features significantly improve the assessment of faults regardless of the operating condition.

**Table 11.** Summary of the highlights found of the analyzed entropy features versus classical frequency analysis through the FFT.

Approach	Fault Detection in IMs Two Conditions: HLT and BRB	
	Advantages	Disadvantages
Classical approach FFT analysis	<ul style="list-style-type: none"> <li>• Low computational burden.</li> <li>• Simple formulation.</li> </ul>	<ul style="list-style-type: none"> <li>• Prior knowledge is required.</li> <li>• High dependency of spectral frequency content.</li> <li>• Faults with similar frequency behavior yield misleading fault detection.</li> <li>• Resolution restriction in the sampling frequency.</li> </ul>
Data-driven approach Entropy features and LDA	<ul style="list-style-type: none"> <li>• Entropy features provide wide energy behavior of the signal.</li> <li>• Most of the indicators provide useful information.</li> <li>• Could be helpful for signals with non-linearities and non-stationary behavior.</li> <li>• The general results have clear cluster separation.</li> </ul>	<ul style="list-style-type: none"> <li>• Complex computation and medium computational burden.</li> </ul>

**Table 12.** Summary of the highlights found of the analyzed entropy features versus statistical features, both used in the same data-driven approach with LDA.

Approach	Fault Detection in GBs Four Conditions: GB0, GB25, GB50, and GB75	
	Advantages	Disadvantages
Data-driven approach Statistical features and LDA	<ul style="list-style-type: none"> <li>• Statistical features are ease of computing work and require a low computational burden.</li> <li>• Statistical features provide wide behavior in geometry, asymmetry, shape, central tendencies, trends, and dispersion.</li> </ul>	<ul style="list-style-type: none"> <li>• Most of the indicators provide redundant information.</li> <li>• Sensible to non-linearities and non-stationary behavior.</li> <li>• The general results have clusters overlapping.</li> </ul>
Data-driven approach Entropy features and LDA	<ul style="list-style-type: none"> <li>• Entropy features provide wide energy behavior of the signal.</li> <li>• Most of the indicators provide useful information.</li> <li>• Could be helpful for signals with non-linearities and non-stationary behavior.</li> <li>• The general results have clear cluster separation.</li> </ul>	<ul style="list-style-type: none"> <li>• Complex computation and medium computational burden.</li> </ul>

## 6. Conclusions

In this work, an analysis of entropy-based features was performed with the objective of identifying changes and trends produced by the unexpected occurrence of faults in a

kinematic chain. Certainly, the proposed analysis is applied to vibration signals and stator current signatures acquired from real laboratory experiments, where different severities of uniform wear are tested in a GB and five different faulty conditions are tested in an IM. There are some important aspects that must be highlighted: the first one is that the calculation of the six proposed entropy features from available vibrations and stator currents leads to a high-performance characterization to assess the occurrence of faults in both GB and IM linked in a kinematic chain that operates at different supply electrical frequencies of operation. Indeed, each different condition under study produces different amplitude values during the calculation of entropy features; thus, any variation obtained in regard to the healthy condition is produced due to the occurrence of an undesirable faulty condition. The second aspect to be highlighted is that, in the analysis of entropy features using the KW test and the FDS technique, both approaches allow us to determine whether the proposed entropy features have statistically significant information for distinguishing between different conditions (classes). Certainly, through boxplots resulting from the KW test, it can be inferred when data points of the analyzed conditions can be separated from each other where the separation between classes is related to the vertical separation that appears between medians in the boxplots. On the other hand, it is demonstrated that overlapping problems are avoided when the values of  $FS_r$  scores are higher than 1; additionally, it has been proved that the combination of entropy features contributes to a better separation between assessed conditions. The final aspect is that the use of entropy features provides significant information to perform the fault assessment for the detection of different severities of uniform wear in a GB regardless of the rotating frequency; on the contrary, when entropy features are applied to the diagnosis of faults in IMs, it may be possible to produce accurate detection of faults whether the IM operates at high supply frequencies but a lack of discriminant capabilities is present at extremely low operating frequencies. Thus, from the experimental tests, it can be concluded that entropy features have the advantage of providing wide energy behavior into a measured signal and, therefore, most of the indicators analyzed provided useful discriminatory information of the conditions; thus, the feature values were robust and consistent. This way, a question appears: why and when should the entropy-based features be used? The answer is that they can be applied in systems under unfavorable conditions like low values of the load (where most of the proposed approaches present difficulties for the diagnosis), also with different types of faults, not necessarily related to the bearing defects, like unbalance, misalignment, broken rotor bars, and gear wear. Thereby, entropy features represent a suitable option to be implemented as a part of condition monitoring strategies for the detection and identification of faults in different elements involved in a kinematic chain such as GBs and IMs. Future work should consider the combination of techniques in which entropy-based features can be exploited for diagnosing faults in non-stationary signals for combined faults, or even under unfavorable noise conditions combined with variable operating conditions.

**Author Contributions:** Conceptualization, D.A.E.-O. and R.d.J.R.-T.; Data Curation, D.A.E.-O. and R.d.J.R.-T.; Formal Analysis, J.J.S.-D. and R.d.J.R.-T.; Funding Acquisition, A.Y.J.-C.; Investigation, A.Y.J.-C. and R.d.J.R.-T.; Methodology, A.Y.J.-C., J.J.S.-D. and D.A.E.-O.; Project Administration, A.Y.J.-C. and R.d.J.R.-T.; Resources, A.Y.J.-C.; Supervision, A.Y.J.-C., J.J.S.-D. and R.d.J.R.-T.; Validation, J.J.S.-D.; Visualization, D.A.E.-O. and R.d.J.R.-T.; Writing—Original Draft, J.J.S.-D. and D.A.E.-O.; Writing—Review and Editing, A.Y.J.-C. and J.J.S.-D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project has been partially supported by the Universidad Autonoma de Querétaro through “Fondo para el Desarrollo del Conocimiento” (FONDEC-UAQ-2022) under project number 20205007071601, and partially by the CONAHcyT 487599, 331141, 558497 and 120260 SNII grants.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. These data can be found here: [<https://grouper.ieee.org/groups/1159/3/docs.html>], (accessed on 14 April 2024)].

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

## References

1. Sardar, M.U.; Vaimann, T.; Kütt, L.; Kallaste, A.; Asad, B.; Akbar, S.; Kudelina, K. Inverter-Fed Motor Drive System: A Systematic Analysis of Condition Monitoring and Practical Diagnostic Techniques. *Energies* **2023**, *16*, 5628. [CrossRef]
2. Drakaki, M.; Karnavas, Y.L.; Tzionas, P.; Chasiotis, I.D. Recent Developments Towards Industry 4.0 Oriented Predictive Maintenance in Induction Motors. *Procedia Comput. Sci.* **2021**, *180*, 943–949. [CrossRef]
3. Torrent, M.; Blanqué, B.; Monjo, L. Replacing Induction Motors without Defined Efficiency Class by IE Class: Example of Energy, Economic, and Environmental Evaluation in 1.5 kW—IE3 Motors. *Machines* **2023**, *11*, 567. [CrossRef]
4. Zhang, J.; Wang, P.; Gao, R.X.; Sun, C.; Yan, R. Induction Motor Condition Monitoring for Sustainable Manufacturing. *Procedia Manuf.* **2019**, *33*, 802–809. [CrossRef]
5. Gangsar, P.; Tiwari, R. Signal Based Condition Monitoring Techniques for Fault Detection and Diagnosis of Induction Motors: A State-of-the-Art Review. *Mech. Syst. Signal Process.* **2020**, *144*, 106908. [CrossRef]
6. Workineh, T.G.; Jember, Y.B.; Kassie, A.T. Evaluation of Intelligent PPI Controller for the Performance Enhancement of Speed Control of Induction Motor. *Sci. Afr.* **2023**, *22*, e01982. [CrossRef]
7. Chen, C.; Yuan, Y.; Sun, W.; Zhao, F. Multivariate Multi-Step Time Series Prediction of Induction Motor Situation Based on Fused Temporal-Spatial Features. *Int. J. Hydrog. Energy* **2024**, *50*, 1386–1394. [CrossRef]
8. Singh, G.; Anil Kumar, T.C.; Naikan, V.N.A. Efficiency Monitoring as a Strategy for Cost Effective Maintenance of Induction Motors for Minimizing Carbon Emission and Energy Consumption. *Reliab. Eng. Syst. Saf.* **2019**, *184*, 193–201. [CrossRef]
9. Rangel-Magdaleno, J.; Peregrina-Barreto, H.; Ramirez-Cortes, J.; Cruz-Vega, I. Hilbert Spectrum Analysis of Induction Motors for the Detection of Incipient Broken Rotor Bars. *Measurement* **2017**, *109*, 247–255. [CrossRef]
10. Kumar, S.; Mukherjee, D.; Guchhait, P.K.; Banerjee, R.; Srivastava, A.K.; Vishwakarma, D.N.; Saket, R.K. A Comprehensive Review of Condition Based Prognostic Maintenance (CBPM) for Induction Motor. *IEEE Access* **2019**, *7*, 90690–90704. [CrossRef]
11. Mallik, S.; Mallik, K.; Barman, A.; Maiti, D.; Biswas, S.K.; Deb, N.K.; Basu, S. Efficiency and Cost Optimized Design of an Induction Motor Using Genetic Algorithm. *IEEE Trans. Ind. Electron.* **2017**, *64*, 9854–9863. [CrossRef]
12. IEA. World Energy Outlook 2023—Analysis, International Energy Agency. Available online: <https://www.iea.org/reports/world-energy-outlook-2023> (accessed on 28 February 2024).
13. Bessous, N.; Sbaa, S.; Toumi, A. Experimental Investigation on Broken Rotor Bar Faults in Three Phase Induction Motors Using MVSA-FFT Method. In Proceedings of the 2018 6th International Conference on Control Engineering & Information Technology (CEIT), Istanbul, Turkey, 25–27 October 2018; pp. 1–7.
14. Kompella, K.C.D.; Mannam, V.G.R.; Rayapudi, S.R. DWT Based Bearing Fault Detection in Induction Motor Using Noise Cancellation. *J. Electr. Syst. Inf. Technol.* **2016**, *3*, 411–427. [CrossRef]
15. Jain, S.K.; Singh, S.N. Harmonics Estimation in Emerging Power System: Key Issues and Challenges. *Electr. Power Syst. Res.* **2011**, *81*, 1754–1766. [CrossRef]
16. Chen, Z.; Liang, D.; Jia, S.; Yang, L.; Yang, S. Incipient Interturn Short-Circuit Fault Diagnosis of Permanent Magnet Synchronous Motors Based on the Data-Driven Digital Twin Model. *IEEE J. Emerg. Sel. Top. Power Electron.* **2023**, *11*, 3514–3524. [CrossRef]
17. Bittencourt, A.C.; Saarinen, K.; Sander-Tavallaey, S.; Gunnarsson, S.; Norrlöf, M. A Data-Driven Approach to Diagnostics of Repetitive Processes in the Distribution Domain—Applications to Gearbox Diagnostics in Industrial Robots and Rotating Machines. *Mechatronics* **2014**, *24*, 1032–1041. [CrossRef]
18. Rosa, T.E.; de Paula Carvalho, L.; Gleizer, G.A.; Jayawardhana, B. Fault Detection for LTI Systems Using Data-Driven Dissipativity Analysis. *Mechatronics* **2024**, *97*, 103111. [CrossRef]
19. Jaen-Cuellar, A.Y.; Elvira-Ortiz, D.A.; Saucedo-Dorantes, J.J. Statistical Machine Learning Strategy and Data Fusion for Detecting Incipient ITSC Faults in IM. *Machines* **2023**, *11*, 720. [CrossRef]
20. Wang, H.; Zheng, J.; Xiang, J. Online Bearing Fault Diagnosis Using Numerical Simulation Models and Machine Learning Classifications. *Reliab. Eng. Syst. Saf.* **2023**, *234*, 109142. [CrossRef]
21. Das, O.; Bagci Das, D.; Birant, D. Machine Learning for Fault Analysis in Rotating Machinery: A Comprehensive Review. *Heliyon* **2023**, *9*, e17584. [CrossRef]
22. Saucedo-Dorantes, J.J.; Jaen-Cuellar, A.Y.; Perez-Cruz, A.; Elvira-Ortiz, D.A. Detection of Inter-Turn Short Circuits in Induction Motors under the Start-Up Transient by Means of an Empirical Wavelet Transform and Self-Organizing Map. *Machines* **2023**, *11*, 958. [CrossRef]
23. Arellano-Espitia, F.; Delgado-Prieto, M.; Martínez-Viol, V.; Saucedo-Dorantes, J.-J.; Osornio-Rios, R.A. Diagnosis Electromechanical System by Means CNN and SAE: An Interpretable-Learning Study. In Proceedings of the 2022 IEEE 5th International Conference on Industrial Cyber-Physical Systems (ICPS), Coventry, UK, 24–26 May 2022; pp. 1–6.
24. Chen, M.; Shao, H.; Dou, H.; Li, W.; Liu, B. Data Augmentation and Intelligent Fault Diagnosis of Planetary Gearbox Using ILoFGAN Under Extremely Limited Samples. *IEEE Trans. Reliab.* **2023**, *72*, 1029–1037. [CrossRef]
25. Ding, A.; Qin, Y.; Wang, B.; Cheng, X.; Jia, L. An Elastic Expandable Fault Diagnosis Method of Three-Phase Motors Using Continual Learning for Class-Added Sample Accumulations. *IEEE Trans. Ind. Electron.* **2023**, *71*, 7896–7905. [CrossRef]
26. Choudhary, A.; Mian, T.; Fatima, S.; Panigrahi, B.K. Passive Thermography Based Bearing Fault Diagnosis Using Transfer Learning With Varying Working Conditions. *IEEE Sens. J.* **2023**, *23*, 4628–4637. [CrossRef]
27. Akhtar, S.; Adeel, M.; Iqbal, M.; Namoun, A.; Tufail, A.; Kim, K.-H. Deep Learning Methods Utilization in Electric Power Systems. *Energy Rep.* **2023**, *10*, 2138–2151. [CrossRef]

28. Liu, L.; Zhi, Z.; Zhang, H.; Guo, Q.; Peng, Y.; Liu, D. Related Entropy Theories Application in Condition Monitoring of Rotating Machineries. *Entropy* **2019**, *21*, 1061. [[CrossRef](#)]
29. Li, Y.; Wang, X.; Liu, Z.; Liang, X.; Si, S. The Entropy Algorithm and Its Variants in the Fault Diagnosis of Rotating Machinery: A Review. *IEEE Access* **2018**, *6*, 66723–66741. [[CrossRef](#)]
30. Li, Y.; Wang, X.; Si, S.; Huang, S. Entropy Based Fault Classification Using the Case Western Reserve University Data: A Benchmark Study. *IEEE Trans. Reliab.* **2020**, *69*, 754–767. [[CrossRef](#)]
31. Huo, Z.; Martinez-Garcia, M.; Zhang, Y.; Yan, R.; Shu, L. Entropy Measures in Machine Fault Diagnosis: Insights and Applications. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 2607–2620. [[CrossRef](#)]
32. Leite, G.D.N.P.; Araújo, A.M.; Rosas, P.A.C.; Stosic, T.; Stosic, B. Entropy Measures for Early Detection of Bearing Faults. *Phys. Stat. Mech. Its Appl.* **2019**, *514*, 458–472. [[CrossRef](#)]
33. Wan, S.; Zhang, X. Teager Energy Entropy Ratio of Wavelet Packet Transform and Its Application in Bearing Fault Diagnosis. *Entropy* **2018**, *20*, 388. [[CrossRef](#)]
34. Fu, L.; Zhu, T.; Zhu, K.; Yang, Y. Condition Monitoring for the Roller Bearings of Wind Turbines under Variable Working Conditions Based on the Fisher Score and Permutation Entropy. *Energies* **2019**, *12*, 3085. [[CrossRef](#)]
35. Ramteke, D.S.; Pachori, R.B.; Parey, A. Automated Gearbox Fault Diagnosis Using Entropy-Based Features in Flexible Analytic Wavelet Transform (FAWT) Domain. *J. Vib. Eng. Technol.* **2021**, *9*, 1703–1713. [[CrossRef](#)]
36. Wang, X.; Si, S.; Li, Y. Hierarchical Diversity Entropy for the Early Fault Diagnosis of Rolling Bearing. *Nonlinear Dyn.* **2022**, *108*, 1447–1462. [[CrossRef](#)]
37. Yuan, X.; Fan, Y.; Zhou, C.; Wang, X.; Zhang, G. Research on Twin Extreme Learning Fault Diagnosis Method Based on Multi-Scale Weighted Permutation Entropy. *Entropy* **2022**, *24*, 1181. [[CrossRef](#)] [[PubMed](#)]
38. Cao, W.; Li, C.; Li, Z.; Wang, S.; Pang, H. Research on Low-Resistance Grounding Fault Line Selection Based on VMD with PE and K-Means Clustering Algorithm. *Math. Probl. Eng.* **2022**, *2022*, e5360302. [[CrossRef](#)]
39. Kamiński, M.M. On Shannon Entropy Computations in Selected Plasticity Problems. *Int. J. Numer. Methods Eng.* **2021**, *122*, 5128–5143. [[CrossRef](#)]
40. Altaf, S.; Soomro, M.W.; Mehmood, M.S. Fault Diagnosis and Detection in Industrial Motor Network Environment Using Knowledge-Level Modelling Technique. *Model. Simul. Eng.* **2017**, *2017*, e1292190. [[CrossRef](#)]
41. Hassan, O.E.; Amer, M.; Abdelsalam, A.K.; Williams, B.W. Induction Motor Broken Rotor Bar Fault Detection Techniques Based on Fault Signature Analysis—A Review. *IET Electr. Power Appl.* **2018**, *12*, 895–907. [[CrossRef](#)]
42. Feng, G.; Hu, N.; Mones, Z.; Gu, F.; Ball, A.D. An Investigation of the Orthogonal Outputs from an On-Rotor MEMS Accelerometer for Reciprocating Compressor Condition Monitoring. *Mech. Syst. Signal Process.* **2016**, *76–77*, 228–241. [[CrossRef](#)]
43. Xu, Y.; Tang, X.; Feng, G.; Wang, D.; Ashworth, C.; Gu, F.; Ball, A. Orthogonal On-Rotor Sensing Vibrations for Condition Monitoring of Rotating Machines. *J. Dyn. Monit. Diagn.* **2021**, *1*, 29–36. [[CrossRef](#)]
44. Ramteke, D.S.; Parey, A.; Pachori, R.B. A New Automated Classification Framework for Gear Fault Diagnosis Using Fourier–Bessel Domain-Based Empirical Wavelet Transform. *Machines* **2023**, *11*, 1055. [[CrossRef](#)]
45. Jamil, M.A.; Khanam, S. Influence of One-Way ANOVA and Kruskal–Wallis Based Feature Ranking on the Performance of ML Classifiers for Bearing Fault Diagnosis. *J. Vib. Eng. Technol.* **2023**. [[CrossRef](#)]
46. Braga, J.A.P.; Andrade, A.R. Multivariate Statistical Aggregation and Dimensionality Reduction Techniques to Improve Monitoring and Maintenance in Railways: The Wheelset Component. *Reliab. Eng. Syst. Saf.* **2021**, *216*, 107932. [[CrossRef](#)]
47. Hu, Q.; Si, X.-S.; Zhang, Q.-H.; Qin, A.-S. A Rotating Machinery Fault Diagnosis Method Based on Multi-Scale Dimensionless Indicators and Random Forests. *Mech. Syst. Signal Process.* **2020**, *139*, 106609. [[CrossRef](#)]
48. Tripanpitak, K.; Viriyavit, W.; Huang, S.Y.; Yu, W. Classification of Pain Event Related Potential for Evaluation of Pain Perception Induced by Electrical Stimulation. *Sensors* **2020**, *20*, 1491. [[CrossRef](#)] [[PubMed](#)]
49. Rostami, M.; Forouzandeh, S.; Berahmand, K.; Soltani, M. Integration of Multi-Objective PSO Based Feature Selection and Node Centrality for Medical Datasets. *Genomics* **2020**, *112*, 4370–4384. [[CrossRef](#)]
50. Wan, H.; Wang, H.; Scotney, B.; Liu, J.; Ng, W.W.Y. Within-Class Multimodal Classification. *Multimed. Tools Appl.* **2020**, *79*, 29327–29352. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.