MEMS Gyroscope Temperature Compensation Based on Improved Complete Ensemble Empirical Mode Decomposition and Optimized Extreme Learning Machine

Zhihao Zhang [1], Jintao Zhang [1], Xiaohan Zhu [2], Yanchao Ren [3], Jingfeng Yu [3] and Huiliang Cao [4],*

1   Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China; 18890037230@163.com (Z.Z.)
2   College of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China
3   Quanzhou Yunjian Measurement Control and Perception Technology Innovation Research Institute, Quanzhou 362000, China; yujingfeng@yjqz.org (J.Y.)
4   Key Laboratory of Instrumentation Science & Dynamic Measurement, Ministry of Education, North University of China, Taiyuan 030051, China
*   Correspondence: caohuiliang@nuc.edu.cn

Abstract: Herein, we investigate the temperature compensation for a dual-mass MEMS gyroscope. After introducing and simulating the dual-mass MEMS gyroscope's working modes, we propose a hybrid algorithm for temperature compensation relying on improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN), sample entropy, time–frequency peak filtering, non-dominated sorting genetic algorithm-II (NSGA II) and extreme learning machine. Firstly, we use ICEEMDAN to decompose the gyroscope's output signal, and then we use sample entropy to classify the decomposed signals. For noise segments and mixed segments with different levels of noise, we use time–frequency peak filtering with different window lengths to achieve a trade-off between noise removal and signal retention. For the feature segment with temperature drift, we build a compensation model using extreme learning machine. To improve the compensation accuracy, NSGA II is used to optimize extreme learning machine, with the prediction error and the 2-norm of the output-layer connection weight as the optimization objectives. Enormous simulation experiments prove the excellent performance of our proposed scheme, which can achieve trade-offs in signal decomposition, classification, denoising and compensation. The improvement in the compensated gyroscope's output signal is analyzed based on Allen variance; its angle random walk is decreased from $0.531076°/h/\sqrt{Hz}$ to $6.65894 \times 10^{-3}°/h/\sqrt{Hz}$ and its bias stability is decreased from $32.7364°/h$ to $0.259247°/h$.

Keywords: dual-mass MEMS gyroscope; temperature compensation; improved complete ensemble empirical mode decomposition with adaptive noise; sample entropy; time–frequency peak filtering; non-dominated sorting genetic algorithm-II; extreme learning machine

1. Introduction

As a very important micro inertial device, the MEMS gyroscope is widely used in aerospace, intelligent electronics, virtual reality and other fields [1–4]. It can measure the angular velocity of objects and provide data support for navigation, attitude control and other applications. However, a temperature change in the internal thermal-sensitive material leads to the temperature drift phenomenon, which affects the stability and accuracy of its output signal [5]. Temperature compensation can offset the effect of temperature on the MEMS gyroscope and can improve its accuracy and long-term stability [6]. The research and development of temperature compensation methods is one of the current research hotspots for MEMS gyroscopes.

The compensation methods for MEMS gyroscopes are broadly categorized into two approaches: hardware compensation and software compensation. These methods aim to

mitigate the detrimental impacts of temperature fluctuations on gyroscope performance, ultimately enhancing the gyroscope's accuracy and stability across diverse operational scenarios. Hardware compensation methods primarily entail the design and integration of physical structures or components within the gyroscope system to counterbalance temperature-induced errors. This approach often includes embedding temperature sensors directly into the gyroscope package to monitor ambient temperature changes. Through measuring these fluctuations, the sensors offer feedback that can be used for adjusting the gyroscope's operational parameters, such as deviation and scale factors, in real time to counteract temperature effects [7]. Furthermore, innovative structural designs, such as the design of an I-shaped bulk acoustic resonator, are leveraged to enhance temperature stability and enable dynamic frequency adjustments across a broad spectrum [8]. Tao et al. explored a dual closed-loop control strategy featuring differential modulation for resonant integrated optic gyroscopes, enhancing detection precision and minimizing noise interference, which significantly improves gyroscope reliability and reduces control inaccuracies across varying conditions [9]. Reference [10] introduced a MEMS gyroscope with a unique design enhancing resilience to manufacturing and environmental changes. It employs a single-DOF drive and a two-DOF sensing mechanism, enabling independent frequency and bandwidth adjustments, thus mitigating previous limitations in gyroscopic robustness. In [11], a real-time temperature self-sensing hardware compensation algorithm was designed for high-Q MEMS gyroscopes, which greatly improves the measurement accuracy of gyroscopes. Although hardware compensation methods are more accurate and reliable, such methods have a long development cycle, have limited adaptability, and are not easy to promote. On the other hand, the software compensation method based on the use of various kinds of artificial intelligence algorithms to learn the complex temperature–error relationship and provide adaptive compensation solutions has high flexibility and adaptability and can be easily realized and optimized through algorithm adjustment [12]. The gyroscope's output is often mixed with noise and temperature drift, which need to be dealt with separately. The software compensation algorithms can be divided into two architecture types: serial and parallel. In a serial architecture, the signal is first denoised, and then temperature compensation is performed, which leads to the loss of useful signals. On the contrary, a parallel architecture decomposes the signal first and then selects the corresponding processing means for different signals [13]. Software compensation algorithms usually rely on adaptive signal decomposition methods, denoising methods, neural networks, etc.

The parallel processing architecture carries out noise reduction and temperature compensation at the same time. The most commonly used signal decomposition algorithms include empirical mode decomposition [14], local mean decomposition [15] and variational mode decomposition [16]. Among them, empirical mode decomposition and local mean decomposition suffer from the interference of mode aliases, while variational mode decomposition relies on the researcher's experience in setting reasonable decomposition parameters, which have a great impact on the decomposition results. Based on empirical mode decomposition, researchers have introduced advanced techniques such as ensemble empirical mode decomposition [17] and complete ensemble empirical mode decomposition with adaptive noise [18], which integrate noise to facilitate decomposition and alleviate modal mixing. Furthermore, Colominas et al. proposed a refined approach known as improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN). Through the incorporation of specialized white noise during decomposition, experiments have demonstrated that ICEEMDAN exhibits superior decomposition performance and versatility, effectively mitigating mode aliases [19]. Consequently, ICEEMDAN is employed in this study to decompose the gyroscope output signal for subsequent analysis. The mainstream denoising algorithms include wavelet transform, Kalman filtering, Fourier transform, etc., all of which have different degrees of drawbacks. Fourier transform lacks time-domain localization ability and is not suitable for non-stationary signals. Kalman filtering involves matrix operations, leading to high computational complexity and sig-

nificant signal distortion. Although the wavelet threshold denoising algorithm proposed by Johnstone et al. has excellent denoising performance, the selection of parameters is challenging [20,21]. Based on instantaneous frequency estimation of the Wigner–Ville distribution, time–frequency peak filtering was proposed by Boashash et al. It features high time–frequency positioning accuracy, high computational efficiency, and strong noise suppression. In addition, its denoising capability can be adjusted by modifying the window length [22]. In addition, high-precision compensation models are key in parallel temperature compensation algorithms, and various artificial intelligence technologies can establish prediction models between the temperature and the gyroscope's output. Shen et al. fit a compensation model for temperature drift based on a genetic-algorithm-optimized Elman neural network [23]. Song et al. introduced a novel approach that merges the artificial fish swarm algorithm with a back-propagation neural network, aiming to enhance the accuracy of fiber optic gyroscopes' output [24]. Li et al. proposed a novel temperature error model based on a radial basis function neural network, improved with particle swarm optimization and regularization methods, to enhance the accuracy and environmental adaptability of the ring laser gyroscope [25]. In addition, support vector regression, wavelet neural networks, fuzzy control theory and so on are often used to build predictive models. Among them, extreme learning machine is widely used because of its simple structure and superior generalization ability [26–28]. Since the input weights of extreme learning machine are generated randomly, the instability of the output matrix of the hidden layer results in large connection weights of the output layer, thus increasing the structural risk [29,30].

Based on the above analysis, this study examines the temperature compensation problem for a dual-mass MEMS gyroscope. We introduce the working principle and working mode of the gyroscope, and we formulate a temperature compensation experiment. To solve the mixing problem of noise and temperature drift in the gyroscope's output signal, we propose an innovative parallel temperature compensation and noise reduction scheme. Firstly, the output signal of the gyroscope is decomposed by improved complete ensemble empirical mode decomposition with adaptive noise, which has a high identification rate. Then, the sample entropy (SE) is used to classify a series of intrinsic mode functions, which can be divided into the noise segment, mixed segment and feature segment. For the noise segment and the mixed segment with different noise levels, the time–frequency peak filtering (TFPF) method with different window lengths is used to achieve a compromise between noise removal and signal retention. For the feature segment with temperature drift, we use non-dominated sorting genetic algorithm-II (NSGA II) to optimize extreme learning machine (ELM) with the prediction error and the 2-norm of the output-layer connection weight as optimization objectives to establish a high-precision temperature compensation model. Finally, the processed components are reconstructed to obtain the final output signal. Through comparative analysis of the experimental results, we prove that the proposed algorithm has excellent performance and effect in gyroscope temperature compensation.

The rest of this paper is organized as follows: Section 2 introduces the working principle of our developed dual-mass MEMS gyroscope. Our proposed algorithm is presented in Section 3. In Section 4, the temperature experiment and algorithm verification are described. Section 5 presents the conclusions.

2. Introduction of the Dual-Mass MEMS Gyroscope

In this article, a dual-mass MEMS gyroscope [5] is introduced and utilized for experiments, as illustrated in Figure 1. The gyroscope operates in drive mode and sense mode. The drive mode consists of components such as a driving comb and a driving spring, facilitating movement along the X-axis. Conversely, the sense mode, comprising the sense comb and sense spring, enables movement along the Y-axis. The masses shown in Figure 1 oscillate along the negative direction of the X-axis in both modes. Additionally, to prevent displacement coupling, the two modes of the gyroscope are intentionally isolated. When angular velocity Ωz is applied around the Z-axis, the resulting Coriolis force from

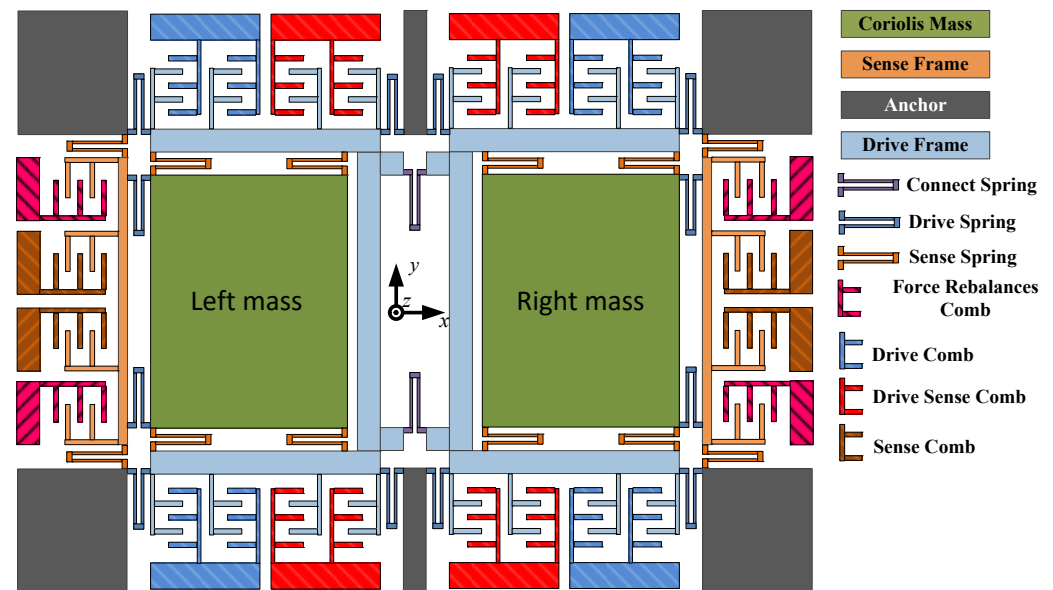the vibrating mass is transmitted along the Y-axis to the frame and monitored through an electrical circuit.

Figure 1. Structure of the dual-mass MEMS gyroscope.

The gyroscope functions on the tuning fork principle, featuring a U-type connecting spring linking the dual-drive mass block and a drive spring connecting the dual-sense mass block. To examine its operational modes, Ansys software (Ansys 2022 R1) was employed for the simulation, as depicted in Figure 2. Notably, there is a significant frequency gap of over 1000 Hz between the first and fourth operational modes, with the fourth mode exhibiting a quality factor exceeding 2000, identifying it as the gyroscope's drive mode. Figure 2a–d sequentially present the simulation results for the gyroscope's operational modes. The first mode, the drive in-phase mode, involves the gyroscope's double mass vibrating consistently with the X-axis. Following this, the second mode, the sensing in-phase mode, sees both masses vibrating simultaneously along the Y-axis direction. In contrast, the third mode, the sensing anti-phase mode, exhibits oscillations in the Y-axis contrary to the direction of the two mass blocks. Lastly, the fourth mode, the drive anti-phase mode, demonstrates oscillations of the gyroscope's two mass blocks counter to the Y-axis. Additionally, the resonant frequencies for these four modes are observed at 2623 Hz, 3342 Hz, 3468 Hz and 3484 Hz. Remarkably, the fourth mode acts not only as the drive anti-phase mode but also as the primary driving mode. From this analysis, it is evident that the gyroscope's two mass blocks have two degrees of freedom, while the drive mode and frame each exhibit only a single degree of freedom.

In Figure 3, the drive comb measures the displacement of the drive frame $x(t)$ using a split amplifier. To achieve the required phase alignment with the signal $V_{dac}Sin(\omega_d t)$, a 90° phase delay is applied to the signal. Subsequently, V_{dac} is extracted via a full-wave rectifier and low-pass filter and concurrently compared to the reference voltage $V_{ref.}$. A control signal is generated when the comparator's output passes through the integrator controller. This control signal drives the DC signal V_{DC} to accumulate into $V_{dac}Sin(\omega_d t)$, thereby exciting the drive mode. The left and right sensitive mass blocks' motion signals are captured by a differential detection amplifier. Afterward, a second differential amplifier processes the output signal to produce the total sense motion signal, V_{stotal}. This signal is then demodulated with $V_{dac}Sin(\omega_d t)$ and subsequently filtered using a low-pass filter to obtain the ultimate sensitive motion signal.
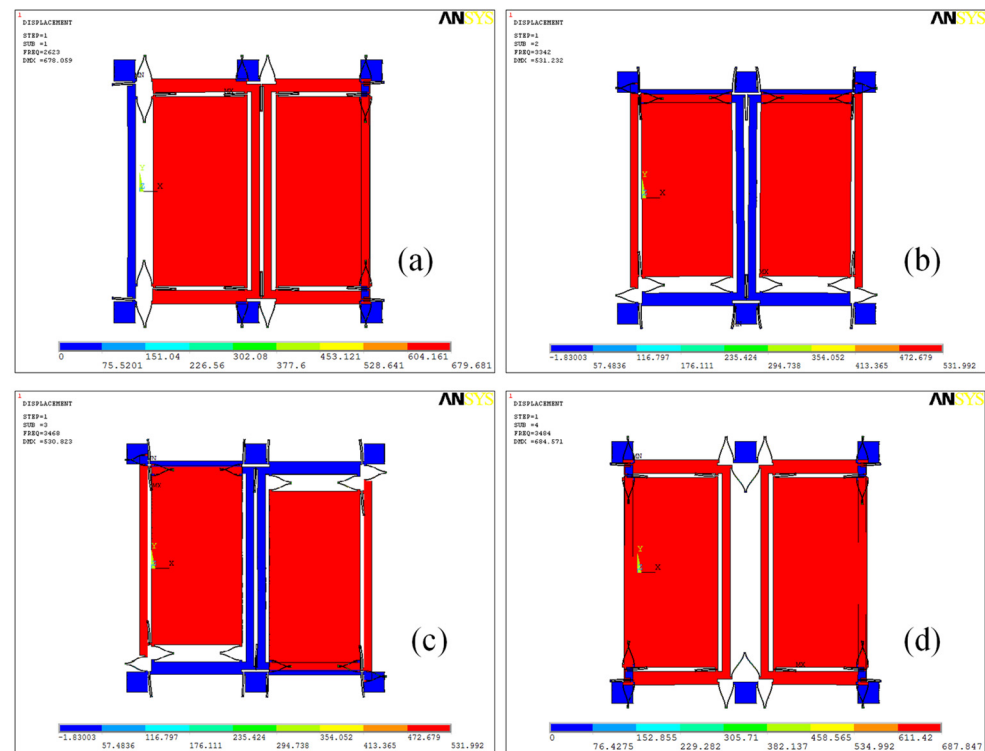
Figure 2. The gyroscope's four operational modes. Subfigures (**a**–**d**) represent each mode, with their corresponding frequencies denoted as $\omega_1 = 2623 \times 2\pi$ rad/s, $\omega_2 = 3342 \times 2\pi$ rad/s, $\omega_3 = 3468 \times 2\pi$ rad/s and $\omega_4 = 3484 \times 2\pi$ rad/s.
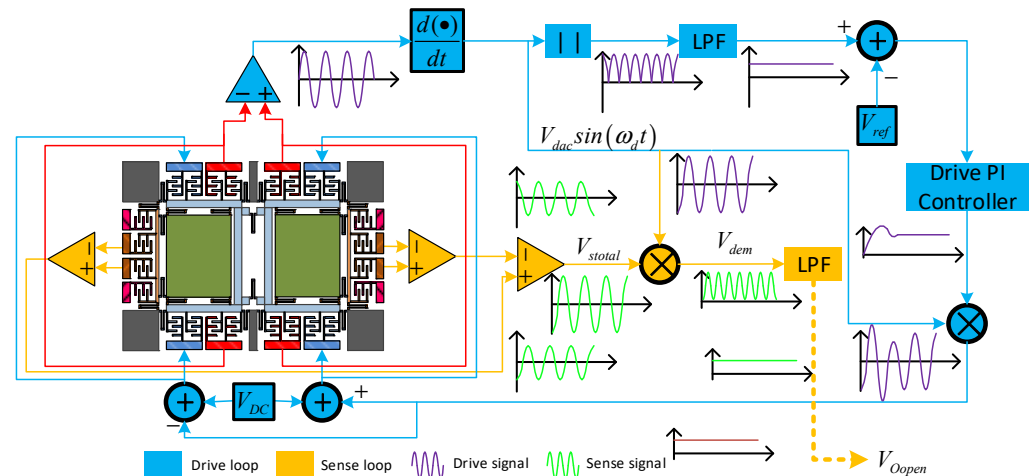

Figure 3. A description of the gyroscope's monitoring system.

3. Algorithms and Models

3.1. Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise

EMD is a classical algorithm for complex signal analysis and processing. EMD works by decomposing the signal into some amplitude–frequency modulation functions and one monotonic trend signal. Although EMD has good effects in the analysis of complex signals, it leads to the phenomenon of mode mixing due to its local characteristics. To solve this challenge, EEMD was proposed to mitigate the mode mixing phenomenon by incorporating noise into the decomposition process. Subsequently, CEEMDAN based on adaptive noise was proposed, which is an important improvement on EEMD and solves the influence of auxiliary noise on the mode number. However, CEEMDAN still has deficiencies compared to EEMD, such as residual noise in the modes and signal characteristics appearing after

the false mode. The above problems were solved through ICEEMDAN. Different from the previous methods, which add Gaussian white noise in the auxiliary decomposition, ICEEMDAN uses special white noise $E_k[w^{(i)}]$ in the auxiliary decomposition; the noise originates from the kth component of white Gaussian noise, which undergoes decomposition via EMD. A brief introduction of ICEEMDAN follows [31,32].

Before describing the algorithm steps, several operation symbols are first defined. $M(.)$ is the operator that generates the local mean value, and the kth IMF value after EMD decomposition can be obtained through the $E_k(.)$ operator. Obviously, $E_1(x) = x - M(x)$.

The decomposition process proceeds as follows:

Step 1. Adding certain noise. In Equation (1), $w^{(i)}$ represents the ith added noise, and β_0 is the standard deviation of adding noise.

$$\begin{cases} x^{(1)} = x + \beta_0 E_1\left[w^{(1)}\right] \\ x^{(2)} = x + \beta_0 E_1\left[w^{(2)}\right] \\ \quad\quad \vdots \\ x^{(i)} = x + \beta_0 E_1\left[w^{(i)}\right] \end{cases}. \tag{1}$$

Step 2. EMD calculates the local mean for each $x^{(i)}$, and the mean of their local means is taken as the first residual r_1; then, we can obtain the first IMF value, c_1:

$$\begin{cases} r_1 = \frac{1}{I}\sum_{i=1}^{I} M\left[x^{(i)}\right] \\ c_1 = x - r_1 \end{cases}. \tag{2}$$

Step 3. The second IMF value, c_2, is calculated immediately:

$$\begin{cases} c_2 = r_1 - r_2 \\ r_2 = \frac{1}{I}\sum_{i=1}^{I} M\left\{r_1 + \beta_1 E_2\left[x^{(i)}\right]\right\} \end{cases}. \tag{3}$$

Step 4. Similarly, k IMF values are calculated according to the above steps, where

$$\begin{cases} c_k = r_{k-1} - r_k, k = 2, 3, \ldots, N \\ r_k = \frac{1}{I}\sum_{i=1}^{I} M\left\{r_{k-1} + \beta_{k-1} E_k\left[x^{(i)}\right]\right\} \end{cases}. \tag{4}$$

After obtaining k modes and adding the residuals, the original signal can be rewritten as

$$x = \sum_{i=1}^{k} c_k + r_k. \tag{5}$$

ICEEMDAN is used to decompose the gyroscope's output into a series of intrinsic mode functions (IMFs). Our proposed hybrid algorithm needs to classify the obtained IMFs to identify different components; thus, it uses SE as a decision rule to distinguish these IMFs.

3.2. Sample Entropy

SE is an effective measurement method that can measure the regularity and complexity of complex signals by calculating the possibility of generating different patterns in the sequence [33]. The SE is not dependent on the data length and has high detection accuracy. A stable estimate can be obtained with a relatively short data sequence. The principle of sample entropy is as follows:

Step. 1 Suppose that time series h is composed of n data, that is, $h = \{h(1), h(2), \ldots, h(n)\}$, and then construct a m-dimensional vector based on the original signal, which is

$H_m(1), \cdots, H_m(n-m+1)$, where $H_m(i) = \{h(i), h(i+1), \cdots, h(i+m-1)\}$, $1 \le i \le n - m + 1$.

Step. 2 The distance d_{ij} between $H_m(i)$ and $H_m(j)$ is

$$d_{ij} = d[h(i), h(j)] = \max_{l=0,...,m-1} |h(i+l) - h(j+l)|, \tag{6}$$

where $j = 1, 2, \ldots, n - m$, and $j \ne i$.

Step. 3 For a given $H_m(i)$, compute the distances between $H_m(i)$ and $H_m(j)$ and calculate the number C_i of them that are less than or equal to p:

$$C_i^m(p) = \frac{1}{n-m-1} C_i. \tag{7}$$

Step. 4 $C^m(p)$ is defined as

$$C^m(p) = \frac{\sum_{i=1}^{n-m} C_i^m(p)}{n-m+1}. \tag{8}$$

Step. 5 When the dimension of the sequence is increased from m to $m + 1$, compute the distances between $H_{m+1}(i)$ and $H_{m+1}(j)$ and count the number D_i of them that are less than or equal to p:

$$D_i^m(p) = \frac{1}{n-m+1} D_i. \tag{9}$$

Step. 6 Then, $D^{(m)}(p)$ can be obtained:

$$D^m(p) = \frac{\sum_{i=1}^{n-m} D_i^m(p)}{n-m+1}. \tag{10}$$

In the above definition, under the condition of similar tolerance p, $C^m(p)$ and $D^{(m)}(p)$ denote the probability of two sequences matching either m or $m + 1$ points individually. Subsequently, SE is defined as

$$SE(m, p) = -\lim_{N \to \infty} \{\ln D^m(p)/C^m(p)\}. \tag{11}$$

Equation (11) can be rewritten into Equation (12) if n is a finite number.

$$SE(m, p) = -\ln \left[\frac{D^m(p)}{C^m(p)} \right]. \tag{12}$$

The smaller the value of SE, the higher the similarity of the signal sequence, which means that the sequence is regular and can be considered a useful signal, and vice versa.

3.3. Time–Frequency Peak Filtering

TFPF stands as a pivotal noise elimination technology. Renowned for its prowess in extracting meaningful signals amidst noisy environments, it finds extensive applications across various engineering domains. Operating primarily on the principles of WVD and instantaneous frequency estimation theory, the TFPF algorithm serves to filter and denoise signals. Despite the widespread utilization of WVD in engineering, owing to its commendable time–frequency focusing capabilities, its efficacy diminishes when processing multi-component signals due to the emergence of cross terms, which compromise its time–frequency resolution. To address the challenges posed by cross terms in TFPF, we integrate the pseudo-Wiener–Ville distribution (PWVD). Essentially, TFPF involves encoding the noisy signal to convert it into an analytic signal of the instantaneous frequency, which

enables an estimation of the effective signal through instantaneous frequency estimation. Let us consider a signal $y(t)$ contaminated with noise:

$$y(t) = x(t) + n(t), \tag{13}$$

where $x(t)$ represents the useful signal in $y(t)$ and $n(t)$ represents the noise in $y(t)$. The steps for denoising $y(t)$ by TFPF are as follows [34]:

Step 1. By subjecting the signal $y(t)$, which incorporates noise, to frequency modulation, the analytic signal $z(t)$ is derived:

$$z(t) = e^{j2\pi\mu \int_0^t y(\lambda)d\lambda}, \tag{14}$$

where μ is the frequency modulation index.

Step 2. The spectrum of the PWVD for $z(t)$ is

$$PW_2(t,f) = \int_{-\infty}^{\infty} h(\tau)z\left(t+\frac{\tau}{2}\right)z^*\left(t-\frac{\tau}{2}\right)e^{-j2\pi ft}d\tau. \tag{15}$$

where t represents time, τ is the integral variable, f denotes frequency, z^* represents the conjugate operator of Z and $h(\tau)$ is the window function.

Step 3. We compute the peak value of the PWVD spectrum of the analytic signal to estimate its instantaneous frequency based on the maximum likelihood estimation principle. Consequently, the amplitude of the original effective signal is estimated:

$$f_z(t) = \frac{\text{argmax}[PW_z(t,f)]}{\mu}. \tag{16}$$

The window length of TFPF determines its denoising ability. The denoising effect of long-window TFPF is obvious but will cause signal distortion. On the contrary, the denoising ability of short-window TFPF is poor.

3.4. Extreme Learning Machine

ELM is a feedforward neural network characterized by a single hidden layer. Its key innovation lies in the random generation of connection weights and biases specifically for the hidden layer, and no adjustment is needed after setting [28]. Compared to most neural networks, which need constant adjustment of their weights and thresholds, this method reduces the amount of computation by half. In addition, the output-layer connection weight β is determined once by solving the equations without iterative adjustment. Studies show that ELM has a high speed while ensuring learning accuracy [29]. The ELM network structure is given in Figure 4, and its theory is introduced as follows [30]:
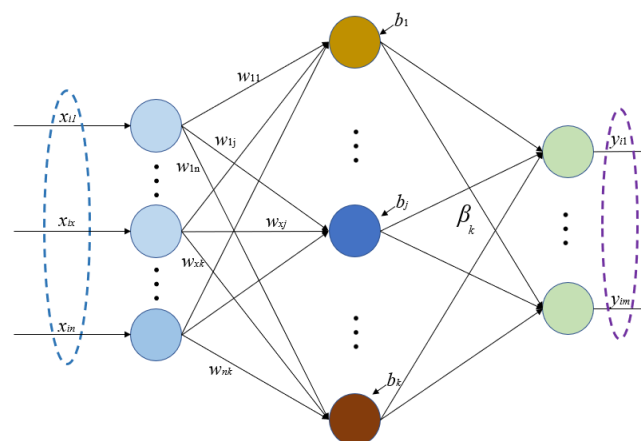

Figure 4. The network structure of ELM.

For any given N training samples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, where $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\mathrm{T}} \in \mathrm{R}^n$, $y_i = [y_{i1}, y_{i2}, \ldots, y_{im}]^{\mathrm{T}} \in \mathrm{R}^n$, for ELM with k hidden nodes, we have

$$\sum_{i=1}^{k} \beta_i g(w_i \cdot x_j + b_i) = y_j \quad j = 1, \ldots, N, \tag{17}$$

where $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{in}]^{\mathrm{T}}$ represents the output connection weights between the ith hidden node and the output nodes, $g(.)$ represents the activation function and $w_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^{\mathrm{T}}$ represents the input connection weights.

$$H = \begin{bmatrix} g(w_1 \bullet x_1 + b_1) & \cdots & g(w_k \bullet x_1 + b_k) \\ \vdots & \cdots & \vdots \\ g(w_1 \bullet x_N + b_1) & & g(w_k \bullet x_N + b_k) \end{bmatrix}. \tag{18}$$

With the introduction of matrix H, Equation (17) is expressed as $H\beta = Y$, in which $Y = [y_1^{\mathrm{T}}, y_2^{\mathrm{T}}, \ldots, y_N^{\mathrm{T}}]^{\mathrm{T}}$, $\beta = [\beta_1^{\mathrm{T}}, \beta_2^{\mathrm{T}}, \ldots, \beta_k^{\mathrm{T}}]^{\mathrm{T}}$; thus, the learning objective function for ELM can be expressed as

$$\min\|H\beta - Y\|. \tag{19}$$

If $g(.)$ is infinitely differentiable, no adjustments are necessary for the output parameters of ELM. The smallest and unique solution β satisfying Equation (19) can be directly calculated:

$$\beta = H^{\dagger}Y, \tag{20}$$

where $H^{\dagger}$ is the Moore–Penrose generalized inverse of the hidden-layer output matrix H.

However, before using ELM, the number of neurons in the hidden layer, type of activation function, and range of input-layer connection weights need to be determined. Improper parameter settings will affect the performance of ELM, which will further affect the effect of temperature compensation.

3.5. Non-Dominated Sorting Genetic Algorithm II

NSGA is an improved genetic algorithm based on Pareto optimality. In this algorithm, the population is stratified according to the non-dominant relationships between individuals, and then selection, crossover, mutation and other operations are carried out. NSGA II [35] introduces the fast non-dominated sorting (FNS) method, congestion-level comparison operators, and an elite strategy to improve NSGA. This not only reduces the complexity of NSGA but also makes the Pareto frontier optimal solution evenly distributed in the whole Pareto domain, so as to alleviate the phenomenon of local optimal solutions and enhance its search ability. The following introduces FNS, the congestion comparison operator, and the elite strategy.

A. Fast non-dominated sorting.

The essence of a multi-objective optimization problem lies in obtaining the Pareto frontier optimal solution. Thus, NSGA II employs an FNS technique to explore the entire population. The population is stratified according to the non-inferiority level of individual solutions, in which the individual solutions in the next layer are all dominated by any solution in the previous layer and there is no dominant relationship between individuals in the same layer. According to this principle, the search is directed to the Pareto-optimal front.

B. Congestion comparison operator.

The crowding degree i_d is defined as the density of individuals in the area where the given point is located. Firstly, the crowding degree i_d of each individual is set to 0.

Following FNS of the population, the crowding degree of individuals on the boundary is assigned infinity, while for others, the crowding degree is

$$i_d = \sum_{j=1}^{m} \left| f_j^{i+1} - f_j^{i-1} \right|, \tag{21}$$

where f_j^{i+1} is the *j*th objective function value of the *i*th point.

C. Elite strategy.

After FNS of the population, the crowding degrees of individuals within the generated non-dominant layers are computed. "Elite" individuals are selected based on a low non-dominant level and high crowding degree, representing the optimal solutions of the Pareto front among these elites.

Figure 5 presents the algorithmic process of NSGA II: Firstly, the initial population P_t, comprising *N* individuals, is randomly generated. Subsequently, the offspring population Q_t is derived using the fundamental operations of the genetic algorithm. Then, the new population R_t with a population size of 2*N* is obtained by reconstructing P_t and Q_t. According to the elite strategy, P_{t+1} is obtained by selecting the optimal *N* individuals in R_t, and their offspring population Q_{t+1} is obtained again through the genetic algorithm. The new population R_{t+1} is then constructed, and the described steps are iteratively executed until the stop condition is satisfied.
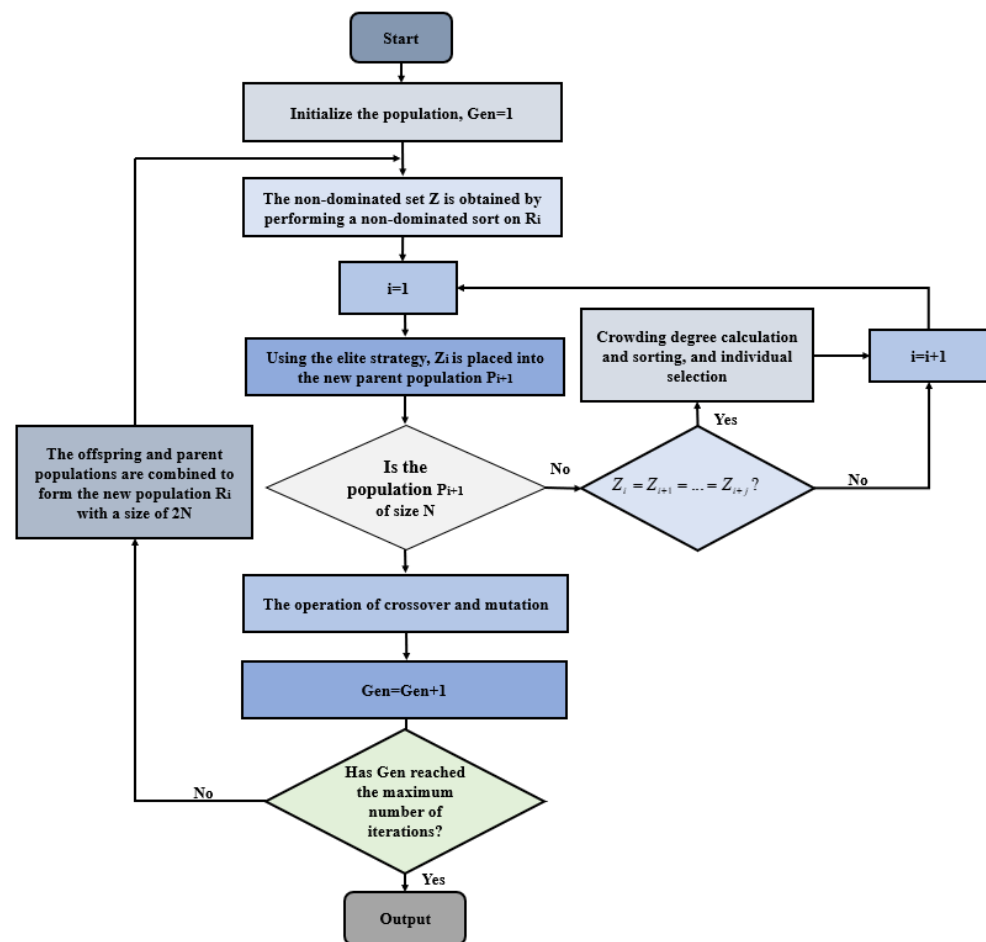

Figure 5. The algorithm process of NSGA II.

3.6. ELM Optimization

To enhance the generalization ability and search ability of ELM, we use NSGA II to optimize the number of neurons in the hidden layer, the type of activation function, and the range of input-layer connection weights. The key technology of NSGA II-ELM lies in the following.

(1) The definition of population individuals.

The mapping relationship between the parameters to be optimized in ELM and the population individuals of the NSGA II algorithm is established. Here, the optimization interval of the hidden neurons was set to [0, 50], and the five most commonly used activation functions—'sigmoid', 'sin', 'hardlim', 'tribas' and 'radbas'—were selected for testing. In order to optimize the input-layer connection weights, the scaling factor λ is introduced to limit the range of the input-layer connection weights. Therefore, the value range of the input-layer connection weights can be expressed as

$$w_{ij} \in \lambda \cdot [-1, 1] \quad \lambda \in [0, 1], \tag{22}$$

where w_{ij} represents the input connection weights, and the randomly set range of the input-layer weights will change with the value of the scaling factor λ, so as to achieve the purpose of limiting their range.

(2) The determination of the objective function.

For ELM, the smaller the output weights and prediction error, the better the network generalization ability. Therefore, the RMSE between the predicted and true values and the 2-norm of the output weight were taken as the two objective functions in this paper.

$$\begin{cases} f_1 = \min \sqrt{\dfrac{\sum\limits_{i=1}^{N} \left(y_{p,i} - y_{t,i}\right)^2}{N}}, \\ f_2 = \min \|\beta\|_2^2 \end{cases} \tag{23}$$

Here, N is the number of input samples, and $y_{p,i}$ and $y_{t,i}$ are the predicted value and true value, respectively, of the ith training sample.

3.7. ICEEMDAN-SE-TFPF and NSGA II-ELM

Relying on the aforementioned algorithms, we propose a hybrid algorithm named ICEEMDAN-SE-TFPF and NSGA II-ELM for gyroscope temperature compensation. The algorithm logic diagram is shown in Figure 6, and the specific steps are as follows:
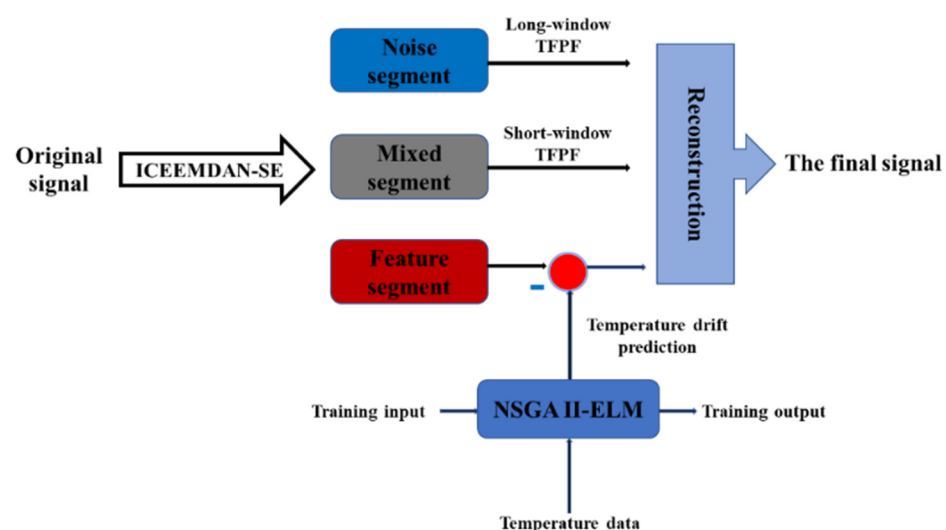

Figure 6. The framework of ICEEMDAN-SE-TFPF and NSGA II-ELM.

Step 1. Firstly, the gyroscope's output containing noise and temperature drift is decomposed into IMFs by ICEEMDAN.

Step 2. Secondly, the decomposed IMFs are classified into a noise segment, mixed segment and feature segment by SE.

Step 3. Long-window TFPF is used to denoise the noise segment with high noise content, short-window TFPF is used to strip the useful signals in the mixed segment, and the feature segment containing temperature drift is compensated by NSGA II-ELM.

Step 4. Finally, the denoised noise segment, denoised mixed segment and compensated feature segment are combined to reconstruct the final gyroscope's output.

4. Experiment

4.1. The Experimental Process

The setup for the temperature experiment on the gyroscope is shown in Figure 7. A detection circuit was distributed across three separate PCBs. Metal pins served not only to interconnect the electronic signals within the detection circuit but also to facilitate the connections between the three PCBs, each of which was encased in a rubber pad. Subsequently, the rubber-wrapped PCBs were placed within a metal shell, a measure that effectively safeguarded the chip's structure while minimizing the risk of severe impact. To mitigate electromagnetic interference, the ground signal was connected to the metal shell. Additionally, one of the three PCBs served as a weak signal interface with the structural chip, while the other two functioned as induction and drive circuits, respectively. The experimental setup included a temperature-controlled oven, a multimeter, a signal generator, and a DC power supply of ± 10 V. The experimental procedure unfolded as follows: Initially, activate the gyroscope and allow it to stabilize at room temperature for one hour. Next, rapidly raise the temperature-controlled oven to 60 °C to ensure that the gyroscope housing reaches a consistent temperature of 60 °C, maintaining this temperature for an additional hour. Subsequently, as the temperature in the oven decreases to 10 °C, the gyroscope continues to operate under these conditions for an hour. Finally, as the temperature drops to −40 °C, the gyroscope operates for one final hour before concluding the experiment.
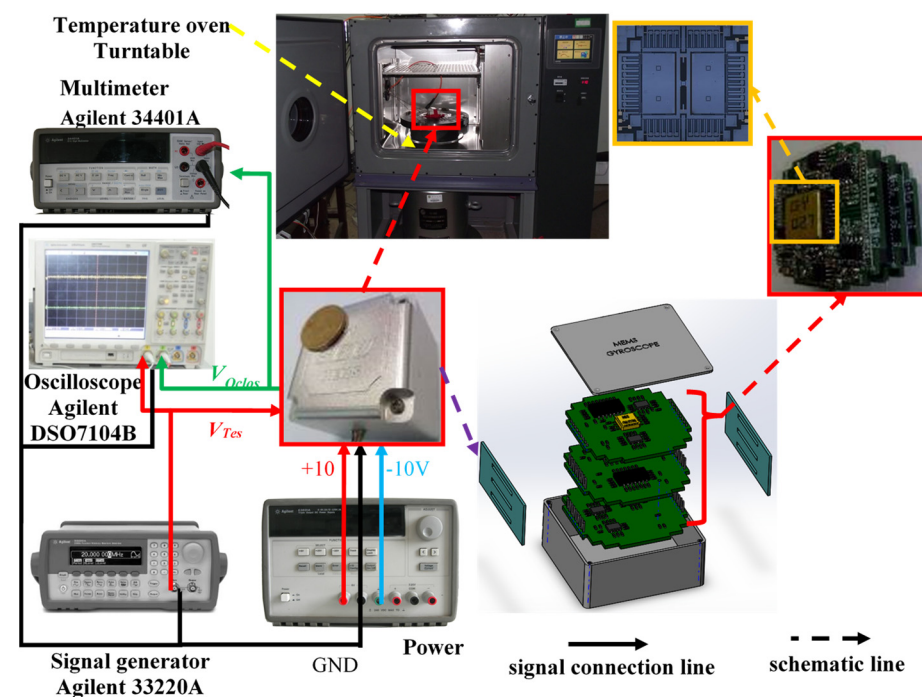

Figure 7. Gyroscope temperature test equipment.

4.2. The Experimental Results

Figure 8 shows the temperature experiment results from the gyroscope. It can be seen that the gyroscope's output presents a nonlinear relationship with the temperature change. As the temperature changed from 60 °C to −40 °C, the output shifted from 0.115° to 0.15°, and the gyroscope's output signal contains not only temperature drift but also noise.
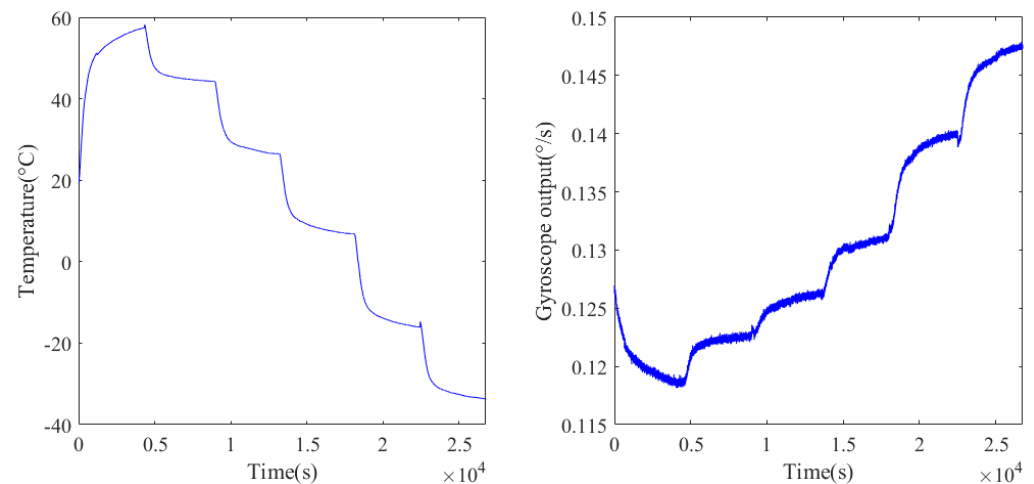

Figure 8. Experimental temperature change and the corresponding gyroscope output.

According to the flow of our proposed hybrid algorithm, the gyroscope's output signal was firstly decomposed by ICEEMDAN. Figure 9 shows that the gyroscope's output signal was decomposed into 12 IMFs. The SE value of each IMF was calculated and classified, and the results are shown in Figure 10. The SE values of the first, second and third IMFs were greater than 0.6, so these IMFs were divided into the noise segment, which contains a lot of noise. Therefore, the long-window TFPF with better denoising ability was selected for denoising this segment. The SE values of the fourth, fifth, sixth and seventh IMFs were between 0.1 and 0.6, so these IMFs were divided into the mixed segment containing noise and useful signals. Because short-window TFPF can well preserve useful components during denoising, it was selected for denoising this segment. The SE values of the 8th, 9th, 10th, 11th and 12th IMFs were below 0.1, so these IMFs were divided into the feature segment, which comprises a drift term caused by the temperature change. This segment was compensated by NSGA II-ELM.

Before establishing the temperature compensation model to compensate the feature segment, NSGA II was used to optimize ELM to obtain the best network parameters. As mentioned before, the optimal searching range for hidden neurons was $[0, 50]$; the activation functions were 'sigmoid', 'sin', 'hardlim', 'tribas' and 'radbas'; and the range of connection weights of the input layer was set to $[-1, 1]$. In addition, the population number was set to 40, the crossover probability was 0.8, the mutation probability was 0.15, and the number of iterations was 15. After the iterations, the optimization result of the NSGA II particle swarm was obtained. Figure 11 shows the distribution of particles in three-dimensional space, where each blue circle represents a particle. The position of the particle marked by the red circle represents the optimal network parameters for ELM, that is, the number of hidden neural networks was 11, the activation function type was the 'sin' function, and the range of input-layer connection weights was $[-0.65, 0.65]$. Figure 12 presents a comparison of each segment before and after processing. In order to investigate the denoising capability of ICEEMDAN-SE-TFPF, we reconstructed the noise segment after denoising, the mixed segment after denoising, and the feature segment. From Figure 13, we can see that the noise was basically eliminated. On this basis, the final signal was obtained by compensating the feature segment. In Figure 14, we can see that the noise component and temperature drift in the gyroscope's output signal were eliminated. To further evaluate the proposed algorithm, a quantitative analysis was performed, as described in the next subsection.
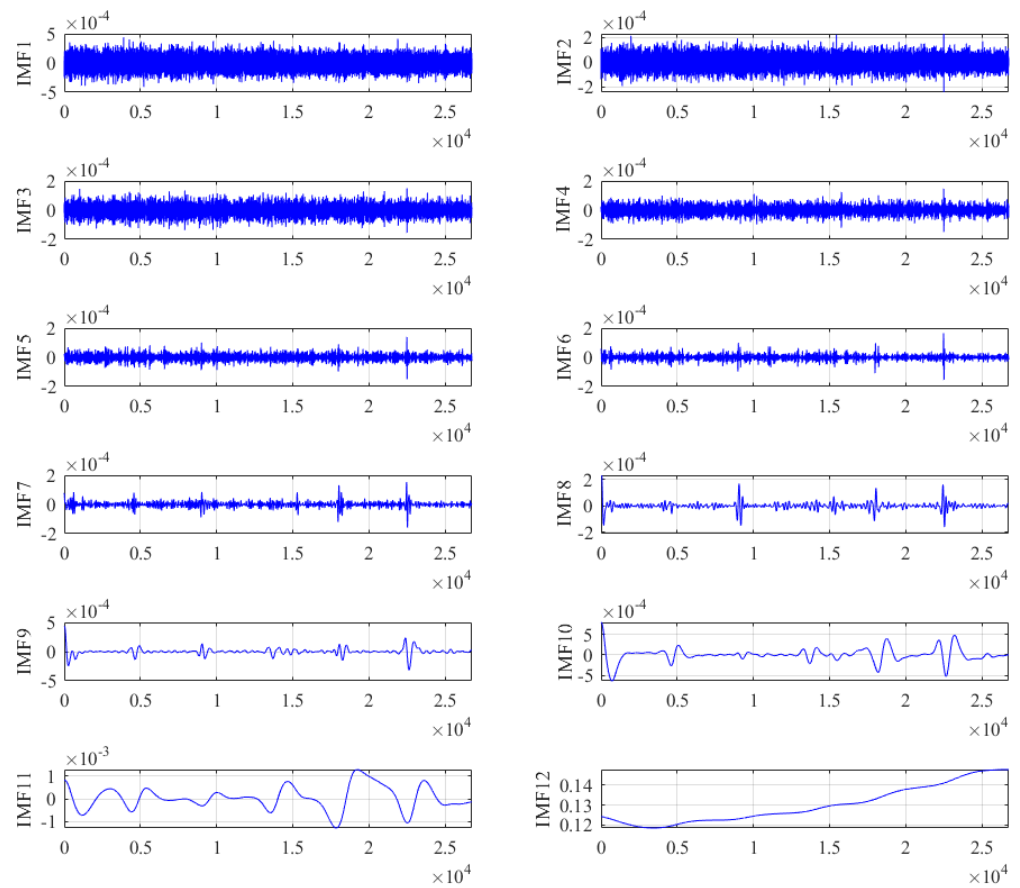
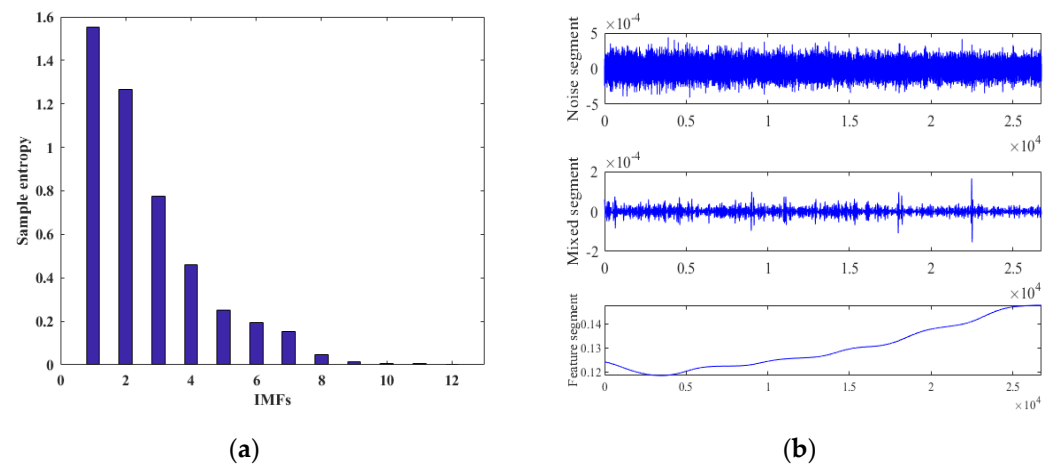**Figure 9.** Decomposition results of the gyroscope's output signal based on ICEEMDAN.

(**a**)

(**b**)

Figure 10. Classification results. (**a**) SE value of each IMF. (**b**) Examples of different segments.

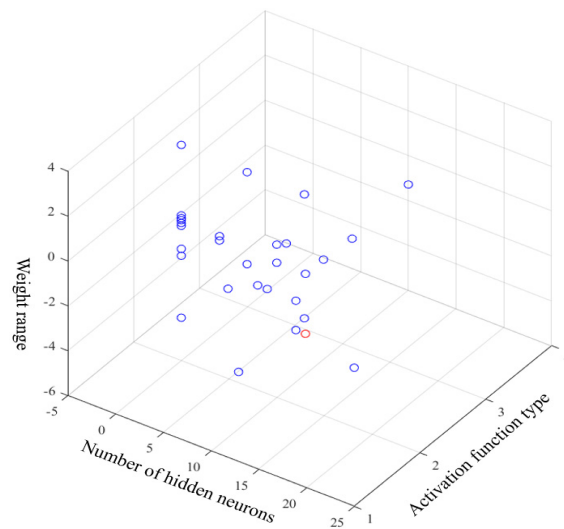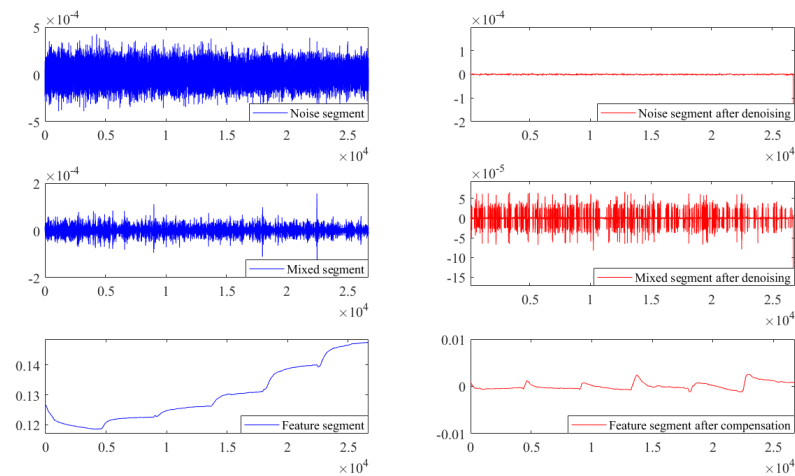**Figure 11.** Distribution of particles in the NSGA II search.

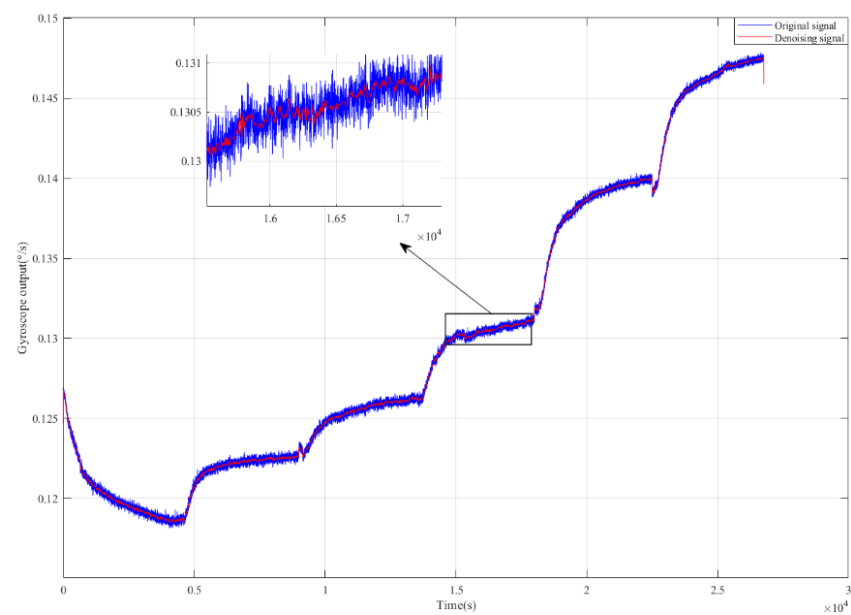**Figure 12.** Noise reduction or compensation results for each segment.

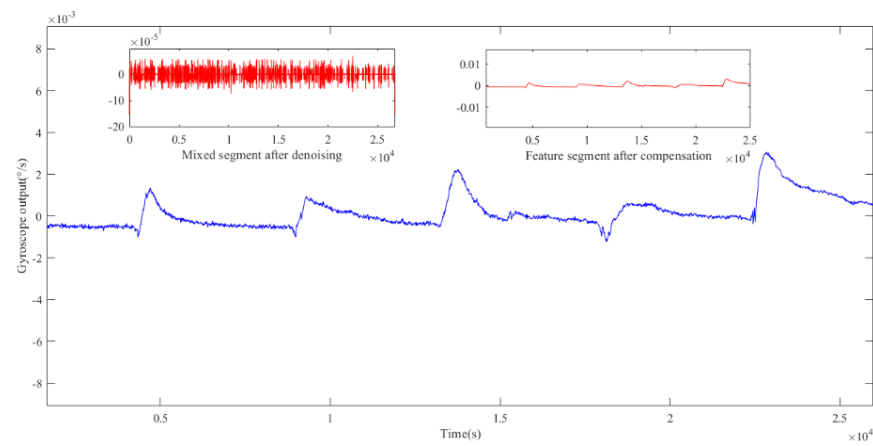**Figure 13.** A comparison between the gyroscope's outputs before and after denoising.

Figure 14. Final output from the gyroscope based on ICEEMDAN-SE-TFPF and NSGA II-ELM.

4.3. Comparative Analysis

Inertial devices commonly employ Allan variance analysis, as outlined in the IEEE standard [36], for error analysis. In this study, we utilized Allan variance analysis to quantitatively assess the angle random walk and bias stability of both the original and compensated gyroscope outputs. Figure 15 illustrates the comparison results. Following denoising and compensation, the angle random walk of the original output decreased from 0.531076 to $6.65894 \times 10^{-3\circ}/\text{h}/\sqrt{\text{Hz}}$, while the bias stability was reduced from $32.7364°/\text{h}$ to $0.259247°/\text{h}$. These findings underscore the superior performance of our proposed hybrid algorithm. A high-precision model is key to improving the compensation effect, so we adopted NSGA II to optimize ELM. To verify the superiority of the improved ELM, a comparative analysis was performed here with a BP neural network and unoptimized ELM, taking the experimental temperature change as the input and the drift part in the gyroscope output as the output for training. In Figure 16, the error of ELM at the inflection point is large when compared with the real output, and the predicted output of the improved ELM is closest to the real value when compared with BP. After the input parameters of ELM were optimized by NSGA II, ELM had stronger generalization ability and learning accuracy.
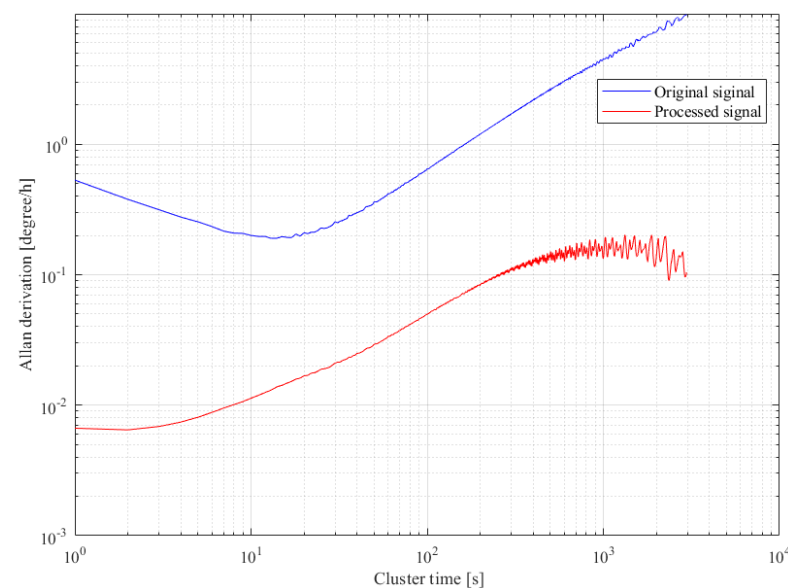

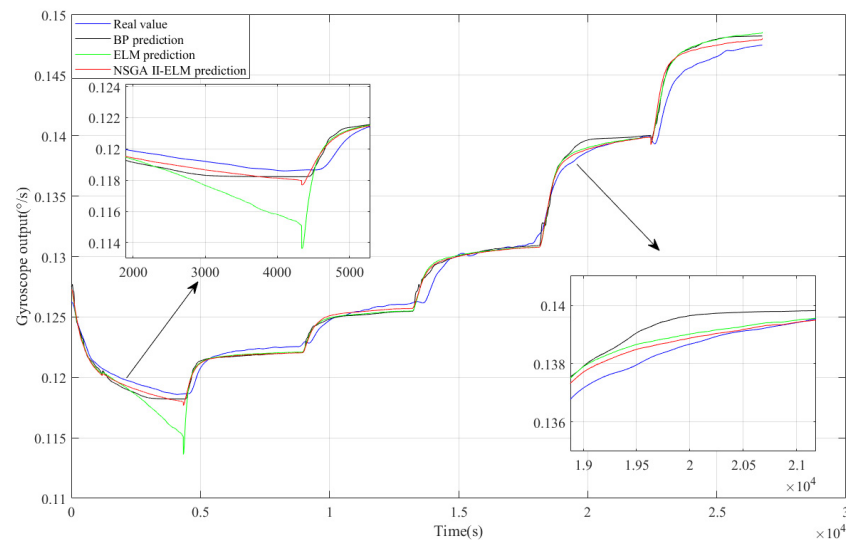**Figure 15.** Allan variance analysis for output signals.

Figure 16. Comparison result for the BP neural network, ELM and NSGA II-ELM.

5. Conclusions

In this paper, we proposed a hybrid algorithm based on ICEEMDAN-SE-TFFPF and NSGA II-ELM for gyroscope temperature compensation. After ICEEMDAN decomposition and SE classification, the gyroscope's output signal was divided into a noise segment, mixed segment and feature segment; then, long-window and short-window TFPF were selected to denoise the noise segment and the mixed segment, respectively. Before processing the feature segment, NSGA II was utilized to optimize ELM, including the number of hidden-layer neurons, the type of activation function, and the range of input-layer connection weights. The optimization objectives were determined by the RMSE of the predicted output and the 2-norm of the output-layer connection weight. This gave ELM good generalization ability and higher prediction accuracy; when compared with the BP neural network and ELM, the predicted value of improved ELM was closest to the real value, displaying improved compensation accuracy. The experimental results show that the angle random walk and bias stability of the compensated output signal were reduced from 0.531076 to $6.65894 \times 10^{-3\circ}/\text{h}/\sqrt{\text{Hz}}$ and from 32.7364 to $0.259247^{\circ}/\text{h}$, respectively.

Author Contributions: Conceptualization, Z.Z.; Methodology, X.Z. and H.C.; Software, X.Z., Z.Z. and J.Z.; Resources, J.Z. and X.Z.; Data curation, J.Z.; Formal analysis, H.C. and Y.R.; Writing—original draft preparation, Z.Z. and J.Z.; Writing—review and editing, Y.R. and J.Y.; Funding acquisition, H.C. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data are contained within the article.

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

References

1. Shen, Q.; Yang, D.; Li, J.; Chang, H. Bias Accuracy Maintenance under Unknown Disturbances by Multiple Homogeneous MEMS Gyroscopes Fusion. *IEEE Trans. Ind. Electron.* **2023**, *70*, 3178–3187. [CrossRef]
2. Ding, X.; Ruan, Z.; Jia, J.; Huang, L.; Li, H.; Zhao, L. In-Run Mode-Matching of MEMS Gyroscopes Based on Power Symmetry of Readout Signal in Sense Mode. *IEEE Sens. J.* **2021**, *21*, 23806–23817. [CrossRef]
3. Pentek, Z.; Hiller, T.; Czmerk, A. Algorithmic Enhancement of Automotive MEMS Gyroscopes with Consumer-Type Redundancy. *IEEE Sens. J.* **2021**, *21*, 2092–2103. [CrossRef]

4. Antonio, J.A.D.; Longo, M.; Zaninelli, D.; Ferrise, F.; Labombarda, A. MEMS-based Measurements in Virtual Reality: Setup an Electric Vehicle. In Proceedings of the 2021 56th International Universities Power Engineering Conference (UPEC), Middlesbrough, UK, 31 August–3 September 2021; pp. 1–5.

5. Chen, L.; Miao, T.; Li, Q.; Wang, P.; Wu, X.; Xi, X.; Xiao, D. A Temperature Drift Suppression Method of Mode-Matched MEMS Gyroscope Based on a Combination of Mode Reversal and Multiple Regression. *Micromachines* **2022**, *13*, 1557. [CrossRef]

6. Cao, H.; Wei, W.; Liu, L.; Ma, T.; Zhang, Z.; Zhang, W.; Shen, C.; Duan, X. A Temperature Compensation Approach for Dual-Mass MEMS Gyroscope Based on PE-LCD and ANFIS. *IEEE Access* **2021**, *9*, 95180–95193. [CrossRef]

7. Cui, J.; Yan, G.; Zhao, Q. Enhanced temperature stability of scale factor in MEMS gyroscope based on multi parameters fusion compensation method. *Measurement* **2019**, *148*, 106947. [CrossRef]

8. Ho, G.K.; Pourkamali, S. Micromechanical IBARs: Tunable high-Q resonators for temperature-compensated reference oscillators. *J. Microelectromech. Syst.* **2010**, *19*, 503–515. [CrossRef]

9. Tao, Y.; Liu, W.; Nan, C.; Wang, S.; Ding, X.; Liu, L.; Bai, Y.; Zhou, Y.; Xing, E.; Chen, J.; et al. Dual closed-loop control method for resonant integrated optic gyroscopes with combination differential modulation. *Opt. Commun.* **2024**, *556*, 130264. [CrossRef]

10. Trusov, A.A.; Schofield, A.R.; Shkel, A.M. Performance characterization of a new temperature-robust gain-bandwidth improved MEMS gyroscope operated in air. *Sens. Actuators A Phys.* **2008**, *155*, 16–22. [CrossRef]

11. Prikhodko, I.P.; Trusov, A.A.; Shkel, A.M. Compensation of drifts in high-Q MEMS gyroscopes using temperature self-sensing. *Sens. Actuators A Phys.* **2013**, *201*, 517–524. [CrossRef]

12. Cao, H.; Cui, R.; Liu, W.; Ma, T.; Zhang, Z.; Shen, C.; Shi, Y. Dual mass MEMS gyroscope temperature drift compensation based on TFPF-MEA-BP algorithm. *Sens. Rev.* **2021**, *41*, 162–175. [CrossRef]

13. Cao, H.; Zhang, Z.; Zheng, Y.; Guo, H.; Zhao, R.; Shi, Y.; Chou, X. A New Joint Denoising Algorithm for High-G Calibration of MEMS Accelerometer Based on VMD-PE-Wavelet Threshold. *Shock Vib.* **2021**, *2021*, 8855878. [CrossRef]

14. Liu, Y.; Chen, G.; Wei, Z.; Yang, J.; Xing, D. Denoising Method of MEMS Gyroscope Based on Interval Empirical Mode Decomposition. *Math. Probl. Eng.* **2020**, *2020*, 3019152. [CrossRef]

15. Xu, L.; Li, W.; Zhang, B.; Zhu, Y.; Lang, C. Fault Diagnosis of Mine Truck Hub Drive System Based on LMD Multi-Component Sample Entropy Fusion and LS-SVM. *Actuators* **2023**, *12*, 468. [CrossRef]

16. Wei, J.; Zhang, Z.; Cao, H.; Duan, X. Hybrid Temperature Compensation Model of MEMS Gyroscope Based on Genetic Particle Swarm Optimization Variational Modal Decomposition and Improved Backpropagation. *Sens. Mater.* **2021**, *33*, 2835–2856. [CrossRef]

17. Motin, M.A.; Karmakar, C.K.; Palaniswami, M. Ensemble Empirical Mode Decomposition with Principal Component Analysis: A Novel Approach for Extracting Respiratory Rate and Heart Rate From Photoplethysmographic Signal. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 766–774. [CrossRef]

18. Tian, S.; Bian, X.; Tang, Z.; Yang, K.; Li, L. Fault Diagnosis of Gas Pressure Regulators Based on CEEMDAN and Feature Clustering. *IEEE Access* **2019**, *7*, 132492–132502. [CrossRef]

19. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomed. Signal Process. Control* **2014**, *14*, 19–29. [CrossRef]

20. Donoho, D.L.; Johnstone, I.M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Am. Stat. Assoc.* **1995**, *90*, 1200–1224. [CrossRef]

21. Donoho, D.L.; Johnstone, I.M. Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **1994**, *81*, 425–455. [CrossRef]

22. Boashash, B.; Mesbah, M. Signal enhancement by time-frequency peak filtering. *IEEE Trans. Signal Process.* **2004**, *52*, 929–937. [CrossRef]

23. Shen, C.; Song, R.; Li, J.; Zhang, X.; Tang, J.; Shi, Y.; Liu, J.; Cao, H. Temperature drift modeling of MEMS gyroscope based on genetic-Elman neural network. *Mech. Syst. Signal Process.* **2016**, *72–73*, 897–905.

24. Song, R.; Chen, X.; Shen, C.; Zhang, H. Modeling FOG Drift Using Back-Propagation Neural Network Optimized by Artificial Fish Swarm Algorithm. *J. Sens.* **2014**, *2014*, 276043. [CrossRef]

25. Li, J.; Jiao, F.; Fang, J.; Cheng, J. Temperature Error Modeling of RLG Based on Neural Network Optimized by PSO and Regularization. *IEEE Sens. J.* **2014**, *14*, 912–919.

26. Gu, H.; Liu, X.; Zhao, B.; Zhou, H. The In-Operation Drift Compensation of MEMS Gyroscope Based on Bagging-ELM and Improved CEEMDAN. *IEEE Sens. J.* **2019**, *19*, 5070–5077. [CrossRef]

27. Wang, Z.; Yan, J.; Chen, F.; Peng, X.; Zhang, Y.; Wang, Z.; Duan, S. Sensor Drift Compensation of E-Nose Systems with Discriminative Domain ReCconstruction Based on an Extreme Learning Machine. *IEEE Sens. J.* **2021**, *21*, 17144–17153. [CrossRef]

28. Bacanin, N.; Stoean, C.; Zivkovic, M.; Jovanovic, D.; Antonijevic, M.; Mladenovic, D. Multi-Swarm Algorithm for Extreme Learning Machine Optimization. *Sensors* **2022**, *22*, 4204. [CrossRef]

29. Han, F.; Yao, H.-F.; Ling, Q.-H. An improved evolutionary extreme learning machine based on particle swarm optimization. *Neurocomputing* **2013**, *116*, 87–93. [CrossRef]

30. Wu, Y.; Zhang, Y.; Liu, X.; Cai, Z.; Cai, Y. A multiobjective optimization-based sparse extreme learning machine algorithm. *Neurocomputing* **2018**, *317*, 88–100. [CrossRef]

31. Xu, L.; Su, H.; Cai, D.; Zhou, R. RDTS Noise Reduction Method Based on ICEEMDAN-FE-WSTD. *IEEE Sens. J.* **2022**, *22*, 17854–17863. [CrossRef]

32. Sun, Q.; Liu, C.; Chen, T.; Zhang, A.; Liu, C.; Tao, Y. Adaptive Decomposition and Multitimescale Analysis of Long Time Series of Climatic Factors and Vegetation Index Based on ICEEMDAN-SVM. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 6203–6219. [CrossRef]

33. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef]

34. Bai, L.; Han, Z.; Li, Y.; Ning, S. A Hybrid De-Noising Algorithm for the Gear Transmission System Based on CEEMDAN-PE-TFPF. *Entropy* **2018**, *20*, 361. [CrossRef]

35. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]

36. Wang, L.; Zhang, C.; Lin, T.; Li, X.; Wang, T. Characterization of a fiber optic gyroscope in a measurement while drilling system with the dynamic Allan variance. *Measurement* **2015**, *75*, 263–272. [CrossRef]