

Article

Physics-Informed Neural Networks for the Condition Monitoring of Rotating Shafts

Marc Parziale , Luca Lomazzi , Marco Giglio and Francesco Cadini *

Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156 Milan, Italy; marc.parziale@polimi.it (M.P.); luca.lomazzi@polimi.it (L.L.); marco.giglio@polimi.it (M.G.)

* Correspondence: francesco.cadini@polimi.it

Abstract: Condition monitoring of rotating shafts is essential for ensuring the reliability and optimal performance of machinery in diverse industries. In this context, as industrial systems become increasingly complex, the need for efficient data processing techniques is paramount. Deep learning has emerged as a dominant approach due to its capacity to capture intricate data patterns and relationships. However, a prevalent challenge lies in the black-box nature of many deep learning algorithms, which often operate without adhering to the underlying physical characteristics intrinsic to the studied phenomena. To address this limitation and enhance the fusion of data-driven methodologies with the fundamental physics of the system under study, this paper leverages physics-informed neural networks (PINNs). Specifically, a simple but realistic numerical case study of an extended Jeffcott rotor model, encompassing damping effects and anisotropic supports for a more comprehensive modelling, is considered. PINNs are used for the estimation of five parameters that characterize the health state of the system. These parameters encompass the radial and angular position of the static unbalance due to the disk installed on the shaft, the stiffness along the principal axes of elasticity, and the non-rotating damping coefficient. The estimation is conducted solely by exploiting the displacement signals from the centre of the disk and, to showcase the efficacy and precision provided by this novel methodology, various scenarios involving different constant rotational speeds are examined. Additionally, the impact of noisy input data is also taken into account within the analysis and the performance is compared to that of traditional optimization algorithms used for parameters estimation.

Keywords: condition monitoring; rotating shaft; physics-informed neural network; parameters estimation



Citation: Parziale, M.; Lomazzi, L.; Giglio, M.; Cadini, F. Physics-Informed Neural Networks for the Condition Monitoring of Rotating Shafts. *Sensors* **2024**, *24*, 207. <https://doi.org/10.3390/s24010207>

Academic Editors: Yongfeng Yang, Jin Zhou, Rafael Morales, Alexandre Presas and Saleh Mobayen

Received: 21 November 2023
Revised: 19 December 2023
Accepted: 28 December 2023
Published: 29 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rotating shafts are elements of engineering systems that play a paramount role in the transmission of power, encompassing speed and torque, from one point to another [1,2]. They are typically designed to endure substantial loads and to operate at high velocities, underscoring the need for precise alignment, equilibrium, and freedom from imperfections. These considerations are significant not only for enhancing the overall system performance, but also for improving its safety and reliability [3,4]. To attain this objective, the practice of condition monitoring (CM) for rotating shafts allows continuously evaluating the shaft condition and performance and detecting any indications of malfunction or deterioration [5,6]. Through the application of CM methodologies, potential issues can be promptly identified, and maintenance or repair actions can be driven to prevent accidents. This approach has shown to effectively mitigate the risk of unexpected downtime, reinforcing the overall system reliability [7,8]. Furthermore, the widespread and cost-effective availability of sensors has revolutionized the acquisition of diagnostic signals, such as accelerations, strains, and elastic waves [9]. However, the availability of big data is itself a new layer of complexity, especially in the realm of signal processing. That is, the rapid expansion of the amount of acquired data has introduced the need for (i) improved hardware and software performance,

and (ii) developing tools to deal with confounding factors, including those unrelated to the system health state, such as environmental and operational conditions [10,11].

To tackle these challenges, deep learning has stood as a pivotal technological advancement in the CM of rotating machines, offering multifaced contributions of significant importance [12]. This approach has excelled in automatically extracting intricate patterns and features from raw sensor data, enhancing the precision and reliability of fault detection and anomaly characterization. As an example, the work in [13] applied deep learning to enhance wind turbine CM, addressing the data surge from increased wind farm units. By combining convolutional neural networks (CNNs) [14] and recurrent neural networks (RNNs) [15], it efficiently extracted features, reduced dimensionality, and provided effective CM, offering both real-time unit state checks and early warning capability, even amidst accidental parameter changes. In [16], a CM model based on CNNs for automatic fault detection in rotating equipment was developed. The model, utilizing data from a single vibration sensor on the motor-drive end bearing, achieved accuracies of 99.58% and 97.3% when applied to two different databases under controlled ambient conditions. Another example was presented in [17], where the authors proposed a novel deep learning algorithm for detecting rotor unbalance in industrial machinery. The algorithm, extracting important vibration signatures such as fast Fourier transform (FFT) and short-time Fourier transform (STFT), combined the depth of ResNet [18] and the feature extraction capability of CNN. This hybrid approach surpassed the performance of both individual models. The study involved two analyses: binary detection of balanced vs. unbalanced cases and multilevel detection of the degree of unbalance. The work in [19] addressed planetary gearbox fault detection by representing baseline vibration signals using the varying index coefficient autoregression (VICAR) model. The authors proposed a modified VICAR (MVICAR) model to effectively incorporate rotating speed into the representation while maintaining nonlinear modeling capacity. Experimental results demonstrated the superiority of the MVICAR model over autoencoders, expanded VICAR (EVICAR), and linear parameter-varying autoregression models in planetary gearbox fault detection. In [20], a semi-supervised fault diagnosis approach for wind turbines was introduced. The method utilized a deep neural network with adversarial learning and incorporated a metric-guided feature enhancement technique. Despite having a limited number of annotated samples, the methodology exhibited superior fault diagnosis accuracy in experiments conducted on a wind turbine fault dataset.

However, in the context of CM, the developed methods have predominantly relied on black-box deep learning algorithms, lacking transparency in how input data are processed and whether the network behavior aligns with the physics of the problem [21]. Existing approaches to address this issue involve either post-training explainability algorithms or more intricate physics-based deep learning models. The former, while debunking network behavior, fails to provide evidence of adherence to physical laws [22–25]. On the other hand, the latter ensures predictions align with the physics by incorporating regularization terms representing known physical laws during training. These terms are integrated into the network loss function, specifically at the stage where it quantifies the disparity between predicted and actual outcomes. This critical addition serves to guide the neural network towards solutions that not only capture intricate patterns from data but also adhere rigorously to the established physical laws, enhancing the reliability and interpretability of physics-informed neural network (PINN) predictions. The regularization terms act as foundational constraints, influencing the network learning to prioritize solutions that respect the governing physics throughout the training iterations. Moreover, physics-informed algorithms offer a distinct advantage by providing accurate predictions even in the presence of scarce data, a capability not shared by traditional deep learning methods. Notably, physics-informed algorithms are versatile tools applicable in various contexts, including data-driven solutions for partial differential equations, discovery of physical laws, and parameter estimation [26–28]. However, within the CM domain, few contributions have integrated physical knowledge effectively into the training process of deep learning models.

In [29], the authors introduced a novel approach for fault detection in gearboxes using long-short term memory (LSTM) neural networks. Given a lack of data from faulty states, the authors proposed a physics-informed hyperparameter selection strategy for LSTM identification, emphasizing maximizing the discrepancy between healthy and physics-informed faulty states. Case studies on detecting gear tooth crack and tooth wear demonstrated that the approach outperformed traditional methods based on minimizing validation mean squared error (VAMSE). The work in [30] presented a physics-informed deep learning method for bearing fault detection that combined a threshold model and a CNN. The approach was validated using data from bearings on an agricultural machine and a laboratory test stand in the Case Western Reserve University Bearing Data Centre. In [31], a method for identifying unbalance faults in rotary systems using physics-guided neural networks (PGNNs) was proposed. The approach involved the use of a standard neural network to localize the nodal position of the experimental fault, followed by PGNN to quantify the unbalance magnitude and phase angle. Instead, the work in [32] introduced a novel physics-informed convolution long-short-term memory (LSTM-CNN) network for rotor unbalance and shaft cracks detection and localization. In particular, the physics were taken into account through the construction of a neural network model which mimicked a finite element (FE) resolution of the problem.

To the best of the authors' knowledge, still no efforts have been made for the direct estimation of multiple parameters characterizing the health state of a rotating shaft system by leveraging PINNs. In this work, PINNs are utilized to estimate critical health state parameters in a simple but realistic numerical case of an extended Jeffcott rotor model. This model incorporates damping effects and anisotropic supports for a more comprehensive representation. The parameters under consideration include the radial and angular position of the static unbalance caused by the disk on the shaft, stiffness along the principal axes of elasticity, and the non-rotating damping coefficient. The estimation is exclusively based on the displacement signals from the disk centre. Note that this estimation not only optimizes the performance of machineries, enhancing efficiency and reliability, but also enables predictive maintenance by identifying potential faults early on. To highlight the effectiveness and precision of the proposed methodology, various scenarios with different constant rotational speeds are examined, and the performance is compared to that of traditional optimization algorithms used for parameters estimation. Furthermore, the analysis accounts for the impact of noisy input data. It is important to note that the proposed work presents a proof of concept, demonstrating the effectiveness of the proposed methodology through simulation experiments in a controlled environment. The transition from simulations to real-world applications is highlighted, emphasizing the commitment to practicality. Subsequent efforts will focus on rigorous experimental validation and testing on more complex systems to enhance the approach versatility and robustness.

The main innovation of this work lies in integrating established physical knowledge, describing the fundamental dynamics of rotating shaft systems, into the neural network training process. This incorporation serves to guide the training, enhancing the robustness and reliability in the system health state parameter estimation. Furthermore, the estimation relies exclusively on raw time-domain displacements at the disk centre, minimizing the requirement for numerous sensors and simplifying the overall preprocessing steps.

The paper is organised as follows: Section 2 offers a brief overview of the necessary theoretical foundations about PINNs for parameter estimation; Section 3 shortly presents the case study and then shows in detail the implementation and the results of the PINN for the system health state characterization. Finally, Section 4 provides some concluding remarks.

2. Methodology

The proposed framework hinges upon the use of PINNs to estimate the unknown parameters characterizing the dynamics of a rotating shaft system. The innovative aspect in this methodology stems from the tailored and specific application of PINNs, addressing the challenges and requirements associated with accurately estimating health parameters

in the context of rotating shaft systems. Notably, PINNs represent deep learning tools that combine NNs with the system governing equations, and are particularly useful when data might be limited or noisy, and where the underlying physics of the problem is well understood [26].

Assume that a generic physical system is governed by the n -th order ordinary differential equations (ODEs) shown in Equation (1):

$$\mathbf{u}^{(n)} = F(t, \mathbf{u}, \mathbf{u}', \mathbf{u}'', \dots, \mathbf{u}^{(n-1)}, \boldsymbol{\lambda}) \quad (1)$$

where t refers to the system independent variable, $\mathbf{u} = [u_1(t), u_2(t), \dots, u_p(t)]$ denotes the state vector consisting of p components defined in the domain $[t_0, t_f]$, and $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_k]$ represents the vector made of the k unknown parameters describing the system state. Subsequently, considering the NN universal approximation theorem [33], an NN can be exploited to obtain an approximation $\hat{\mathbf{u}} = N(\mathbf{W}, \mathbf{b}, \boldsymbol{\lambda})$ of the state vector \mathbf{u} , such that $\hat{\mathbf{u}} \approx \mathbf{u}$. More specifically, \mathbf{W} and \mathbf{b} denote the weight and bias matrices of the NN, respectively, and their values are the result of a training process [34], as well as for the parameter vector $\boldsymbol{\lambda}$. Note that, since $\hat{\mathbf{u}}$ is a function, its derivatives concerning the independent variable t can be computed during the training process through automatic differentiation (AD) [35,36]. Then, a function \mathbf{g} outlining the approximation of Equation (1) can be defined, as reported in Equation (2):

$$\mathbf{g} = \hat{\mathbf{u}}^{(n)} - F(t, \hat{\mathbf{u}}, \hat{\mathbf{u}}', \hat{\mathbf{u}}'', \dots, \hat{\mathbf{u}}^{(n-1)}, \boldsymbol{\lambda}) \quad (2)$$

To enable the neural network to fine-tune the parameters \mathbf{W} , \mathbf{b} , and $\boldsymbol{\lambda}$ in order to (i) fulfil the underlying ODEs describing the system behaviour and (ii) to fit the available data (i.e., gathered measurements, in which the state vector \mathbf{u} is known), two different loss functions are considered, as shown in the following Equations (3) and (4).

$$\mathcal{L}_f = \frac{1}{N_e} \cdot \sum_{i=1}^{N_e} |\mathbf{g}(t_e^i)|^2 \quad (3)$$

$$\mathcal{L}_u = \frac{1}{N_e} \cdot \sum_{i=1}^{N_e} |\mathbf{u}(t_e^i) - \hat{\mathbf{u}}(t_e^i)|^2 \quad (4)$$

where \mathcal{L}_f denotes the loss for the ODEs fulfilling while \mathcal{L}_u is the loss of the observed data; t_e^i indicates the generic i -th element of the vector \mathbf{t}_e , made of N_f elements inside the domain $[t_0, t_f]$, in which \mathcal{L}_f and \mathcal{L}_u are evaluated. Specifically, the acquisition time of the measured state vector \mathbf{u} is typically employed as the vector \mathbf{t}_e . These loss functions are subsequently integrated to yield the loss term \mathcal{L} , as presented in Equation (5):

$$\mathcal{L} = \alpha \cdot \mathcal{L}_f + \beta \cdot \mathcal{L}_u \quad (5)$$

where α and β denote two coefficients employed to assign greater weight either to the contributions derived from the accessible data or those related to the system physics. Consequently, the objective of minimizing \mathcal{L} is enforced, enabling the PINN to infer the unidentified parameters that define the system dynamics. Notably, the appropriate values for the coefficients α and β are determined through an iterative trial-and-error process. A scheme showing how the PINN is trained is presented in Figure 1. The input of the PINN is represented by the generic time instant t_e^i , and its output is the corresponding approximation of the components of the measured state vector \mathbf{u} . In each training iteration, the PINN output is compared with the actual value of the components of the state vector for all the time instants t_e^i within the vector \mathbf{t}_e , resulting in the loss term \mathcal{L}_u . Simultaneously, the PINN outputs are differentiated automatically to obtain the various terms of the n -th order ODEs described in Equation (1). This process enables the derivation of the residual \mathbf{g} of the physics equation, from which the loss term \mathcal{L}_f is computed. The two losses are then

combined to form the total loss term \mathcal{L} , which is the metric to be minimized. Note that the hyperparameters of the PINN to be optimized are not only the weights W and biases b but also, and significantly, the parameter vector λ describing the system health state.

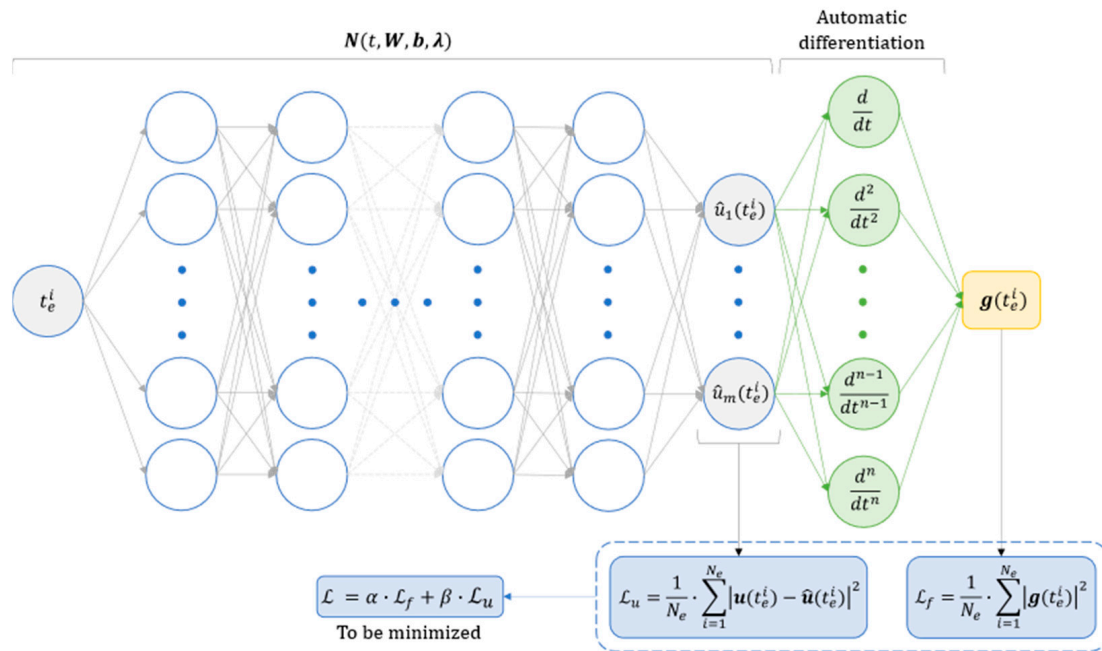


Figure 1. Scheme showing how a physics-informed neural network is trained to estimate the parameters vector λ of a system described by ODEs.

3. Case Study

3.1. Extended Jeffcott Rotor with Unknown System Parameters

The PINN approach described in the Section 2 is here applied on a numerical case study of a rotating shaft system in which the dynamics are simulated with the use of an extended Jeffcott rotor model [37]. The system, illustrated in Figure 2, consists of a 1 m long rotating shaft (i.e., $l = 1$ m) made of aluminium (Young's modulus E set to 70,000 MPa, and density ρ equal to $2700 \text{ kg}\cdot\text{m}^{-3}$) and supported at both ends. It has a circular cross section with a diameter of 20 mm, and it incorporates a disk (representing, for instance, a flywheel, fan, turbine, gear, etc.) that is mounted at distances l_1 and l_2 from the respective supports. The shaft rotates at a velocity $\Omega = \frac{d\vartheta(t)}{dt}$, in which $\vartheta(t)$ denotes the angle defined with respect to the x axis of the right-handed xyz reference frame which is fixed in space. In this reference frame, the disk lies within the xy plane, and the z axis is aligned with the line connecting the two supports. Without any loss of generality, the disk is here considered to be positioned at the midpoint of the shaft, i.e., $l_1 = l_2 = l/2$. Moreover, the disk centre of mass P is displaced from the axis of rotation, whose trace in the disk plane is identified with the point C , generating a static unbalance defined by the distance ε and the angle $\varphi = \vartheta(0)$. This unbalance causes the point C to displace from the line joining the supports leading the shaft to whirl around it.

The numerical model used to compute the rotor dynamics consists of a system of two second-order ODEs with a state vector of two components (making reference to Section 2, $n = 2$ and $p = 2$, respectively), as reported in Equations (6) and (7):

$$\begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{Bmatrix} \ddot{x}_c \\ \ddot{y}_c \end{Bmatrix} + \begin{bmatrix} c_n + c_r & 0 \\ 0 & c_n + c_r \end{bmatrix} \begin{Bmatrix} \dot{x}_c \\ \dot{y}_c \end{Bmatrix} + \begin{bmatrix} k_x & \Omega \cdot c_r \\ -\Omega \cdot c_r & k_y \end{bmatrix} \begin{Bmatrix} x_c \\ y_c \end{Bmatrix} = \begin{Bmatrix} f_x \\ f_y \end{Bmatrix} \quad (6)$$

$$\begin{Bmatrix} f_x \\ f_y \end{Bmatrix} = \begin{Bmatrix} m_d \cdot \varepsilon \cdot (\Omega^2 \cdot \cos(\Omega \cdot t + \varphi) + \dot{\Omega} \cdot \sin(\Omega \cdot t + \varphi)) \\ m_d \cdot \varepsilon \cdot (\Omega^2 \cdot \sin(\Omega \cdot t + \varphi) - \dot{\Omega} \cdot \cos(\Omega \cdot t + \varphi)) - m \cdot g \end{Bmatrix} \quad (7)$$

where x_c and y_c denote the x and y coordinates of the point C , respectively, and represent the components of the state vector (i.e., $\mathbf{u} = [x_c, y_c]$); c_n and c_r are the equivalent viscous damping terms of the stationary and rotating parts of the system, respectively; m represents the summation of the disk mass m_d and the shaft equivalent mass at the disk location m_s^e ; g indicates the gravity acceleration; k_x and k_y are the system stiffnesses along the x and y axes, respectively, that are assumed to coincide with the axes of the ellipse of elasticity, i.e., the principal axes of elasticity of the supporting structure. Notably, the rotating damping c_r is here considered to be coincident to the contribution given by the shaft material properties, thus neglecting any potential additional term, and it is computed exploiting the approximation of a linear system [37], i.e., $c_r = 2\zeta_r\sqrt{k_s \cdot m_s}$. Here, ζ_r denotes the rotational damping coefficient, that is assumed to be 0.001, m_s represents the mass of the shaft, while $k_s = \frac{48 \cdot E \cdot I}{l^3}$ denotes the shaft flexural stiffness, with I indicating the moment of inertia of the shaft cross section. The disk mass m_d is considered to be equal to 2 kg, while the shaft equivalent mass at the disk location m_s^e is computed as $m_s^e = \frac{k_s \cdot \delta}{g}$, in which $\delta = \frac{5 \cdot m_s \cdot g \cdot l^3}{384 \cdot E \cdot I}$ denotes the static displacement of the shaft at the disk location due to the shaft weight. The stiffnesses k_x and k_y are the combination of the shaft stiffness k_s and those of the supports k_b^x and k_b^y along the x and y axes, respectively. That is, $\frac{1}{k_x} = \frac{1}{k_s} + \frac{1}{k_b^x}$ and $\frac{1}{k_y} = \frac{1}{k_s} + \frac{1}{k_b^y}$. However, determining the stiffness values of the supports can be challenging for different reasons, e.g., when they have complex geometries and interactions, due to misalignments or imperfections in installation, when the support materials are not well defined or uniform, and they can change due to degradation over time [38]. A similar reasoning applies for the non-rotating damping and for the static unbalance. Hence, in this work, ε , φ , c_n , k_x , and k_y represent the components of the parameter vector λ , and their values are estimated through a PINN. Such parameters are selected because tracking their value is essential for maintaining the performance, reliability, and safety of rotating machinery.

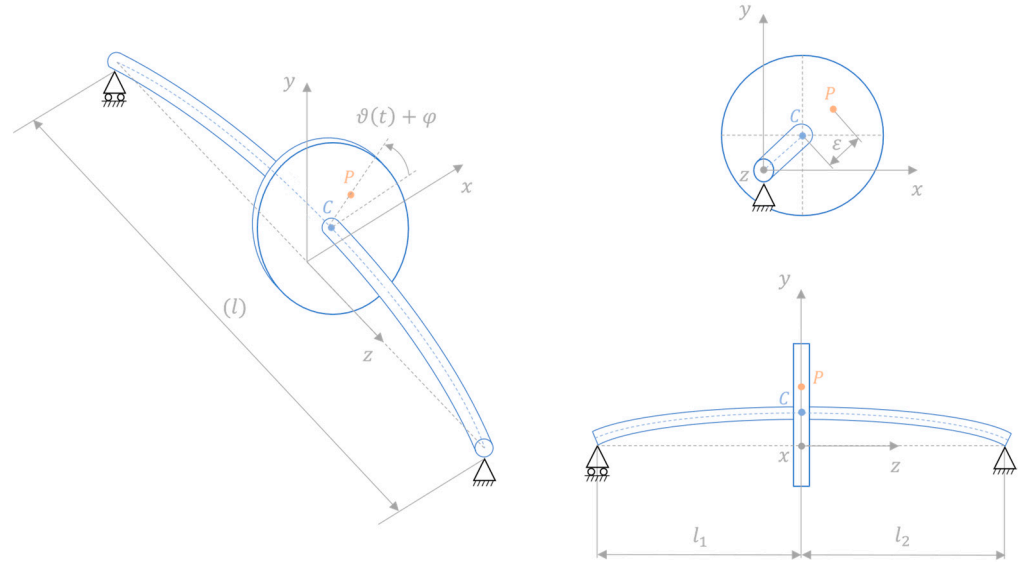


Figure 2. Scheme of the considered rotating shaft system with the disk geometrical centre (point C) and static unbalance (point P) highlighted.

Figure 3 shows the model responses x_c and y_c obtained by solving Equations (6) and (7) with a Runge-Kutta 4/5 integration method [39] in the time range $[0, 10]$ s. In the simulated scenario, a constant rotational speed of $\Omega = 30 \text{ rad} \cdot \text{s}^{-1}$ is imposed, and $k_x = 7.76 \text{ N} \cdot \text{mm}^{-1}$, $k_y = 6.71 \text{ N} \cdot \text{mm}^{-1}$, $c_n = 7.0 \times 10^{-3} \text{ N} \cdot \text{s} \cdot \text{mm}^{-1}$, $\varepsilon = 8 \text{ mm}$, $\varphi = 10 \text{ deg}$. No artificial noise is added. Figure 4 shows the scatter plot of the position of the disk centre C in the xy plane over time.

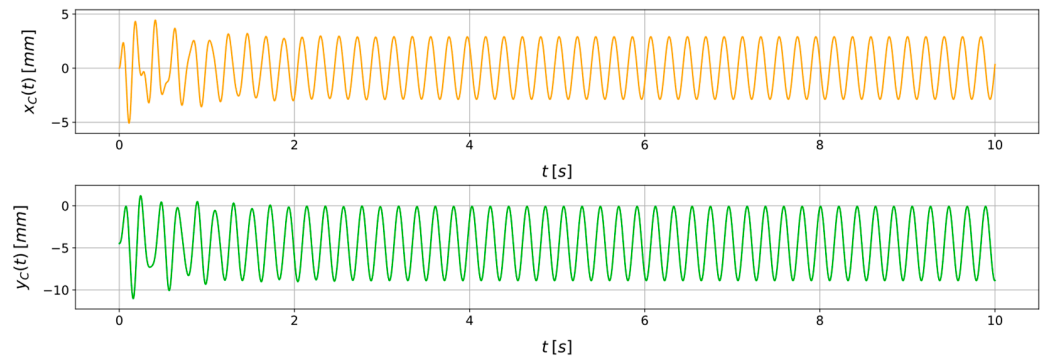


Figure 3. Example of the model responses x_C and y_C in a representative scenario with $\Omega = 30 \text{ rad}\cdot\text{s}^{-1}$, $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 7.76 \text{ N}\cdot\text{mm}^{-1}$, and $k_y = 6.71 \text{ N}\cdot\text{mm}^{-1}$, without adding artificial noise.

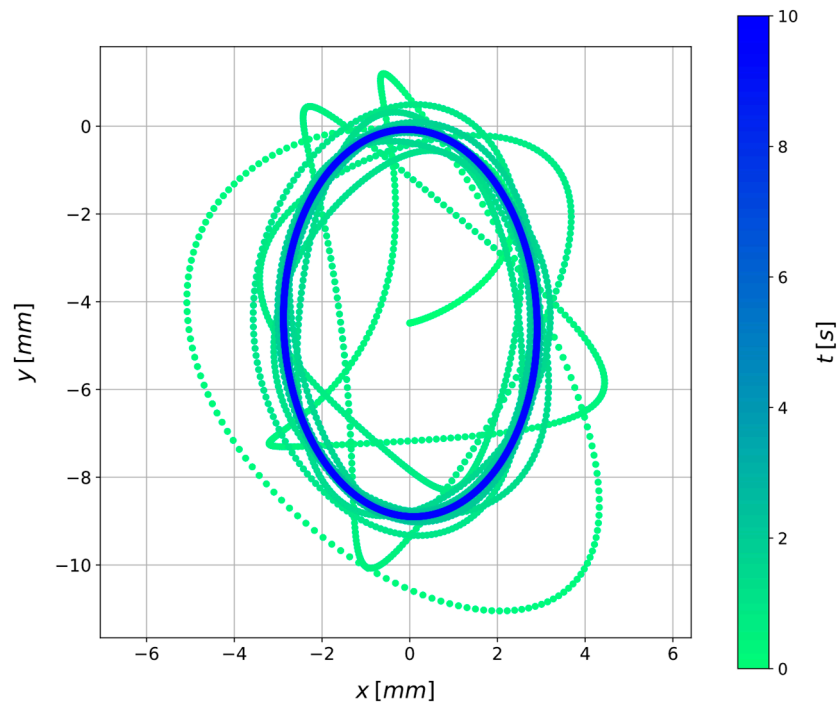


Figure 4. Visualization in the $x - y$ plane of the disk centre position (x_C, y_C) over time t for a scenario in which a constant rotational speed of $\Omega = 30 \text{ rad}\cdot\text{s}^{-1}$ is imposed and where $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 7.76 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 6.71 \text{ N}\cdot\text{mm}^{-1}$, without adding artificial noise.

3.2. System State Characterization through Physics-Informed Neural Networks

A PINN is then exploited to estimate the unknown parameters of the analysed rotating shaft system, i.e., to estimate the parameter vector $\lambda = [\varepsilon, \varphi, c_n, k_x, k_y]$. Notably, the employed NN architecture consists of 1 input neuron that takes in the generic time instant t , 1 hidden layer made of 200 neurons, and 2 output neurons to predict the values of x_C and y_C . Moreover, the hyperbolic tangent \tanh activation function [40] is used in the hidden layer, while the output layer embeds a linear activation function. The true system responses x_C and y_C required for training the PINN are numerically obtained with the Runge-Kutta 4/5 integration method in the time range $[0, 1] \text{ s}$ (i.e., $t_0 = 0 \text{ s}$ and $t_f = 1 \text{ s}$). Within this range, the true solution is sampled with a sampling frequency of 10 kHz, which means that 10,001 equally spaced points are considered in time. The same time instants are also considered for building the vector \mathbf{t}_e that is used for training (i.e., $N_f = 10,001$). It is worth noting that various sampling frequency values were examined in this study. Specifically,

the investigation covered a sampling frequency range of [1, 10] kHz with increments of 3 kHz. Due to the similarity in outcomes from the analysis, these results are omitted here for the sake of conciseness.

A representative scenario is examined to assess the efficacy of the proposed methodology. The scenario involves the imposition of a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$, alongside the following unknown parameters: $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 5.53 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 7.25 \text{ N}\cdot\text{mm}^{-1}$. No artificial noise is added to the system responses. The PINN is trained on an AMD Ryzen 9 5900HX 3.30 GHz processor using a limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) optimization algorithm [41]. The learning rate is set to $\eta = 0.01$, while both the coefficients of the loss, α and β , are fixed at 1. The training process comprises 10,000 iterations. Moreover, the arbitrary initial guess set for the unknown parameters is $\varepsilon = 0.00 \text{ mm}$, $\varphi = 0.00 \text{ deg}$, $c_n = 0.001 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 1.00 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 2.00 \text{ N}\cdot\text{mm}^{-1}$. The training process, as depicted in Figure 5, illustrates the progressive reduction of the training loss as the number of iterations increases. Instead, Figure 6 shows how the trained PINN is able to fit the available data to estimate the unknown parameters, showing a comparison between the true solution and the one obtained with the PINN. The outcome reveals that the approximation of the state vector offered by the PINN closely aligns with the actual vector, showcasing marginal disparities primarily observed in the state variable y_C during the initial time instants. Finally, the true value and the correspondent PINN estimation for all the system unknown parameters is shown in Table 1. What emerges is that all the parameters are estimated by the PINN with a remarkable accuracy, presenting the best performance in identifying the system stiffness values k_x and k_y , where the relative error in the estimation remains below 0.70%. Notably, among the parameters under estimation, the angle φ of the static unbalance and the non-rotating damping c_n prove to be the most challenging. This observation finds potential justification in the relatively subdued impact these parameters exert on the state variables compared to their counterparts. Nevertheless, even in these instances, the relative error remains constrained within 5.80%.

In order to assess the robustness of the PINN to external influences affecting input data (e.g., measurement noise, environmental vibrations, etc.), the same scenario is revisited while maintaining consistent training process conditions (i.e., algorithm used, learning rate, initial parameters guess, etc.). In this context, the input variables x_C and y_C , employed for the estimation of system parameters, are subjected to a perturbation through the introduction of supplementary numerical noise with a signal-to-noise ratio (SNR) of 20 dB. The data fitting performed by the PINN and the parameters estimation are reported in Figure 7 and Table 2, respectively. The PINN approximation of the state variables x_C and y_C appears to closely match the true solution. That is, the PINN manages to smooth out all the perturbances introduced by the added numerical noise, thus acting as a filter. The outcomes demonstrate that the PINN continues to function as a dependable tool for characterizing the health state of the system, even when confronted with external disturbances within the input data. Notably, no indications of a significant diminished algorithm performance are discernible, with all parameter estimations retaining a relative error of less than 8.00%.

Table 1. True value and the correspondent PINN estimation for all the system unknown parameters related to a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed; no noise added to the processed system responses.

Unknown Parameter	True Value	PINN Estimation	Relative Error (%)
ε [mm]	8.00	7.87	1.62
φ [deg]	10.00	9.42	5.80
c_n [$\text{N}\cdot\text{s}\cdot\text{mm}^{-1}$]	7.00×10^{-3}	6.70×10^{-3}	4.29
k_x [$\text{N}\cdot\text{mm}^{-1}$]	5.53	5.54	0.18
k_y [$\text{N}\cdot\text{mm}^{-1}$]	7.25	7.20	0.69

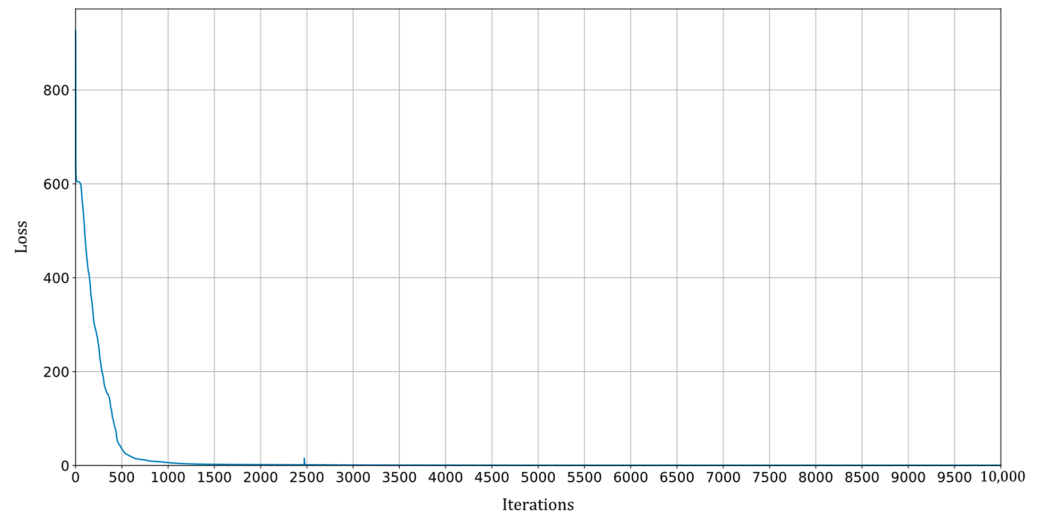


Figure 5. PINN training loss over the training iterations for a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed and where $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 5.53 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 7.25 \text{ N}\cdot\text{mm}^{-1}$, without adding artificial noise.

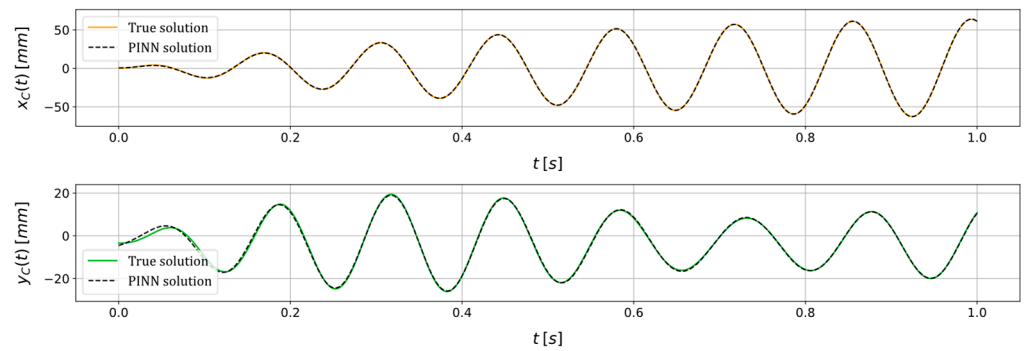


Figure 6. True and PINN solutions of the state variables x_C and y_C for a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed and where $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 5.53 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 7.25 \text{ N}\cdot\text{mm}^{-1}$, without adding artificial noise.

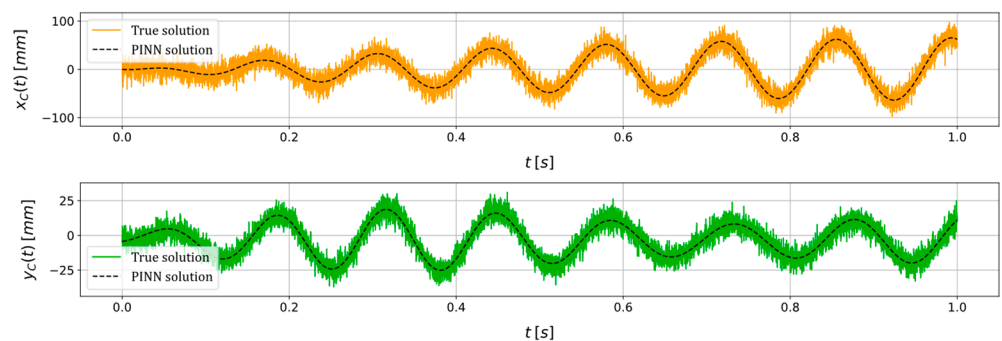


Figure 7. True and PINN solutions of the state variables x_C and y_C for a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed and where $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 5.53 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 7.25 \text{ N}\cdot\text{mm}^{-1}$, adding an artificial noise with a SNR = 20 dB.

Table 2. True value and the correspondent PINN estimation for all the system unknown parameters related to a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed; noise added to the processed system responses with an $\text{SNR} = 20 \text{ dB}$.

Unknown Parameter	True Value	PINN Estimation	Relative Error (%)
ε [mm]	8.00	7.91	1.12
φ [deg]	10.00	9.30	7.00
c_n [$\text{N}\cdot\text{s}\cdot\text{mm}^{-1}$]	7.00×10^{-3}	6.45×10^{-3}	7.86
k_x [$\text{N}\cdot\text{mm}^{-1}$]	5.53	5.54	0.18
k_y [$\text{N}\cdot\text{mm}^{-1}$]	7.25	7.29	0.55

Lastly, an additional example is presented below to assess the estimation capabilities of the PINN when fed with data pertaining to a scenario marked by distinct imposed conditions and health states. The scenario encompasses the application of a constant rotational speed of $\Omega = 35 \text{ rad}\cdot\text{s}^{-1}$, accompanied by the following unknown parameters: $\varepsilon = 12.00 \text{ mm}$, $\varphi = 20.00 \text{ deg}$, $c_n = 5.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 7.76 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 6.14 \text{ N}\cdot\text{mm}^{-1}$. Moreover, an artificial noise with an SNR of 30 dB is added to the input data, and the same training conditions of the previous scenarios are kept. The true and PINN solutions of the state variables x_C and y_C are reported in Figure 8, and the estimation of the unknown parameters is shown in Table 3. As for the previous scenario, the PINN approximation of the state variables x_C and y_C shows a strong correspondence with the true solution. The unknown parameters are satisfactorily estimated even in this scenario. The lowest relative estimation error characterizes the system stiffnesses k_x and k_y , i.e., 0.52% and 0.16%, respectively, while the non-rotating damping c_n is identified with a 9.80% error.

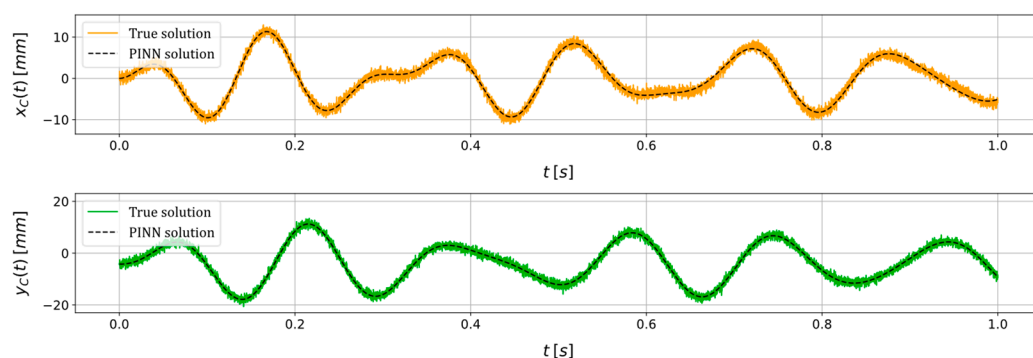


Figure 8. True and PINN solutions of the state variables x_C and y_C for a scenario in which a constant rotational speed of $\Omega = 35 \text{ rad}\cdot\text{s}^{-1}$ is imposed and where $\varepsilon = 12.00 \text{ mm}$, $\varphi = 20.00 \text{ deg}$, $c_n = 5.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 7.76 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 6.14 \text{ N}\cdot\text{mm}^{-1}$, adding an artificial noise with an $\text{SNR} = 30 \text{ dB}$.

Table 3. True value and the correspondent PINN estimation for all the system unknown parameters related to a scenario in which a constant rotational speed of $\Omega = 35 \text{ rad}\cdot\text{s}^{-1}$ is imposed; noise added to the processed system responses with an $\text{SNR} = 30 \text{ dB}$.

Unknown Parameter	True Value	PINN Estimation	Relative Error (%)
ε [mm]	12.00	11.96	0.33
φ [deg]	20.00	21.17	5.85
c_n [$\text{N}\cdot\text{s}\cdot\text{mm}^{-1}$]	5.00×10^{-3}	5.49×10^{-3}	9.80
k_x [$\text{N}\cdot\text{mm}^{-1}$]	7.76	7.72	0.52
k_y [$\text{N}\cdot\text{mm}^{-1}$]	6.14	6.13	0.16

3.3. Comparison with Traditional Optimization Algorithms for Parameters Estimation

The potentialities and limitations of the proposed framework are then identified by comparing its performance to that of traditional optimization algorithms. To this purpose,

a gradient-based optimization algorithm [42] and a genetic algorithm [43] are employed to estimate the parameter vector $\lambda = [\varepsilon, \varphi, c_n, k_x, k_y]$. The gradient-based optimization algorithm leverages the *fmincon* nonlinear solver implemented in MATLAB. The optimization problem searches the target parameters within prescribed bounds, and a step tolerance of 1×10^{-10} is used for improved performance. Instead, the genetic algorithm used for estimating the unknown parameters is based on the *ga* MATLAB function. Population size of 200 and elite count of 2 are selected through trial and error. Crossover is applied through the built-in function *crossoverscattered*, while the selected mutation function is *mutationadaptfeasible*. Optimization is stopped according to the default early stopping criteria, or when 2000 generations are reached. Regardless of the optimization algorithm employed, the loss function used to drive the optimization process involves the solution of the ODEs in Equations (6) and (7) to minimize the error between the computed and the observed displacement time history. As done for the PINN, the ODEs are solved using the Runge-Kutta 4/5 integration method.

First, the representative scenario involving the imposition of a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$, alongside the following unknown parameters: $\varepsilon = 8.00 \text{ mm}$, $\varphi = 10.00 \text{ deg}$, $c_n = 7.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 5.53 \text{ N}\cdot\text{mm}^{-1}$, and $k_y = 7.25 \text{ N}\cdot\text{mm}^{-1}$ is analysed. The results are shown in Table 4. Parameter φ is estimated with less accuracy than the PINN estimate shown in Table 1, while the accuracy is preserved for all the other variables. However, the optimization algorithms are much faster than the neural network-based framework, and the gradient-based algorithm allows for real-time estimation.

Table 4. True value and the correspondent optimization algorithms estimations related to a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed; no noise added to the processed system responses.

Unknown Parameter	True Value	Gradient-Based Method		Genetic Algorithm	
		Estimation	Relative Error (%)	Estimation	Relative Error (%)
ε [mm]	8.00	8.00	0.00	7.60	5.00
φ [deg]	10.00	0.02	99.80	20.53	105.30
c_n [$\text{N}\cdot\text{s}\cdot\text{mm}^{-1}$]	7.00×10^{-3}	6.09×10^{-3}	13.00	7.06×10^{-3}	0.86
k_x [$\text{N}\cdot\text{mm}^{-1}$]	5.53	5.66	2.35	5.39	2.53
k_y [$\text{N}\cdot\text{mm}^{-1}$]	7.25	7.36	1.52	7.25	0.00

The parameters are then estimated in the case of noise with *SNR* of 20 dB affecting the input variables x_C and y_C . The results are shown in Table 5. Similar considerations to those already reported above regarding the same scenario, but unaffected by noise, can be drawn out.

Table 5. True value and the correspondent optimization algorithms estimation related to a scenario in which a constant rotational speed of $\Omega = 45 \text{ rad}\cdot\text{s}^{-1}$ is imposed; noise added to the processed system responses with an *SNR* = 20 dB.

Unknown Parameter	True Value	Gradient-Based Method		Genetic Algorithm	
		Estimation	Relative Error (%)	Estimation	Relative Error (%)
ε [mm]	8.00	7.99	0.12	7.40	7.50
φ [deg]	10.00	0.01	99.90	1.92	80.80
c_n [$\text{N}\cdot\text{s}\cdot\text{mm}^{-1}$]	7.00×10^{-3}	6.09×10^{-3}	13.00	4.84×10^{-3}	30.86
k_x [$\text{N}\cdot\text{mm}^{-1}$]	5.53	5.66	2.35	5.66	2.35
k_y [$\text{N}\cdot\text{mm}^{-1}$]	7.25	7.35	1.38	7.30	0.69

Finally, the last scenario presented in Section 3.2 is also analysed. That is, a constant rotational speed of $\Omega = 35 \text{ rad}\cdot\text{s}^{-1}$ is considered, accompanied by the following unknown parameters: $\varepsilon = 12.00 \text{ mm}$, $\varphi = 20.00 \text{ deg}$, $c_n = 5.00 \times 10^{-3} \text{ N}\cdot\text{s}\cdot\text{mm}^{-1}$, $k_x = 7.76 \text{ N}\cdot\text{mm}^{-1}$, $k_y = 6.14 \text{ N}\cdot\text{mm}^{-1}$. An artificial noise with an *SNR* of 30 dB is added to the input data.

The results are shown in Table 6. The optimization algorithms perform similarly to the PINN-based framework in this scenario, with the advantage of allowing for real-time parameters estimation.

Table 6. True value and the correspondent optimization algorithms estimation related to a scenario in which a constant rotational speed of $\Omega = 35 \text{ rad}\cdot\text{s}^{-1}$ is imposed; noise added to the processed system responses with an $\text{SNR} = 30 \text{ dB}$.

Unknown Parameter	True Value	Gradient-Based Method		Genetic Algorithm	
		Estimation	Relative Error (%)	Estimation	Relative Error (%)
ε [mm]	12.00	11.8	1.67	10.2	15.00
φ [deg]	20.00	19.56	2.20	19.8	1.00
c_n [$\text{N}\cdot\text{s}\cdot\text{mm}^{-1}$]	5.00×10^{-3}	5.41×10^{-3}	8.20	4.98×10^{-3}	0.40
k_x [$\text{N}\cdot\text{mm}^{-1}$]	7.76	7.76	0.00	7.53	2.96
k_y [$\text{N}\cdot\text{mm}^{-1}$]	6.14	6.10	0.65	6.12	0.33

4. Conclusions

This paper has introduced a novel approach employing PINNs for estimating unknown parameters characterizing the health state of rotating shaft systems. The investigation has focused on a realistic numerical case study involving an extended Jeffcott rotor model, which has incorporated damping effects and anisotropic supports. The parameters considered have encompassed the radial and angular position of static unbalance caused by a shaft-mounted disk, stiffness values along the principal axes of elasticity, and the non-rotating damping coefficient. The estimation has relied exclusively on displacement signals from the disk centre, and various scenarios, incorporating different constant rotational speeds, have been thoroughly examined. Results have revealed the implemented PINN accuracy in estimating these parameters, demonstrating minimal relative errors even in the presence of substantial data noise. Moreover, the comparison with the estimates obtained using traditional optimization methods have revealed that PINNs slightly outperform gradient-based and genetic methods in terms of estimation accuracy, despite the longer processing time. Beyond optimizing machinery performance and enhancing efficiency and reliability, the proposed estimation method has facilitated predictive maintenance by early fault identification.

The simulation experiments outlined in this paper establish a compelling proof of concept, showcasing the effectiveness of our proposed approach within a controlled environment. It is crucial to acknowledge that, while these simulations offer valuable insights, the next step involves experimental verification to ensure the real-world applicability of our methodology. Subsequent efforts will be dedicated to conducting experimental studies on more intricate case scenarios, aiming to provide a robust validation and refinement of our proposed approach.

Author Contributions: Conceptualization, M.P.; methodology, M.P.; software, M.P.; validation, M.P.; formal analysis, M.P.; investigation, M.P.; resources, M.P.; data curation, M.P.; writing—original draft preparation, M.P.; writing—review and editing, M.P. and L.L.; visualization, M.P.; supervision, F.C.; project administration, M.P., L.L. and F.C.; funding acquisition, M.G. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Tandon, N.; Parey, A. Condition Monitoring of Rotary Machines. In *Condition Monitoring and Control for Intelligent Manufacturing*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 109–136. [CrossRef]
2. Jeong, H.; Park, S.; Woo, S.; Lee, S. Rotating Machinery Diagnostics Using Deep Learning on Orbit Plot Images. *Procedia Manuf.* **2016**, *5*, 1107–1118. [CrossRef]
3. Tiboni, M.; Remino, C.; Bussola, R.; Amici, C. A Review on Vibration-Based Condition Monitoring of Rotating Machinery. *Appl. Sci.* **2022**, *12*, 972. [CrossRef]
4. Silva, D.; Mendes, J.C.; Pereira, A.B.; Gégot, F.; Alves, L.N. Measuring Torque and Temperature in a Rotating Shaft Using Commercial SAW Sensors. *Sensors* **2017**, *17*, 1547. [CrossRef] [PubMed]
5. Nandi, S.; Toliyat, H.A.; Li, X. Condition Monitoring and Fault Diagnosis of Electrical Motors—A Review. *IEEE Trans. Energy Convers.* **2005**, *20*, 719–729. [CrossRef]
6. Farrar, C.R.; Worden, K. An Introduction to Structural Health Monitoring. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2006**, *365*, 303–315. [CrossRef] [PubMed]
7. Hameed, Z.; Hong, Y.S.; Cho, Y.M.; Ahn, S.H.; Song, C.K. Condition Monitoring and Fault Detection of Wind Turbines and Related Algorithms: A Review. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1–39. [CrossRef]
8. Zhou, H.; Huang, X.; Wen, G.; Lei, Z.; Dong, S.; Zhang, P.; Chen, X. Construction of Health Indicators for Condition Monitoring of Rotating Machinery: A Review of the Research. *Expert Syst. Appl.* **2022**, *203*, 117297. [CrossRef]
9. Bogue, R. Sensors for Condition Monitoring: A Review of Technologies and Applications. *Sens. Rev.* **2013**, *33*, 295–299. [CrossRef]
10. Sohn, H. Effects of Environmental and Operational Variability on Structural Health Monitoring. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2006**, *365*, 539–560. [CrossRef]
11. Parziale, M.; Lomazzi, L.; Giglio, M.; Cadini, F. Vibration-Based Structural Health Monitoring Exploiting a Combination of Convolutional Neural Networks and Autoencoders for Temperature Effects Neutralization. *Struct. Control Health Monit.* **2022**, *29*, e3076. [CrossRef]
12. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep Learning and Its Applications to Machine Health Monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
13. Fu, J.; Chu, J.; Guo, P.; Chen, Z. Condition Monitoring of Wind Turbine Gearbox Bearing Based on Deep Learning Model. *IEEE Access* **2019**, *7*, 57078–57087. [CrossRef]
14. O’Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. *Int. J. Res. Appl. Sci. Eng. Technol.* **2015**, *10*, 943–947. [CrossRef]
15. Schmidt, R.M. Recurrent Neural Networks (RNNs): A Gentle Introduction and Overview. *arXiv* **2019**, arXiv:1912.05911.
16. Souza, R.M.; Nascimento, E.G.S.; Miranda, U.A.; Silva, W.J.D.; Lepikson, H.A. Deep Learning for Diagnosis and Classification of Faults in Industrial Rotating Machinery. *Comput. Ind. Eng.* **2021**, *153*, 107060. [CrossRef]
17. Wisal, M.; Oh, K.Y. A New Deep Learning Framework for Imbalance Detection of a Rotating Shaft. *Sensors* **2023**, *23*, 7141. [CrossRef] [PubMed]
18. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
19. Chen, Y.; Rao, M.; Feng, K.; Niu, G. Modified Varying Index Coefficient Autoregression Model for Representation of the Nonstationary Vibration from a Planetary Gearbox. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3511812. [CrossRef]
20. Han, T.; Xie, W.; Pei, Z. Semi-Supervised Adversarial Discriminative Learning Approach for Intelligent Fault Diagnosis of Wind Turbine. *Inf. Sci.* **2023**, *648*, 119496. [CrossRef]
21. Dosilovic, F.K.; Brcic, M.; Hlupic, N. Explainable Artificial Intelligence: A Survey. In Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 21–25 May 2018; pp. 210–215. [CrossRef]
22. Lomazzi, L.; Fabiano, S.; Parziale, M.; Giglio, M.; Cadini, F. On the Explainability of Convolutional Neural Networks Processing Ultrasonic Guided Waves for Damage Diagnosis. *Mech. Syst. Signal Process.* **2023**, *183*, 109642. [CrossRef]
23. Parziale, M.; Lomazzi, L.; Giglio, M.; Cadini, F. Transmissibility Functions-Based Structural Damage Assessment with the Use of Explainable Convolutional Neural Networks. In *International Conference on Experimental Vibration Analysis for Civil Engineering Structures*; Springer Nature: Cham, Switzerland, 2023; pp. 540–549. [CrossRef]
24. Parziale, M.; Henrique Silva, P.; Giglio, M.; Cadini, F. Explainability of Convolutional Neural Networks for Damage Diagnosis Using Transmissibility Functions. Available online: <http://dx.doi.org/10.2139/ssrn.4545333> (accessed on 27 December 2023).
25. Parziale, M.; Yeung, Y.F.; Youcef-Toumi, K.; Giglio, M.; Cadini, F. Anomaly Characterization for the Condition Monitoring of Rotating Shafts Exploiting Data Fusion and Explainable Convolutional Neural Networks. Available online: <http://dx.doi.org/10.2139/ssrn.4634978> (accessed on 27 December 2023).
26. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
27. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics Informed Deep Learning (Part I): Data-Driven Solutions of Nonlinear Partial Differential Equations. *arXiv* **2017**, arXiv:1711.10561.
28. Haghghat, E.; Raissi, M.; Moure, A.; Gomez, H.; Juanes, R. A Physics-Informed Deep Learning Framework for Inversion and Surrogate Modeling in Solid Mechanics. *Comput. Methods Appl. Mech. Eng.* **2021**, *379*, 113741. [CrossRef]

29. Chen, Y.; Rao, M.; Feng, K.; Zuo, M.J. Physics-Informed LSTM Hyperparameters Selection for Gearbox Fault Detection. *Mech. Syst. Signal Process.* **2022**, *171*, 108907. [[CrossRef](#)]
30. Shen, S.; Lu, H.; Sadoughi, M.; Hu, C.; Nemani, V.; Thelen, A.; Webster, K.; Darr, M.; Sidon, J.; Kenny, S. A Physics-Informed Deep Learning Approach for Bearing Fault Detection. *Eng. Appl. Artif. Intell.* **2021**, *103*, 104295. [[CrossRef](#)]
31. Garpelli, L.N.; Alves, D.S.; Cavalca, K.L.; de Castro, H.F. Physics-Guided Neural Networks Applied in Rotor Unbalance Problems. *Struct. Health Monit.* **2023**, *22*, 4117–4130. [[CrossRef](#)]
32. Deng, W.; Nguyen, K.T.P.; Medjaher, K.; Gogu, C.; Morio, J. Rotor Dynamics Informed Deep Learning for Detection, Identification, and Localization of Shaft Crack and Unbalance Defects. *Adv. Eng. Inform.* **2023**, *58*, 102128. [[CrossRef](#)]
33. Cybenko, G. Approximation by Superpositions of a Sigmoidal Function. *Math. Control. Signals Syst.* **1989**, *2*, 303–314. [[CrossRef](#)]
34. Wang, S.-C. Artificial Neural Network. In *Interdisciplinary Computing in Java Programming*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 81–100. [[CrossRef](#)]
35. Margossian, C.C.; Charles Margossian, C.C. A Review of Automatic Differentiation and Its Efficient Implementation. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2019**, *9*, e1305. [[CrossRef](#)]
36. Al Seyab, R.K.; Cao, Y. Nonlinear System Identification for Predictive Control Using Continuous Time Recurrent Neural Networks and Automatic Differentiation. *J. Process Control* **2008**, *18*, 568–581. [[CrossRef](#)]
37. Genta, G.; Keith, R.H. Vibration Dynamics and Control. *Noise Control Eng. J.* **2009**, *57*, 156. [[CrossRef](#)]
38. Cerrada, M.; Sánchez, R.V.; Li, C.; Pacheco, F.; Cabrera, D.; Valente de Oliveira, J.; Vásquez, R.E. A Review on Data-Driven Fault Severity Assessment in Rolling Bearings. *Mech. Syst. Signal Process.* **2018**, *99*, 169–196. [[CrossRef](#)]
39. Bogacki, P.; Shampine, L.F. An Efficient Runge-Kutta (4,5) Pair. *Comput. Math. Appl.* **1996**, *32*, 15–28. [[CrossRef](#)]
40. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
41. Liu, D.C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Math. Program.* **1989**, *45*, 503–528. [[CrossRef](#)]
42. Bengio, Y. Gradient-Based Optimization of Hyperparameters. *Neural Comput.* **2000**, *12*, 1889–1900. [[CrossRef](#)]
43. Elsayed, S.M.; Sarker, R.A.; Essam, D.L. A New Genetic Algorithm for Solving Optimization Problems. *Eng. Appl. Artif. Intell.* **2014**, *27*, 57–69. [[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.