ISA Transactions 131 (2022) 397–414

Contents lists available at ScienceDirect

ISA Transactions

journal homepage: www.elsevier.com/locate/isatrans

Research article

Incipient chatter fast and reliable detection method in high-speed milling process based on cumulative strategy

Yanqing Zhao [a,b], Kondo H. Adjallah [b,*], Alexandre Sava [b], Zhouhang Wang [b]

[a] Faculty of Transportation Engineering, Huaiyin Institute of Technology, Huai'an, 223003, China
[b] Université de Lorraine, LCOMS, F-57000 Metz, France

ARTICLE INFO

Article history:
Received 18 November 2021
Received in revised form 27 May 2022
Accepted 27 May 2022
Available online 8 June 2022

Keywords:
Adaptive indicator
Cumulative information
Detection
High-speed milling
Incipient chatter
Sampling frequency

ABSTRACT

This paper proposes an incipient chatter detection method to meet high dynamic applications' time and reliability constraints, such as high-speed milling involving heavy noise. The herein introduced method relies on a multiple sampling per revolution (MSPR) technique, coupled with two data preprocessing techniques, a modified adaptive cumulative chatter indicator, and a two-risk levels-based threshold. The MSPR technique enables collecting information-rich enough data to characterize the chatter dynamics thanks to a significant amount of data collected in each revolution. Therefore, the MSPR technique allows for acquiring the data using a short-time window, thus reducing the detection delay. Two data preprocessing techniques, i.e., Z-score normalization and mean-centered, are implemented for data integration and chatter information consolidation. The modified adaptive cumulative chatter indicator has three advantages: (a) it accumulates the information on the chatter feature and highlights the appearance of an incipient chatter; (b) it adapts to the variation of the environmental disturbance noises, resulting in enhanced detection reliability; (c) it is faster than the adaptive cumulative log-likelihood ratio (ACLLR) for decision-making statistically. The two-risk levels-based threshold overcomes the limitations of a unique threshold, and allows simultaneously assessing the two risk levels, thus improving detection reliability. We successfully applied the proposed method to detect incipient chatter in a digital high-speed milling process and assessed its effectiveness by comparing it with several existing chatter detection methods.

1. Introduction

Chatter appears in milling processes frequently, especially in high-speed milling processes. An unsuitable machining parameters combination or the machine tool or cutting tool's degradation can result in chatter [1,2]. Currently, experts set the machining parameters within the stability regions and close to the instability boundary to improve productivity. However, due to slight variations of machining parameters or wear of the machine's cutting tool, incipient or severe chatter may arise. The severe chatter has distinct characteristics and can be detected using traditional methods. Therefore, many research results about severe chatter detection are reported in the literature [1]. Unlike severe chatter, the incipient chatter refers to a slight chatter with a small amplitude and tends to be buried by the environmental disturbance noise. To the best of our knowledge, there are few works on incipient chatter detection in the literature. Some machined products require a high-quality surface so that even an incipient chatter may not be tolerable [3]. Thus, fast and reliable detection of incipient chatter is relevant in high-speed milling processes to ensure the machining quality of the workpiece, prevent premature wear or tool breaking, and improve productivity.

Regenerative chatter and mode coupling chatter have been reported extensively in milling processes. Both types of chatter have different mechanisms. The regenerative chatter stems from the regenerative cutting process and happens locally at the cutting tool or workpiece. Its vibration frequencies are different from the spindle rotating frequency and corresponding harmonics. Besides, these vibration frequencies can exceed hundreds of Hz [4]. Unlike the regenerative chatter, the mode coupling chatter arises due to the machine tool's insufficient stiffness and mode coupled structure. Its vibration frequencies depend on the natural frequencies of the different elements of the machining system. Commonly, the vibration frequencies of the mode coupling chatter are relatively low [5,6]. In other words, the vibration wavelength in the mode coupling chatter is relatively long. The MSPR technique cannot acquire information-rich enough data in a short-time window to characterize the mode coupling chatter dynamics. Our work focuses on regenerative rather than mode coupling chatter, using a short-time window for data collection.

* Corresponding author.
E-mail address: kondo.adjallah@univ-lorraine.fr (K.H. Adjallah).

https://doi.org/10.1016/j.isatra.2022.05.039
0019-0578/

Nomenclature

Acronyms

ACLLR	Adaptive cumulative log-likelihood ratio
CLLR	Cumulative log-likelihood ratio
DOF	Degree of freedom
EMD	Empirical mode decomposition
FHF	Failure hazard function
FP	Failure probability
LLR	Log-likelihood ratio
M-ACLLR	Modified adaptive cumulative log-likelihood ratio
MaxEnt	Maximum entropy
MSPR	Multiple sampling per revolution
OMP	Orthogonal matching pursuit
OSPR	Once sampling per revolution
PDF	Probability density function
SLD	Stability lobe diagram
SNR	Signal-to-noise ratio
SPRT	Sequential probability ratio test
VMD	Variational mode decomposition
WT	Wavelet transform
WTMM	Wavelet transform's modulus maxima

Symbols

α	First risk-level or type I error
β	Second risk-level or type II error
σ_X	Standard deviation of the raw data
$f(x)$	Probability density function
$H(\cdot)$	Shannon information entropy
$g_j(x)$	Specific function
$E(\cdot)$	Statistical expectation
L	Lower-value
$\log\Lambda(\cdot)$	Log-likelihood ratio
R	Upper-value
C_0	Complexity
$p(\cdot)$	Probability
S_n	Adaptive cumulative log-likelihood ratio
S_n'	Modified adaptive cumulative log-likelihood ratio
$x(t)$	Measured milling signal
x_i	ith raw data point
$\overline{X}$	Mean value
z_i	Normalized data point
A	Lower threshold
B	Upper threshold
A_k	Amplitudes of the tooth passing frequency and its harmonics
C_k^+	Complex conjugate amplitudes of the chatter frequency and its harmonics
C_k^-	Complex conjugate amplitudes of the chatter frequency and its harmonics
M	Number of samples collected per revolution
N	Number of revolutions defining the time window width
N_1	Number of significant harmonics of the tooth passing frequency
N_2	Number of significant harmonics of the chatter frequency
T_n	Unique threshold
Z	Number of the cutting teeth
Ω	Spindle revolution speed
w	Basic chatter frequency
$n(t)$	Environmental disturbance noise
x-direction	Feed direction
y-direction	Cross-feed direction
x	Displacement of the cutting tool in the feed direction
$\dot{x}$	Velocity of the cutting tool in the feed direction
$\ddot{x}$	Acceleration of the cutting tool in the feed direction
y	Displacement of the cutting tool in the cross-feed direction
$\dot{y}$	Velocity of the cutting tool in the cross-feed direction
$\ddot{y}$	Acceleration of the cutting tool in the cross-feed direction
y_w	Displacement of the workpiece in the cross-feed direction
$\dot{y}_w$	Velocity of the workpiece in the cross-feed direction
$\ddot{y}_w$	Acceleration of the workpiece in the cross-feed direction
m_x	Model mass of the cutting tool in the feed direction
m_y	Model mass of the cutting tool in the cross-feed direction
m_{wy}	Model mass of the workpiece in the cross-feed direction
c_x	Model damping of the cutting tool in the feed direction
c_y	Model damping of the cutting tool in the cross-feed direction
c_{wy}	Model damping of the workpiece in the cross-feed direction
k_x	Model stiffness of the cutting tool in the feed direction
k_y	Model stiffness of the cutting tool in the cross-feed direction
k_{wy}	Model stiffness of the workpiece in the cross-feed direction
F_x	Cutting force acting on the cutting tool in the feed direction
F_y	Cutting force acting on the cutting tool in the cross-feed direction
F_{wy}	Cutting force acting on the workpiece in the cross-feed direction
k_{tc}	Cutting force coefficient
k_{nc}	Cutting force coefficient
k_{te}	Cutting force coefficient
k_{ne}	Cutting force coefficient
b	Axial cutting depth
$h(t)$	Instantaneous chip thickness
φ	Tooth angle

Zhu and Liu [7] recently proposed a critical review of methods for detecting and identifying chatter, where chatter detection is among the major current research issues in developing smart machine tools or smart spindles. These chatter detection methods involve three main steps: data acquisition, feature extraction, and decision-making, as a general fault detection scheme. Detecting incipient chatter requires identifying its onset, starting from a signature observed in the feature extracted from valid data acquired using suitable sensors.

Different sensors allow chatter data acquisition, which is the first step to incipient chatter detection. A good data acquisition process should capture enough chatter information. The most used types of sensors for collecting data for chatter detection are acceleration sensors, force sensors, displacement sensors, and acoustic emission sensors [1,7–9]. In [7,8], the authors compared and identified different chatter detection sensors' relative strengths and weaknesses. Given that chatter is a self-excited vibration, it is natural to use a vibration sensor. The vibration signal of a cutting process embeds the dynamic behavior directly, and it has been commonly used in chatter detection [10]. Thus, this work chooses the vibration signal for incipient chatter detection.

The feature extraction step is essential for the real-time implementation of incipient chatter detection. Feature extraction methods rely, in general, on the signal processing methods and mathematical models of chatter physics.

The chatter signal is a complex nonlinear and non-stationary signal. Thus, one uses nonlinear and non-stationary signal processing methods to extract chatter features. One can categorize these signal processing methods into the time, frequency, and time–frequency domains. Among them, time–frequency domain methods such as wavelet transform (WT) [11,12], empirical mode decomposition (EMD) [13], variational mode decomposition (VMD) [14,15], and related improved methods [16,17] are used widely. These time–frequency domain methods decompose the chatter signal into several components/modes with different frequency bands. Thus, the chatter feature is extracted based on the components/modes incorporating rich chatter information. However, the WT-based chatter detection methods require setting suitable WT's parameters and mother wavelet. The unsuitable setting of the WT's parameters or mother wavelet can reduce the decomposition accuracy and lead to low detection reliability. Aimed at this issue, Tran et al. [18] proposed determining the WT's parameters and mother wavelet using an optimizing process. Although the optimizing process optimizes the WT's parameters or mother wavelet, this chatter detection method's computational cost and detection delay may be inadequate in the high-speed milling process. The time from chatter onset to mature chatter can be less than 100 ms for the high-speed milling process [19]. The EMD-based methods do not have a strengthened theoretical basis. Therefore, the EMD-based methods are limited to detecting established chatter in practice. Besides, the detection strategies based on these methods depend on machining parameters such as the spindle revolution speed. To avoid the influence of machining parameters, some researchers proposed employing the comb filter at the data preprocessing stage to filter the spindle rotating frequency, the tooth passing frequency, and their harmonics [19,20]. However, the appropriate setting of the comb filter parameters is challenging, and the comb filter implementation involves computational cost.

Another class of chatter detection methods relies on the mathematical models of chatter physics, such as the probabilistic process models [21,22], bifurcation and chaos models [23–25], and artificial intelligence models [18,26–28].

Many random factors influence the chatter appearance in the milling process, such as the uncertainties of machining parameters and in-homogeneous materials [29]. Therefore, the probabilistic process models are suitable for chatter detection [30]. The probability distribution of the cutting force in chatter-free conditions is different from that obtained in chatter conditions in the milling process. Du et al. [21] proposed to detect chatter by using the variance ratio of the probability distribution of the cutting force to deal with this phenomenon. Although their method can detect severe chatter, it is not appropriate for incipient chatter detection due to the fuzzy boundary between chatter-free and incipient chatter processes. There is a transient phenomenon between the chatter-free process and the chatter process onset. The wavelet transform's modulus maxima (WTMM) approach allows to effectively detect the transient phenomena in image enhancement [31]. Therefore, Wang and Liang [22] proposed to detect chatter based on the feature extracted using the probability distribution of WTMM. The authors demonstrated that the WTMM probability distribution follows the Weibull law. Experiment results concluded this method could detect chatter in constant and variable speed regimes. Nevertheless, this method fails to detect incipient chatter reliably due to its lightness.

The bifurcation and chaos models incorporating nonlinear dynamic properties of the milling processes are also used to detect chatter. Schmitz et al. [23,24] proposed to detect chatter by employing the once-per-revolution sampling for vibration data acquisition. This sampling technique, also called the "one sampling per revolution" (OSPR) technique, aims to collect synchronously one sample in each spindle revolution. According to Schmitz et al. [23], the OSPR data distribution is tight in chatter-free conditions, while it is broader in chatter conditions. Thus, the difference between the OSPR data distributions in chatter-free and chatter conditions allows chatter detection. The OSPR data is independent of machining parameters. So it can be used for real-time chatter detection in high-speed milling processes with variable spindle speed. Honeycutt and Schmitz [25] used the OSPR to identify the instability in milling simulation. Although OSPR-based methods in [23–25] can detect chatter or identify the instability in the milling process, they may fail to detect incipient chatter in the noisy environment reliably.

Besides probabilistic process models and bifurcation and chaos models, the artificial intelligence models are also used to detect the chatter process during their rapid development. Artificial intelligence models are commonly used to extract chatter features or diagnose chatter. Dun et al. [27] proposed using auto-encoding and hybrid clustering to extract chatter features and diagnose chatter. Also, Wan et al. [28] proposed to diagnose chatter using an adaptive boosting algorithm and support vector machine. Li et al. [26] proposed to use gradient tree boosting for the chatter intelligent diagnosis. Although these artificial intelligence methods-based models can detect chatter effectively, they need massive datasets to train artificial intelligence models. Table 1 summarizes the abovementioned chatter detection methods.

The abovementioned methods mainly focus on severe or early chatter detection and do not consider incipient chatter. Few incipient chatter detection methods are found in the literature. Wang et al. [32] proposed a weak chatter detection method based on a sparse dictionary matrix and orthogonal matching pursuit (OMP). These two techniques are used to extract the incipient chatter frequencies. Although the experimental results verify the effectiveness of this method, it needs the predicted chatter frequencies as prior knowledge for constructing the sparse dictionary matrix. In practice, it is challenging to predict incipient chatter frequencies reliably due to the high-speed milling process complexity. Besides, this method may not be able to detect incipient chatter in real-time due to its relatively high computational cost.

To implement real-time incipient chatter detection, Zhao et al. [33] suggested a cumulative approach combining the OSPR technique, the maximum entropy (MaxEnt) principle, and the sequential probability ratio test (SPRT). The OSPR technique is used to

Table 1
Summary of the abovementioned chatter detection methods.

Methods' category	References	Extracted features	Limitations
Based on signal processing methods	[11]	Standard deviation of wavelet transform and wavelet packet energy ratio	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters
	[12]	Auto-correlation coefficient and improved wavelet packet entropy	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters
	[13]	Mean value and standard deviation of the Hilbert–Huang spectra	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters (4) No strengthened theoretical basis
	[14]	Energy entropy based on variational mode decomposition and wavelet packet decomposition	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters
	[15]	Approximate entropy and sample entropy, based on optimized variational mode decomposition	(1) Failure to detect incipient chatter reliably (2) Dependency on machining parameters
	[16]	Standard deviation and energy ratio based on wavelet packet transform	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters
	[17]	Variance ratio based on Hilbert–Huang transform	(1) Failure to detect incipient chatter reliably (2) Dependency on machining parameters (3) No strengthened theoretical basis
	[20]	C_0 complexity and power spectral entropy based on ensemble empirical modes decomposition	(1) Failure to detect incipient chatter reliably (2) No strengthened theoretical basis
	[19]	33 statistical features in time, frequency, and time–frequency domains	(1) Failure to detect incipient chatter reliably (2) No selection criterion
Based on mathematical models of chatter physics	[21]	Variance ratio based on a low pass filter	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters
	[22]	Distributional property of the wavelet transform modulus maxima	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Dependency on machining parameters
	[23] [24]	Variance based on OSPR	(1) Failure to detect incipient chatter reliably
	[25]	Absolute value of the differences in successive OSPR data	(1) Failure to detect incipient chatter reliably
	[18]	Eight features (Kurtosis, skewness, root mean square et al.)	(1) Failure to detect incipient chatter reliably (2) Need a massive datasets
	[26]	Multiscale permutation entropy and multiscale power spectral entropy	(1) Failure to detect incipient chatter reliably (2) No selection criterion (3) Need a massive datasets
	[27]	Compressed signal in two dimensions	(1) Failure to detect incipient chatter reliably (2) Need a massive datasets
	[28]	Nine features in time and frequency domains and compressed signal in eight dimensions	(1) Failure to detect incipient chatter reliably (2) Need a massive datasets

acquire data independent of the machining parameters, and the MaxEnt principle allows extracting the chatter feature. The SPRT accumulates the weak chatter information on the chatter feature and highlights the incipient chatter appearance. This approach detects incipient chatter regardless of machining parameters and is suitable for remote chatter monitoring using wireless sensors. Nevertheless, its detection delay may not meet the high-speed or ultra-high-speed machining requirements.

Generally, one sets the sampling frequency much higher in chatter detection than the spindle rotating frequency for acquiring high-order harmonics information. Thus, one can collect multiple samples in each revolution and obtain multiple sets of OSPR data rather than one set of OSPR data. Since multiple samples are collected in each revolution, we call these OSPR data sets "multiple sampling per revolution" (MSPR) data. In [30], the authors proposed to detect early chatter based on the MSPR. Although this detection method can detect early chatter, the detection delay is still penalizing since the recommended time window width is big. The recommended time window widths lie between 20 and 30 spindle revolutions. Besides, the detection reliability is relatively poor due to using the unique threshold in the decision-making step. The unique threshold is determined based on the small probability event principle and considers only one risk level [34]. The appearance of incipient chatter is recognized when the chatter index exceeds a preset unique threshold value.

Most chatter detection methods use the unique threshold to decide the chatter onset [7,34]. However, the unique threshold methods cannot simultaneously deal with two risk levels in a noisy environment. In the presence of the disturbance noise, the chatter index has randomness and normal evolving characteristics as well. The chatter index may not exceed the unique threshold value in incipient chatter condition. In this situation, a missing alarm arises. Conversely, the chatter index may exceed the unique threshold value when the operational process is normal, which will trigger a false alarm. Due to these two risks of detection error, the unique threshold cannot recognize incipient chatter reliably in noisy environments.

To improve the incipient chatter detection by overcoming the unique threshold's limitation, Zhao et al. [33,35] proposed

Table 2
Summary of our previous contributions to incipient chatter detection.

References	Data acquisition	Feature extraction	Decision-making	Applications	Limitations
[33]	OSPR	MaxEnt principle	SPRT	Remote detection	(1) Long detection delay (2) Low reliability
[30]	MSPR	MaxEnt feature-based reliability model method	Unique threshold	Fast detection	(1) Low reliability
[35]	OSPR	MaxEnt principle	SPRT with a dynamic threshold variant	Remote detection in a variable noisy environment	(1) Long detection delay

a chatter detection method based on SPRT. The SPRT theory considers the abovementioned two risk levels. After recognizing the chatter-free state, the cumulative log-likelihood ratio (CLLR) is reset to zero for monitoring the milling process condition. However, only incipient chatter detection is helpful in the incipient chatter detection process. In other words, the chatter-free detection procedure is not helpful for incipient chatter detection. The entire detection delay may be reduced statistically by discarding the chatter-free detection procedure. Table 2 summarizes our previous contributions to incipient chatter detection.

In summary, the incipient chatter detection method we proposed in this research paper improves our previous contributions [30,33,35] by shortening the detection delay and enhancing the detection reliability in a variable disturbance noisy environment.

The research work proposed in this paper includes three main contributions. First, we proposed acquiring more data per revolution to consolidate the incipient chatter information using an MSPR technique. This technique allows using a short-time window, reducing the detection delay. Second, starting from implementing the data integration and chatter information consolidation, we propose a new chatter indicator called modified adaptive cumulative log-likelihood ratio (M-ACLLR). This indicator accumulates the chatter feature information sequentially, adapts to the variations in the environmental disturbance noises, and reduces the detection delay statistically. The incipient chatter information embedded in the raw MSPR data is consolidated through mean-centered, and the T-score enables enhancing the data quality. Third, we introduced a two-risk levels-based threshold to enhance the detection reliability while overwhelming the unique threshold's limitations. The proposed detection method consists of six steps: (1) specify prior knowledge, (2) acquire MSPR data, (3) normalize and integrate MSPR data, (4) estimate the maximum entropy (MaxEnt) of the integrated MSPR data by using the MaxEnt principle, (5) calculate the modified adaptive cumulative chatter indicator, and (6) make a decision according to the two-risk levels-based threshold. The effectiveness of the proposed detection method is assessed through a case study of a high-speed milling process.

The rest of the paper is organized as follows. First, Section 2 introduces the MSPR technique, two data normalization approaches, the MaxEnt principle, the modified adaptive cumulative chatter indicator, the two-risk levels-based threshold, and the proposed incipient chatter detection method. Next, Section 3 describes a dynamic milling model and milling signals generation. Then, Section 4 assesses the proposed method's effectiveness through digital simulations and discusses the results. Lastly, Section 5 concludes this paper with a prospect for future research direction.

2. Theoretical background

2.1. Characteristic analysis of the milling vibration signal

The milling vibration signal can be composed of periodic, chatter, and noise components [36]. The periodic component is generated by the revolution motion of the cutting tool. The chatter component is raised from the chip regeneration effect. The noise component stems from the measurement error or random disturbance in the actual milling process. According to Refs. [31,36], the measured milling vibration signal in a steady-state case can be approximately expressed as follows,

$$x(t) = \underbrace{\sum_{k=-N_1}^{N_1} A_k e^{j(\frac{kZ\Omega\pi}{30})t}}_{periodic\ \ component} + \underbrace{\sum_{k=-N_2}^{N_2} \left[C_k^+ e^{j(w+\frac{kZ\Omega\pi}{30})t} + C_k^- e^{j(-w+\frac{kZ\Omega\pi}{30})t}\right]}_{chatter\ \ component} + \underbrace{n(t)}_{noise\ \ component}, \quad (1)$$

where $x(t)$ represents the measured milling signal; A_k indicates the amplitudes of the tooth passing frequency and its harmonics; C_k^+ and C_k^- denotes a pair of complex conjugate amplitudes of the chatter frequency and its harmonics; N_1 and N_2 are the number of significant harmonics of the tooth passing frequency and chatter frequency, respectively; Z represents the number of the cutting teeth; Ω denotes the spindle revolution speed in r.p.m, w is the basic chatter frequency; $n(t)$ represents the environmental disturbance noise.

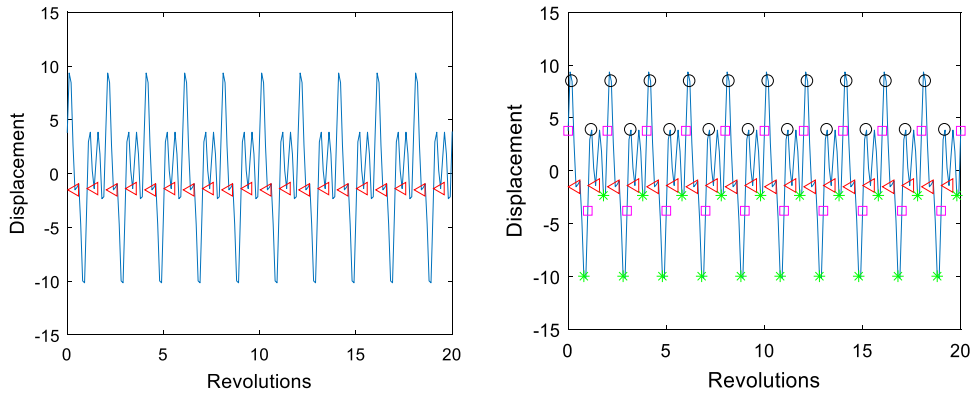
This work aims to detect the appearance of the incipient chatter, which is also called slight chatter and is related to bifurcation phenomena [37]. Thus, bifurcation analysis methods may be suitable for the same purpose in milling processes.

2.2. MSPR technique

Before introducing the MSPR technique, let us recall the OSPR one. The OSPR technique is derived from the Poincaré map, and it aims to acquire one sample per revolution synchronously. This technique has been used to indicate the chatter appearance [23,24]. Nevertheless, using a small amount of OSPR data to estimate the chatter feature leads to low accuracy results, and researchers often use an extended-time window to overcome this shortcoming. Unfortunately, this solution delays the detection. Moreover, a set of OSPR data cannot reveal the chatter reliably. For instance, using a set of OSPR data may not be suitable to detect the chatter resulting in a period-2 bifurcation [38].

Unlike the OSPR technique, the MSPR technique collects multiple samples synchronously rather than one sample per revolution by repeatedly implementing the OSPR at different rotation angles. Thus, the MSPR procedure described hereafter improves the OSPR technique performance:

(1) Set the number of samples M collected per revolution and the number of revolutions N defining the time window width.
(2) Repeat the OSPR M times at different rotation angles to obtain M sets of OSPR data, i.e., MSPR data.



(a) OSPR data. (b) MSPR data (four sets of OSPR data).

Fig. 1. OSPR data (a) and MSPR data (b) on a chatter signal.

Although the MSPR technique is a type of synchronous sampling, it is different from the classical synchronous sampling commonly used for order tracking. The classical synchronous sampling regards synchronous data as a series of data, while the MSPR technique considers multiple independent OSPR data sets [30].

To show the advantages of the MSPR technique, we illustrate the OSPR data and MSPR data using a chatter signal in Fig. 1(a) and (b), respectively. We plot only four sets of OSPR data in MSPR data to clearly show the OSPR data fluctuation. In practice, one should set the number of samples M collected per revolution according to the sampling costs and reliability requirements.

As shown in Fig. 1(a), a set of OSPR data cannot reveal the chatter related to the period-2 bifurcation. This set of OSPR data is almost constant for the chatter related to the period-2 bifurcation. Therefore, we cannot distinguish the chatter-free process and the chatter process related to the period-2 bifurcation using this set of OSPR data. Besides, one set of OSPR data may not be enough to estimate the chatter feature for the chatter related to the Hopf bifurcation. Although an extended-time window allows collecting more samples for estimating the chatter feature, it increases the chatter detection delay and becomes unsuitable for high dynamic applications.

The MSPR technique allows sampling multiple sets of OSPR data, consolidating their chatter information, and reducing the detection delay. Fig. 1(b) shows the MSPR data, including four sets of OSPR data. Each set of OSPR data depicts a piece of chatter information. Compared with the OSPR data shown in Fig. 1(a), the MSPR data have more samples in each revolution. The MSPR technique allows characterizing the chatter dynamics by consolidating the chatter information in OSPR data. The chatter feature can then be estimated using the MSPR data even in a short-time window. Therefore, the MSPR allows using a short-time window while reducing the detection delay. Consequently, the authors proposed to detect incipient chatter using MSPR data rather than OSPR data used in [33,35].

2.3. Data integration and chatter information consolidation

Data integration and chatter information consolidation are critical for detecting incipient chatter [39,40]. To integrate MSPR data and consolidate the chatter information in the OSPR data, we use the Z-score normalization and mean-centered in the data preprocessing step. First, the MSPR is normalized using the Z-score to improve the data consistency. Then, the mean-centered is used to integrate the normalized MSPR data for consolidating the chatter information in multiple sets of OSPR data. The Z-score normalization enhances the data consistency by scaling the data into a common range [39]. It normalizes the raw data by the mean-centered to standard deviation ratio, as expressed in Eq. (2) [41], where x_i is the ith raw data point, z_i represents the normalized data point, $\overline{X}$ and σ_X are the mean and standard deviation of the raw data X, respectively.

$$z_i = \frac{x_i - \overline{X}}{\sigma_X} \tag{2}$$

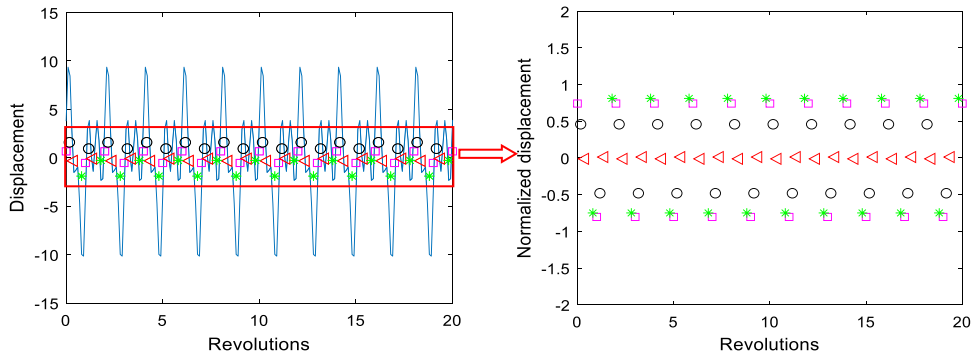
In the milling process, the amplitudes of the vibration signals are different under different machining conditions. So it is challenging to detect incipient chatter reliably due to their different scales and the influence of the noise embedded in the signals. We continue to use the displacement signal, shown in Fig. 1, to introduce the Z-score normalization and mean-centered. Fig. 2(a) displays the MSPR data normalized by the Z-score normalization. The normalized MSPR data are zero-mean and unit variance. Compared with the result shown in Fig. 1(b), one can see that the Z-score normalization reduces the raw data scales influences and removes any offset.

One set of OSPR data cannot characterize the incipient chatter completely since one set of OSPR data has partial chatter information. The mean-centered allows consolidating the chatter information residing in the multiple sets of OSPR data, i.e., MSPR data. Therefore, the mean-centered is proposed to integrate the normalized MSPR data by zero-centering the OSPR data. Fig. 2(b) illustrates the integrated MSPR data. One can find that each set of OSPR data is zero-mean. Compared with one set of OSPR data displayed in Fig. 1(a), the integrated MSPR data have multiple sets of OSPR data on a unified scale. Consequently, the integrated MSPR data characterize the incipient chatter dynamics more accurately than one set of OSPR data.

In addition, we apply the Z-score normalization and mean-centered to a chatter-free signal. Fig. 3(a), (b), and (c) display the MSPR data, normalized MSPR data, and integrated MSPR data on the chatter-free signal, respectively. Each OSPR data set of the chatter-free signal is constant. After the data integration and information consolidation, the integrated MSPR data is zero. The difference between the integrated MSPR data in chatter-free and chatter conditions allows chatter detection.

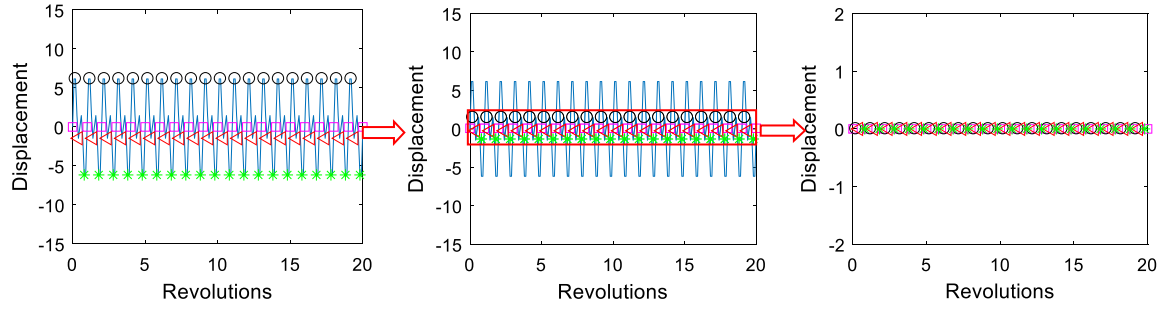
2.4. MaxEnt principle

The maximum entropy (MaxEnt) principle is a flexible and powerful tool for estimating probability distribution. According to Jaynes's theory [42,43], the optimal probability distribution has maximal information entropy under incomplete available information. Thus, finding the maximal information entropy helps



(a) Normalized MSPR data. (b) Integrated MSPR data.

Fig. 2. Normalized MSPR data (a) and integrated MSPR data (b) on the chatter signal.



(a) MSPR data (four sets of OSPR data) (b) Normalized MSPR data. (c) Integrated MSPR data.

Fig. 3. MSPR data (a), normalized MSPR data (b), and integrated MSPR data (c) on the chatter-free signal.

optimize the probability distribution. No assumption is required by the MaxEnt principle to estimate the probability distribution. Therefore, the estimated probability distribution is a maximally unbiased estimation satisfying the available information.

Let us consider the stochastic variable X, and denote x its realization and $f(x)$ its probability density function (PDF). Under certain constraints, the MaxEnt defines the optimal PDF $f(x)$. According to the well-known information entropy theory, Eq. (3) expresses the entropy:

$$H(f(x)) = -\int_R f(x) * \log(f(x))dx. \tag{3}$$

Assuming there are $J+1$ constraints linked to the moments of the stochastic variable X:

$$E(g_j(x)) = \int_R g_j(x) * f(x)dx = \alpha_j, j = 0, 1, 2, \ldots, J, \tag{4}$$

where $g_j(x)$ is a specific function relative to an informational component of X, i.e., $g_j(x) = x^j$, and $E(g_j(x))$ represents the statistical expectation of $g_j(x)$. If $j = 0$ then the zeroth moment $E(g_j(x))$ is the total probability, if $j = 1$ then $g_1(x) = x^1$, $E(g_j(x))$ represents the first-order moment and so forth.

The MaxEnt principle aims to estimate $f(x)$ with no extra assumption by maximizing the information entropy under the given constraint functions of $g_j(x)$ in Eq. (4). Usually, the Lagrange multipliers method allows for solving the MaxEnt optimization problem [42,43].

After estimating $f(x)$ by using the MaxEnt principle, one can calculate the MaxEnt according to Eq. (3). The calculated MaxEnt can be used as a chatter feature. This paper proposes accumulating the information quantity in the sequential calculated MaxEnt feature to obtain the chatter indicator.

2.5. Modified adaptive cumulative chatter indicator

Inspiring from the SPRT theory presented by Wald [44], we proposed an adaptive chatter indicator based on the chatter feature distributions in chatter-free and incipient chatter conditions for incipient chatter detection. The proposed adaptive chatter indicator is robust to variable disturbance noises from the environment by tuning the chatter feature distributions. This chatter indicator allows monitoring and distinguishing sharply chatter-free and incipient chatter conditions.

The proposed adaptive chatter indicator is the "modified adaptive cumulative log-likelihood ratio" (M-ACLLR). According to the SPRT theory, the probability of getting the chatter feature value x_{n-1} in chatter-free condition is $p(x_{n-1}/\text{chatter-free})$, otherwise, this probability is $p(x_{n-1}/\text{chatter})$. The value x_{n-1} represents the realizations of the chatter feature X. The log-likelihood ratio (LLR) defines the logarithm of the likelihood ratio of chatter-free and chatter conditions as follows:

$$\log\Lambda(x_{n-1}) = \log\left(\frac{p(x_{n-1}/\text{chatter})}{p(x_{n-1}/\text{chatter-free})}\right). \tag{5}$$

When a new chatter feature value x_n is obtained, the adaptive cumulative LLR (ACLLR) S_n is calculated as:

$$S_n = \sum_1^n \log\Lambda(x_i) = \log\left(\frac{p(x_1, x_2, \ldots, x_n/\text{chatter})}{p(x_1, x_2, \ldots, x_n/\text{chatter-free})}\right), \tag{6}$$

where $n = 1, 2, 3, \ldots$.

To speed up the incipient chatter detection, we modify the ACLLR S_n to the M-ACLLR S'_n as follows based on the chatter feature distributions to discard the chatter-free detection process:

$$S'_n = \begin{cases} Max\{0, S_n\} & \overline{X}_{\text{chatter-free}} < \overline{X}_{\text{incipient chatter}} \\ Min\{0, S_n\} & \overline{X}_{\text{chatter-free}} > \overline{X}_{\text{incipient chatter}} \end{cases}, \tag{7}$$

where, $\overline{X}_{\text{chatter-free}}$ and $\overline{X}_{\text{incipient chatter}}$ represent the chatter feature's mean values in chatter-free and incipient chatter conditions, respectively.

Implementing the M-ACLLR involves three steps: (1) tuning the chatter feature distributions at the conditions of chatter-free and incipient chatter according to the variable noises; (2) calculating the ACLLR according to the tuned chatter feature distributions; (3) applying the M-ACLLR according to Eq. (7).

The M-ACLLR can accumulate the information in the observed chatter feature data to indicate the incipient chatter. According to the SPRT theory, one can find that the M-ACLLR accumulates each piece of information in the observed chatter feature data sequentially. With the accumulation of the chatter information, the M-ACLLR increases or decreases progressively. In incipient chatter detection, the incipient chatter has a low amplitude, and it is inapparent, especially in heavy disturbance noises where detecting incipient chatter is challenging. One suitable approach is to accumulate the chatter information for making a decision. The M-ACLLR has three major advantages for detecting incipient chatter.

(1) The M-ACLLR is robust to variable disturbance noises. In practice, the disturbance noises vary with changes in uncontrollable influences, such as the measurement instrument and the workpiece's in-homogeneous material [13]. The variation in disturbance noises may lead to a missing or false alarm. The proposed M-ACLLR allows tuning both chatter feature distributions to reduce missing and false alarms. The M-ACLLR allows adapting to the variation in environmental disturbance noises by tuning chatter feature distributions.

(2) Statistically, for detecting incipient chatter, the M-ACLLR is faster than the ACLLR. The latter can be greater or less than zero in incipient chatter detection, i.e., it may locate in the region between zero and the threshold corresponding to the chatter-free condition. For instance, the thresholds corresponding to the chatter-free and incipient chatter conditions are less and greater than zero, respectively. The negative ACLLR data may accumulate, which delays the incipient chatter detection. The M-ACLLR is always greater than or equal to zero, thanks to the barrier of $Max\{0, S_n\}$. Thus, the negative ACLLR data cannot accumulate for detecting incipient chatter. The threshold corresponding to the incipient chatter condition can be crossed more quickly. Consequently, the M-ACLLR is faster than the ACLLR for incipient chatter detection by discarding the chatter-free detection process.

(3) The M-ACLLR is different from the classical indicators that cannot adapt to the variable noises. The classical indicator-based detection approach generally uses the adaptive threshold, which tunes the chatter feature distribution only in chatter-free condition, considers one risk level only, and is determined based on the small probability principle. However, the M-ACLLR coupled with a constant two-risk levels-based threshold can decide on a chatter involving variable disturbance noises. The next part will introduce the two-risk levels-based threshold.

2.6. Two-risk levels-based threshold

As the name suggested, we presented the threshold determined by using two risk levels, i.e., type I and type II errors (α and β) considered commonly in hypothesis testing. The first risk level α represents the probability of accepting a false chatter mode [44]. The second risk level β represents the probability of rejecting a true chatter mode.

Fig. 4 depicts the preset two-risk levels according to the reliability requirement or risk acceptance. The process allows existing chatter faults with the first risk level α when the chatter feature is greater than or equal to the preset upper-value R. The chatter-free process appears with the second risk level β if the chatter feature

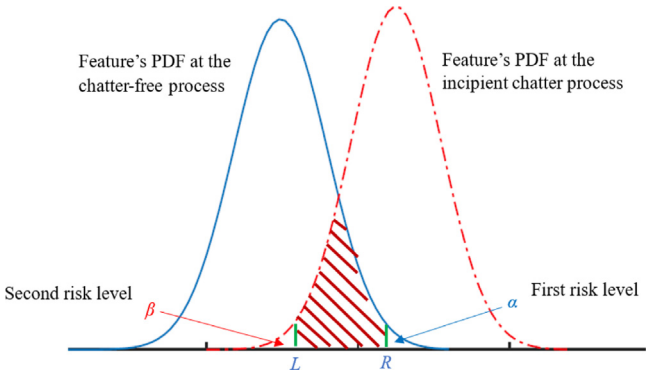


Fig. 4. Two risk levels.

is less than or equal to the preset lower-value L. One sees that the upper-value R and lower-value L determine two risk levels α and β, respectively.

The two-risk levels-based threshold includes the lower and upper thresholds. The required lower, upper, or both thresholds are chosen according to the applications. According to the SPRT theory, both thresholds are calculated based on two risk levels as follows [44,45]:

$$A = \log \frac{\beta}{1-\alpha}, \tag{8}$$

$$B = \log \frac{1-\beta}{\alpha}. \tag{9}$$

Only the upper threshold B is useful for highlighting the chatter condition when the chatter feature's mean value in the chatter condition is great than that in the chatter-free condition, with the following stop criteria:

- Accept chatter process if $S'_n \geq B$,
- Continue monitoring if $S'_n < B$.

Similarly, the M-ACLLR S'_n is $Min\{0, S_n\}$, as shown in (7), if the chatter feature's mean value in chatter condition is less than that in chatter-free condition. Only a lower threshold corresponding to the chatter condition is crossable, with the following stop criteria:

- Accept chatter process if $S'_n \leq A$,
- Continue monitoring if $S'_n > A$,

Thus, one can decide on the chatter appearance by comparing the M-ACLLR S'_n with the threshold value based on two preset risk levels.

The two-risk levels-based threshold is more reliable than the unique threshold designed based on the first risk level. For instance, such a unique threshold is set according to the 3σ principle [19], as Fig. 5 shows. One can set the unique threshold value T_n as follows:

$$T_n = \overline{X}_n + 3\sigma_n, \tag{10}$$

where, $\overline{X}_n$ and σ_n represent the mean and standard deviation of the chatter feature in chatter-free condition subject to specific disturbance noises, respectively. We decide on chatter appearance if the receiving chatter feature value is greater than or equal to the preset unique threshold value T_n. From Fig. 5, one can find that only the first risk level α is considered in decision-making. The second risk level β is ignored. Then, the unique threshold fails to assess the two risk levels simultaneously and cannot detect incipient fault reliably. The two-risk levels-based threshold can deal with these two risk levels simultaneously and overcome the unique threshold's limitation.

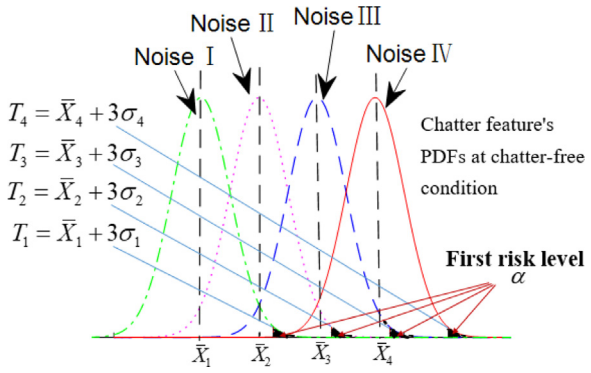


Fig. 5. Unique threshold based on the first risk level.

In addition, the presented threshold is constant and independent of the disturbance noises. In comparison, the unique threshold depends on the disturbance noises and needs to adjust the mean $\bar{X}_n$ and standard deviation σ_n to adapt to the variable disturbance noises.

2.7. Flowchart of the incipient chatter detection method

The reliability and detection delay are essential characteristics of incipient chatter detection methods. On the one hand, the detection delay should comply with the time constraints of the monitored dynamical process. On the other hand, the detection result should not be affected by the machining parameters and the disturbance noises. This subsection proposes a fast and reliable detection method of incipient chatter, as depicted in Fig. 6. To that end, prior knowledge is required about the number of samples per revolution, the time window width, the two risk levels, and the MaxEnt feature's PDFs at chatter-free and incipient chatter conditions. The number of samples per revolution influences the time window width and data acquisition cost. A large number of samples per revolution allows for using a short-time window. However, the large number of samples per revolution leads to high data acquisition costs. Therefore, the number of samples per revolution should be set depending on the application in practice.

The time window width is related to the detection reliability and detection delay. An extended-time window allows a high accuracy of chatter feature extraction, but the detection delay is significant and vice versa. Generally, the incipient chatter arises from the Hopf and period-N bifurcations. In practical terms, we recommend a time window width of 3 to 10 spindle revolutions depending on the number of samples per revolution for the incipient chatter detection. A short-time window is employed when one sets a large number of samples per revolution. Conversely, an extended-time window will be chosen if a small number of samples per revolution is set.

The two risk levels are considered according to the risk acceptance or reliability requirements in practice. For instance, one can set two risk levels α and β to 0.01 conventionally.

The M-ACLLR adapts to the variable disturbance noises by tuning two chatter feature distributions at the conditions of chatter-free and incipient chatter. Determining these two chatter feature distributions is critical for calculating the M-ACLLR. We suggested estimating the chatter feature distributions using the MaxEnt principle based on experimental results.

The proposed detection method relies on six steps depicted in the following:

Step 1: Specify the prior knowledge.

Step 2: Acquire MSPR data using the MSPR technique.

Step 3: Normalize and integrate MSPR data using Z-score normalization and mean-centered:

(a) normalize the MSPR data using the Z-score normalization;
(b) use the mean-centered to integrate the OSPR data in the normalized MSPR data and thus consolidate the chatter information.

Step 4: Extract the chatter feature through the following two substeps:

(a) apply the MaxEnt principle to estimate the probability distribution of the integrated MSPR data;
(b) calculate the MaxEnt of the estimated probability distribution, according to Eq. (3), as a chatter feature.

Step 5: Calculate the M-ACLLR S_n' based on Eqs. (6) and (7).

Step 6: Make a decision by comparing the M-ACLLR S_n' with the upper threshold B calculated according to Eq. (9).

- The incipient chatter is detected if the M-ACLLR S_n' crosses the threshold B.
- Otherwise, the procedure returns to step 2 and continues to detect incipient chatter until the incipient chatter is detected or the milling process is stopped.

3. Dynamic milling model and milling signals generation

We used a dynamic milling model introduced in [37,46] to generate the chatter-free and incipient chatter signals to assess our detection method's effectiveness. Fig. 7 shows the three degrees of freedom (3-DOF) milling dynamic state model. The cutting tool is equivalent to a 2-DOF mass–spring–damper system in the feed (x-) direction and the cross-feed (y-) direction. The workpiece is installed on a specific flexure with a thin-walled structure. The flexure with the workpiece is seen as a single DOF mass–spring–damper system in the y-direction. The flexure's dynamic property in the x-direction is ignored because the flexure's stiffness in the x-direction is far higher than that in the y-direction.

According to machining theory and D'Alembert's principle, the dynamic differential equations of the milling process system are:

$$\begin{bmatrix} m_x & & \\ & m_y & \\ & & m_{wy} \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{y}_w \end{bmatrix} + \begin{bmatrix} c_x & & \\ & c_y & \\ & & c_{wy} \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{y}_w \end{bmatrix} + \begin{bmatrix} k_x & & \\ & k_y & \\ & & k_{wy} \end{bmatrix} \begin{bmatrix} x \\ y \\ y_w \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ F_{wy} \end{bmatrix}, \tag{11}$$

where m_x, m_y, c_x, c_y, k_x, and k_y represent the model mass, damping, and stiffness of the cutting tool in the x- and y-directions, respectively. The variables x, $\dot{x}$, $\ddot{x}$, y, $\dot{y}$, and $\ddot{y}$ describe the displacement, velocity, and acceleration of the cutting tool in the x- and y-directions, respectively. F_x and F_y describe the variable cutting forces acting on the cutting tool in the x- and y-directions, respectively, which are modeled based on the regeneration effect [25,47]. The model parameters m_{wy}, c_{wy}, and k_{wy} stand for the mass, the damping, and the stiffness of the flexure with the workpiece in the y-direction, respectively. The variables y_w, $\dot{y}_w$, and $\ddot{y}_w$ are the workpiece's displacement, velocity, and acceleration in the y-direction, respectively. F_{wy} represents the variable cutting force acting on the workpiece in the y-direction, also known as the F_y's reacting force.

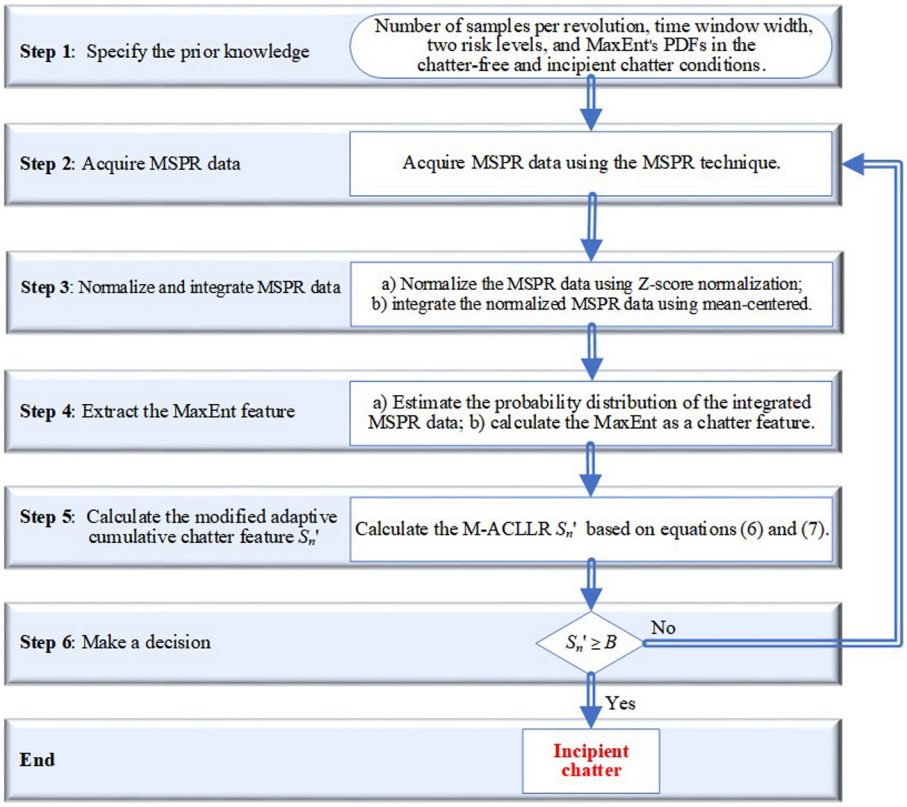


Fig. 6. Flowchart of the proposed detection method.

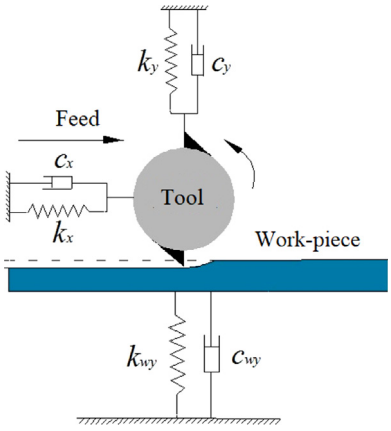


Fig. 7. Dynamic milling model with three degrees of freedom (3-DOF).

Table 3
Equivalent parameters of the cutting tool head.

	Natural frequency (Hz)	Stiffness (N/s)	Damping ratio (%)
x-direction	1055	4.2×10^7	4.5
y-direction	1055	4.2×10^7	4.5

The selected cutting tool is a single-flute end mill with 30 degrees helix angle and 19.1 mm diameter. Table 3 shows the natural frequency, stiffness, and damping ratio of the cutting tool head in the *x*- and *y*-directions [46]. The material of the workpiece is 6061-T6 aluminum.

There are several estimation or calibration methods of the cutting force coefficients [48–50]. According to [46,49], the cutting forces F_x and F_y can be expressed as follows:

$$F_x = (k_{tc}bh(t) + k_{te}b)\cos(\varphi) + (k_{nc}bh(t) + k_{ne}b)\sin(\varphi), \quad (12)$$

$$F_y = (k_{tc}bh(t) + k_{te}b)\sin(\varphi) - (k_{nc}bh(t) + k_{ne}b)\cos(\varphi), \quad (13)$$

where k_{tc}, k_{nc}, k_{te}, and k_{ne} are the cutting force coefficients, the subscript t or n represents the tangential and normal directions, c or e represents the cutting or edge effects; b is the axial cutting depth, $h(t)$ is the instantaneous chip thickness. A dynamometer is used to measure the cutting force, of which coefficients can be estimated using linear regression based on the average measured cutting forces (see [49] for details). Table 4 shows the cutting force coefficients estimated in [46].

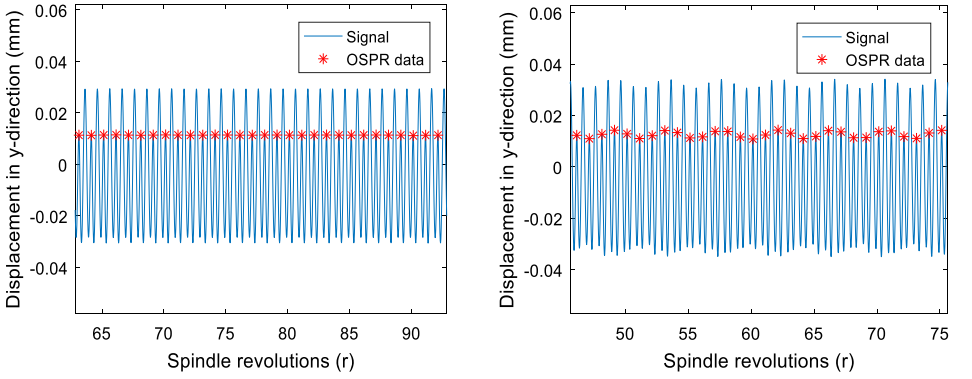
Table 4
Cutting force coefficients.

Cutting force coefficients	Value (N/m^2)
k_{tc}	792×10^6
k_{nc}	352×10^6
k_{te}	26×10^3
k_{ne}	28×10^3

According to the cutting force model, the cutting force coefficients affect the amplitude rather than the frequencies of the cutting force. Nevertheless, the cutting force coefficients may affect the vibration signals' frequencies. For instance, the increase of the cutting force results in an increase in the vibration amplitude of the cutting tool or workpiece. The tooth may not contact the workpiece in its cutting period [51]. In that case, the cutting force coefficients affect the vibration signals' frequencies. No matter how big the cutting force coefficients are, the vibration signals' frequencies are different in chatter-free and incipient chatter stages. From that point of view, the cutting force coefficients cannot affect the performance of the proposed detection method.

Table 5
Machining conditions and flexure dynamics in the *y*-direction.

	Machining conditions			Flexure dynamics in the y-direction		
	Spindle speed (rpm)	Axial depth (mm)	Radial depth (mm)	Natural frequency (Hz)	Stiffness (N/s)	Damping ratio (%)
Chatter-free I	3,800	2.00	5.00	163.00	5.6×10^6	1.08
Incipient chatter I	3,800	2.80	5.00	163.00	5.6×10^6	1.08
Chatter-free II	15,100	6.10	4.50	450.89	2.8×10^7	1.25
Incipient chatter II	15,100	7.70	5.50	450.89	2.8×10^7	0.85
Chatter-free III	16,000	7.20	4.50	466.72	3.0×10^7	1.15
Incipient chatter III	16,000	8.00	4.50	466.72	3.0×10^7	0.72
Chatter-free IV	20,000	6.00	5.80	580.50	3.2×10^7	1.28
Incipient chatter IV	20,000	7.00	5.80	580.50	3.2×10^7	0.78



(a) Displacement in chatter-free condition (b) Displacement in incipient chatter condition

Fig. 8. Displacement responses and OSPR data without the disturbance noise (16,000 rpm).

The machining conditions and flexure dynamics are essential factors causing the incipient chatter. Table 5 shows the chosen machining conditions and flexure dynamics in the *y*-direction to generate chatter-free and incipient chatter signals. The feed per tooth is set to 0.10 mm/tooth. It should be noted that we chose one low-speed milling condition for assessing the effectiveness of the proposed detection method. The assessment results show that our method can apply to the low-speed milling process.

The signals generated using this dynamic milling model can be used to assess the effectiveness of the proposed detection method. We simulated the workpiece dynamic responses using the numerical simulation algorithm [49]. The initial transient signal was discarded to avoid influencing the proposed detection method's assessment. Fig. 8 shows the displacement responses of the workpiece in the *y*-direction at 16,000 rpm. The red stars display the OSPR data points. According to the displacement shown in Fig. 8(a), the OSPR data points are steady, which indicates that the milling process is stable. In contrast, according to the displacement shown in Fig. 8(b), the plot of the OSPR data fluctuates slightly instead of being steady, indicating incipient chatter.

Fig. 9 displays the frequency spectrums of the workpiece's displacement responses in the *y*-direction in chatter-free (a) and incipient chatter (b) conditions, respectively. The frequency spectrum of the chatter-free process has only the spindle's rotation frequency and the corresponding harmonics, as shown in Fig. 9(a). There are no chatter frequencies. Conversely, the frequency spectrum in incipient chatter condition has the spindle's rotation frequency and its corresponding harmonics as well as the chatter frequency, as shown in Fig. 9(b). The small amplitude of the chatter frequency indicates the incipient chatter.

In practice, the disturbance noises caused by uncontrollable influences, such as measuring instrument error and inhomogeneous material of the workpiece, are inevitable. We simulated the disturbance noise influence by adding white noise to the previous signal. The SNR (signal-to-noise ratio) is set to 10 dB for the chatter-free and incipient chatter signals. Fig. 10 illustrates the simulated chatter-free (a) and incipient chatter (b) signals involving the disturbance noises at 16,000 rpm. One can see that the OSPR data in chatter-free condition is similar to that in incipient chatter condition. The disturbance noises reduce their discrimination. It becomes challenging to detect incipient chatter by using the raw signals. These signals are used to assess the effectiveness of the proposed detection method.

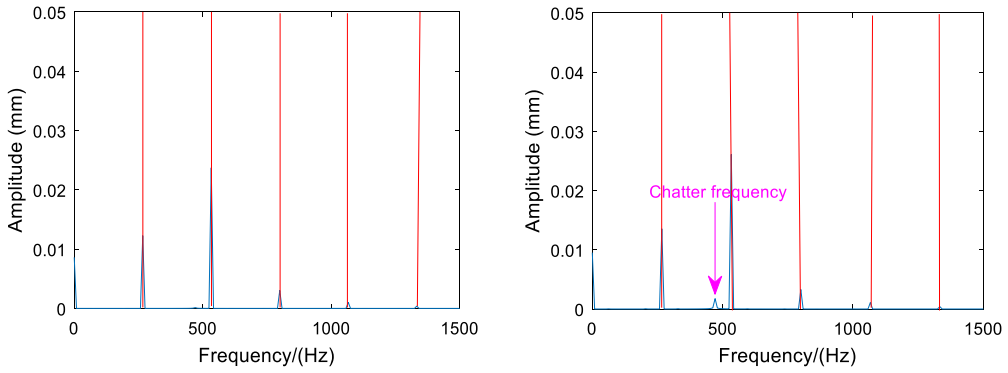
The workpiece is fixed on the flexure which is used to define the system dynamics [52]. The flexure displacement can be measured using a capacitance probe or laser vibrometer. Several kinds of tachometers could help implement MSPR, such as optical encoder, tacho-generator, and magnetic encoder. We suggest using the optical encoder to perform the periodic sampling since it has several advantages, such as safe operation in hazardous environments and insensitivity to electrical and magnetic noise.

4. Results and discussions

We recall that the proposed detection method involves six steps: (1) prior knowledge determination, (2) MSPR data acquisition, (3) MSPR data integration and chatter information consolidation, (4) MaxEnt feature extraction, (5) M-ACLLR chatter indicator calculation, and (6) decision-making using the two-risk levels-based threshold. Now we can assess and discuss the effectiveness of the proposed detection method in the light of these six steps.

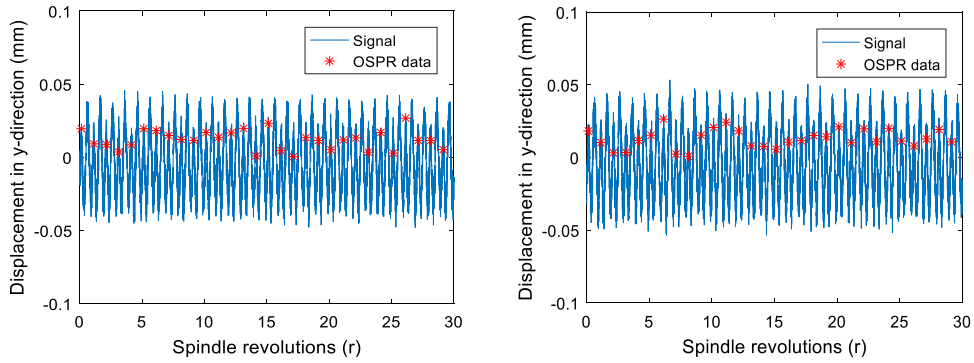
4.1. Prior knowledge determination

The prior knowledge determination is essential to incipient chatter detection. In general, some experimental measurements



(a) Frequency spectrum in chatter-free condition (b) Frequency spectrum in incipient chatter condition

Fig. 9. Frequency spectrums of the displacement responses acquired in the time window width of 30 revolutions.



(a) Displacement in chatter-free condition (b) Displacement in incipient chatter condition

Fig. 10. Displacement responses and OSPR data with a disturbance noise.

or expert knowledge can help to determine prior knowledge. The two risk levels are set to 0.01 for this case study. The threshold can be calculated according to the Eq. (9). Therefore, we set it to log99 in this work.

Setting the number of samples per revolution and the time window width is crucial for incipient chatter detection. As for a specific vibration signal, we tested the influence of these two parameters on the MaxEnt feature. It should be noted that the MaxEnt feature is extracted from the integrated MSPR data based on Eq. (3) by using the MaxEnt principle. To reduce the influence of random characteristics of the disturbance noise, we simulated 100 MaxEnt feature values and then calculated their statistical values.

Fig. 11 shows the MaxEnt feature's standard deviation values on different numbers of samples per revolution in chatter-free and incipient chatter conditions, respectively. The MaxEnt feature's standard deviation values decrease with the increase of the number of samples per revolution. Besides, the decreasing rate also decreases with the increase in the number of samples per revolution. The small MaxEnt feature's standard deviation means the distribution of the MaxEnt feature is tight, and the MaxEnt feature has low randomness. To estimate MaxEnt feature accurately and express the chatter information reliably, we should set a large number of samples per revolution. However, the large number of samples per revolution leads to high data acquisition costs. It is better to set a small number of samples per revolution from this viewpoint. To trade off these two conflicts, we suggest choosing the number of samples per revolution lying between 50 and 150. This work sets the number of samples per revolution to 100.

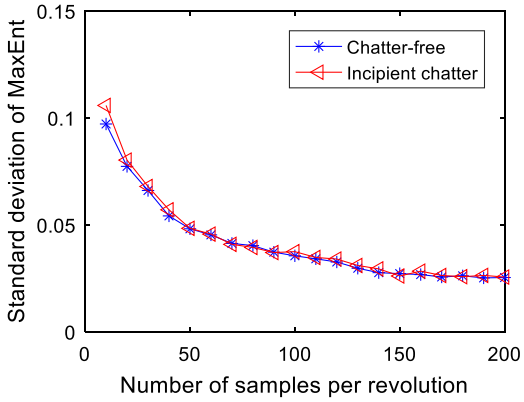


Fig. 11. MaxEnt feature's standard deviation on different numbers of samples per revolution.

Fig. 12 depicts the MaxEnt feature's mean values on different time window widths in chatter-free and incipient chatter conditions. It is difficult to distinguish the chatter-free from incipient chatter on the time window width of two spindle revolutions. The distance between two MaxEnt feature's mean values in chatter-free and incipient chatter conditions is relatively close. MaxEnt feature's mean values increase with the increase of the time window width. The more extended-time window width allows collecting more MSPR data and estimating a more accurate MaxEnt feature. Generally, the MaxEnt feature extracted on a short-time window width is less than that extracted on an extended-time window width since the adding disturbance noise is white noise.

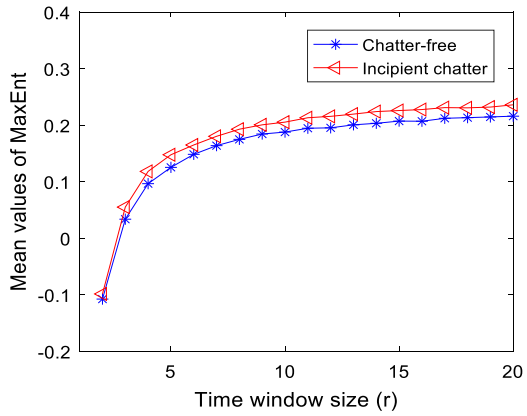


Fig. 12. MaxEnt feature's mean values on different time window widths.

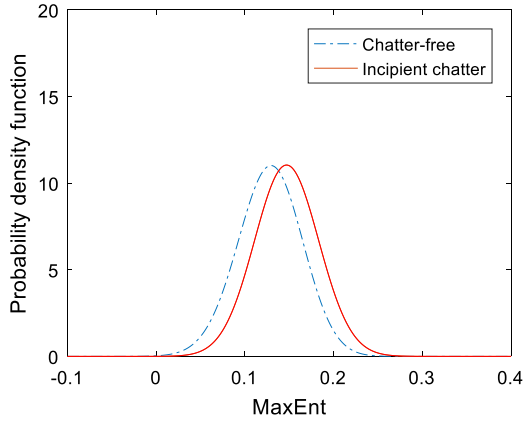


Fig. 13. MaxEnt feature's PDFs.

The distance between the two MaxEnt feature's mean values in chatter-free and incipient chatter conditions is almost constant when the time window width exceeds five spindle revolutions. A short-time window width allows for a short detection delay. Therefore, we set the time window width to five spindle revolutions in this work.

Based on the chosen number of samples per revolution and time window width, one can estimate the MaxEnt feature's PDFs in chatter-free and incipient chatter conditions, as shown in Fig. 13. To estimate the MaxEnt feature's PDFs with high accuracy, we simulate 300 values at each machining condition shown in Table 5. 1200 MaxEnt feature's values are used to estimate the MaxEnt feature's PDFs at the conditions of chatter-free and incipient chatter, respectively.

4.2. MSPR data acquisition

The detection delay is essential for incipient chatter detection and should be shortened as much as possible. The MSPR data reduces the time window width and detection delay due to including multiple sets of OSPR data. This work collects synchronously 100 samples per revolution on the workpiece displacement responses. Thus, the obtained MSPR data include 100 sets of OSPR data, as shown in Fig. 14. Compared with the OSPR data plotted in Fig. 10, much more data are collected in one revolution. A short-time window can acquire data incorporating rich enough information to characterize the chatter dynamics. Therefore, we set the time window width to five revolutions rather than the 25 revolutions used in [33], which has reduced the time window width and detection delay significantly.

Fig. 15 shows the frequency spectrums of the displacement responses obtained on the time window width of five revolutions in chatter-free (a) and incipient chatter (b) conditions. The short-time window leads to a low-frequency resolution, and it is challenging to distinguish the chatter-free and incipient chatter signals due to the low-frequency resolution, the disturbance noise influence, and the weak amplitude of the chatter frequency.

4.3. MSPR data integration and chatter information consolidation

The Z-score normalization and mean-centered were used to integrate the acquired MSPR data and consolidate the chatter information. First, the acquired MSPR data is normalized by using the Z-score normalization. The normalized MSPR data has zero mean and unit variance. Then, the mean-centered is used to zero-center and integrate the OSPR data in the normalized MSPR data. Each set of the integrated OSPR data has a zero mean. Fig. 16(a) and (b) display the integrated MSPR data in chatter-free and incipient chatter conditions, respectively.

4.4. MaxEnt feature extraction

After the MSPR data integration and chatter information consolidation, we need to extract the MaxEnt feature from the integrated MSPR data. The MaxEnt principle was used to estimate the MaxEnt feature of the integrated MSPR data. The MaxEnt feature extraction process includes two steps: (a) estimating the PDF of the integrated MSPR data using the MaxEnt principle; (b) calculating the MaxEnt feature of this estimated PDF according to Eq. (3). The calculated MaxEnt feature is seen as a chatter feature. Due to multiple samples collected per revolution, the calculated MaxEnt feature has relatively high accuracy in a short-time window.

4.5. M-ACLLR chatter indicator calculation

The M-ACLLR chatter indicator was calculated according to the extracted MaxEnt and the preset MaxEnt feature's PDFs in chatter-free and incipient chatter conditions. The calculated M-ACLLR is greater than or equal to zero due to the barrier of $Max\{0, S_n\}$. Figs. 17 and 18 show the decision results obtained respectively at the conditions of chatter-free and incipient chatter. In Fig. 17, all M-ACLLR values are below the threshold in chatter-free condition, indicating the chatter-free milling process. Besides, one observes that some successive M-ACLLR values are equal to zero. This phenomenon, which rarely happens in the ACLLR, commonly occurs in the M-ACLLR since this indicator updates the ACLLR, using $Max\{0, S_n\}$.

Fig. 18 shows the threshold value crossing at the 22nd time window by the M-ACLLR, which indicates the appearance of the incipient chatter. After several information accumulations, the M-ACLLR crosses above the threshold. Due to the information accumulation, the M-ACLLR detects incipient chatter even though the chatter feature is inapparent. The M-ACLLR reduces the detection delay statistically since the $Max\{0, S_n\}$ attenuates the ACLLR's stochastic character. Fig. 18 displays a lower detection delay for the M-ACLLR than that of the ACLLR, which crosses the threshold at the 25th time window. The reason is that the ACLLR accumulates three negative ACLLR values at the beginning. However, in the M-ACLLR calculation process, the first three ACLLRs, which are less than zero, are reset to zero to avoid influence on the incipient chatter detection. Thus, the detection delay is short when one discards the chatter-free detection process. However, if there is no ACLLR less than zero in the detection process, the results of the M-ACLLR and the ACLLR will be identical.

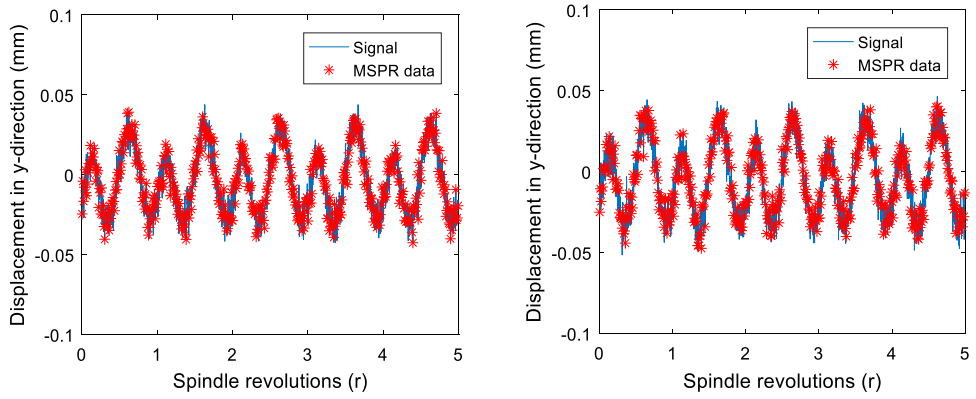


Fig. 14. Displacement and MSPR data with disturbance noises.

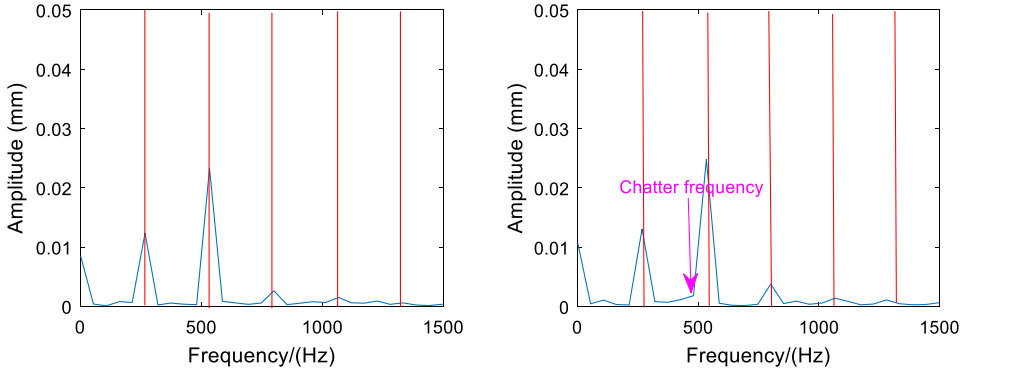


Fig. 15. Frequency spectrums of the displacement responses acquired on the time window width of five revolutions.

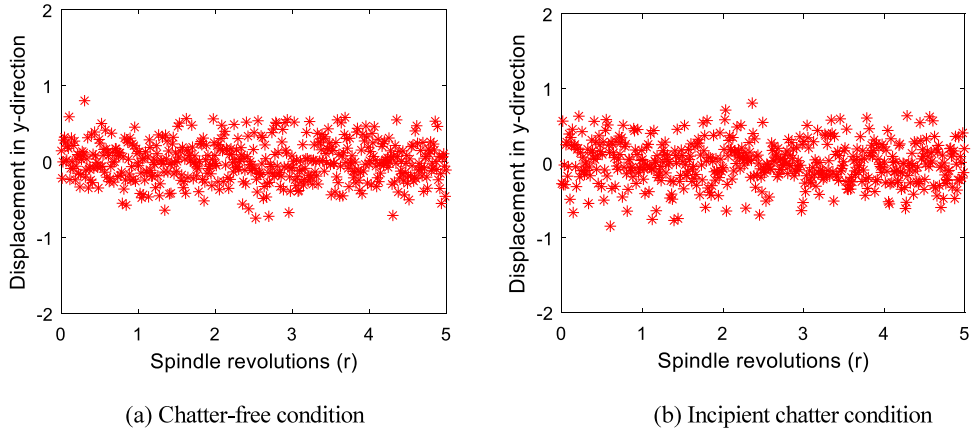


Fig. 16. Integrated MSPR data with disturbance noises.

The different measurement errors and random disturbances in the actual milling process add different disturbance noises to the vibration signal. The variation of the added disturbance noise changes the MaxEnt feature's PDFs in chatter-free and incipient chatter conditions, and reduces the detection reliability. The variation of the added disturbance noises implies an increase or decrease in the signal-to-noise ratio (SNR). For instance, machining more in-homogeneous workpieces leads to an increase in disturbance noise. The disturbance noise increase leads to an increased MaxEnt, and a false alarm could be generated. We assumed the SNR decreases to 9.6 dB from 10 dB to simulate the increase in the disturbance noises. Fig. 19 depicts the false alarm caused by the increase in disturbance noises without the MaxEnt feature's PDFs tuning.

The M-ACLLR can deal with this issue by tuning the MaxEnt feature's PDFs in chatter-free and incipient chatter conditions. The M-ACLLR is calculated according to the tuned MaxEnt feature's PDFs. The calculated M-ACLLR adapts to the variable disturbance noises. Fig. 20 displays the reset MaxEnt feature's PDFs. Compared with the previous MaxEnt feature's PDFs shown in

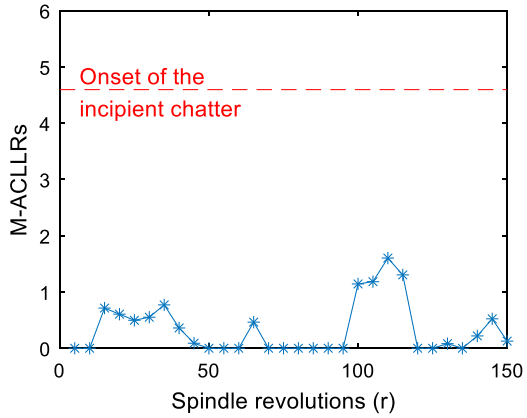


Fig. 17. Decision results of the M-ACLLR in chatter-free condition.

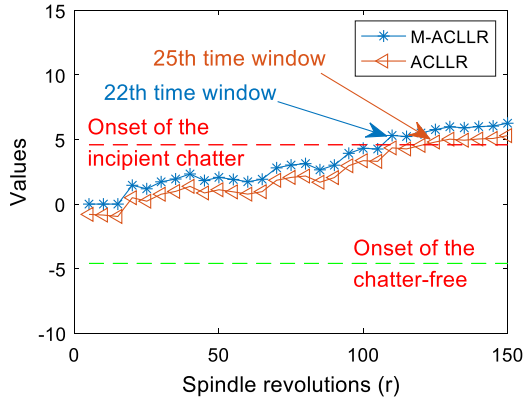


Fig. 18. Decision results of the M-ACLLR and ACLLR in incipient chatter condition.

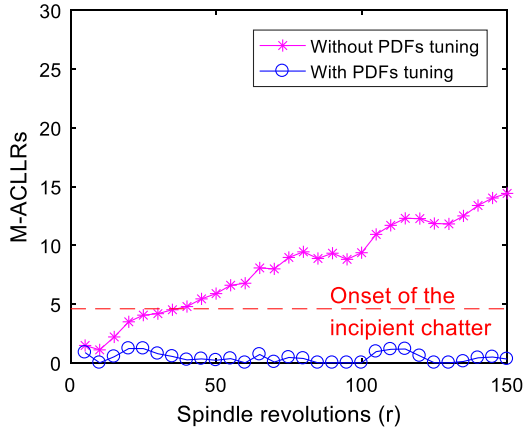


Fig. 19. Detection results of M-ACLLR without and with the MaxEnt feature's PDFs tuning.

Fig. 13, current PDFs are shifted to the right due to the disturbance noise increase. The M-ACLLR avoids the false alarm, as shown in Fig. 19.

4.6. Decision-making using the two-risk levels-based threshold

The two-risk levels-based threshold is used to decide whether the incipient chatter has appeared. The two risk levels are set to 0.01 in this work. The two-risk levels-based threshold is calculated according to Eq. (9), which yields about 4.60. The incipient chatter is detected if the M-ACLLR is greater or equal to the calculated threshold, as shown in Fig. 18. Otherwise, we continued to monitor the incident chatter appearance, as shown in Fig. 17.

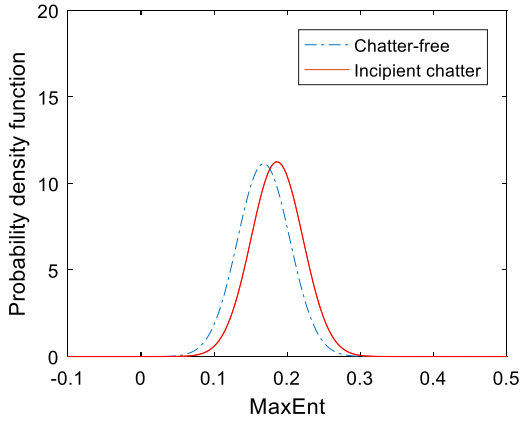


Fig. 20. Turned MaxEnt feature's PDFs in chatter-free and incipient chatter conditions.

The proposed detection method is benchmarked with five detection methods in reliability. Denoted A's first detection method, is based on the OSPR, the MaxEnt, and the SPRT [35]. We denoted B as the second method based on the MSPR, the failure hazard function (FHF), and the unique threshold [30]. In [30], the authors also introduced the failure probability (FP), which can be used as a chatter feature. Therefore, we also chose the detection method based on the MSPR, the FP, and the unique threshold denoted as method C. The proposed F detection method uses the M-ACLLR and the two-risk levels-based threshold. To highlight the advantages of the M-ACLLR and the two-risk levels-based threshold, we also tested two detection methods based on the ACLLR and the unique threshold. These two detection methods are denoted as D and E.

Among the six detection methods, method F has the best detection performance. We have conducted a comparative experimental study of these six detection methods subject to two noise intensities, with the statistical results summarized in Table 6, based on 1200 simulation samples for each case. The time window width was set to 5 revolutions to fairly compare these six detection methods. The false alarm rate, missing alarm rate, and average detection delay are used to assess their detection performance. Values of these three features close to zero indicate high detection performance and vice versa.

Method A cannot detect incipient chatter since the set time window is too short. Only five data points are collected using the OSPR technique. The MaxEnt principle cannot estimate the PDF of the OSPR data and calculate the MaxEnt steadily due to collecting few OSPR data. Methods B, C, and D are based on the unique threshold, which is commonly used in the literature [13,17,19]. These three unique threshold-based detection methods only consider the first risk level and ignore the second risk level. Therefore, they have significantly higher missing alarm rates than method F under both disturbance noises. Besides, these three unique threshold-based detection methods have higher false alarm rates. In general, method F has better performance than these three methods, although the average detection delays of methods C and D are relatively low.

Methods E and F have the same false alarm rate. Nevertheless, method F's missing alarm rate and average detection delay are less than method E. Method F's missing alarm rate is always equal to zero, thanks to the barrier of $Max\{0, S_n\}$ in the M-ACLLR. It means that method F avoids the missing alarm rate completely. The ACLLR does not have this barrier. Therefore, method E cannot

Table 6
Statistical results from 1200 digital experiments of the performance indicators of the six detection methods on a time window width of five spindle revolutions (–: failure to detect incipient chatter).

Detection methods	Noisy environment (SNR)	False alarm rate (%)	Missing alarm rate (%)	Average detection delay (r)
Method A [35]	9.6 dB	–	–	–
(OSPR+MaxEnt+ACLLR+SPRT)	10 dB	–	–	–
Method B [30]	9.6 dB	0.75	98.67	375.00
(MSPR+FHF+unique threshold)	10 dB	1.00	97.50	200.00
Method C [30]	9.6 dB	0.83	96.70	153.84
(MSPR+PF+unique threshold)	10 dB	0.67	96.17	130.43
Method D	9.6 dB	0.50	96.75	153.85
(MSPR+MaxEnt+unique threshold)	10 dB	0.83	96.08	127.66
Method E	9.6 dB	0	0.08	193.55
(MSPR+MaxEnt+ACLLR+two-risk levels-based threshold)	10 dB	0.08	0	181.82
Method F	9.6 dB	**0**	**0**	**162.16**
(Proposed detection method)	10 dB	**0.08**	**0**	**153.85**

Table 7
Statistical results from 1200 digital experiments of the performance indicators of the three detection methods on a time window width of 25 spindle revolutions (–: failure to detect incipient chatter).

Detection methods	Noisy environment (SNR)	False alarm rate (%)	Missing alarm rate (%)	Average detection delay (r)
Method A [35]	9.6 dB	–	–	–
(OSPR+MaxEnt+SPRT)	10 dB	0	0.08	3000.00
Method B [30]	9.6 dB	1.08	91.83	306.12
(MSPR+FHF+unique threshold)	10 dB	1.08	92.92	352.94
Method C [30]	9.6 dB	1.17	91.67	300.00
(MSPR+PF+unique threshold)	10 dB	1.17	89.00	227.27

avoid the missing alarm. Compared with the statistical results of method E, method F also improves the average detection delay by 19.36% and 18.18% under both disturbance noises, respectively. We conclude that the M-ACLLR is better than the ACLLR for incipient chatter detection. Generally, method F has the best performance among the six detection methods.

In methods A, B, and C, the authors recommended that the time window widths lie between 20 and 30 spindle revolutions. We also tested the performance of these three detection methods with the time window width of 25 spindle revolutions. Table 7 shows the tested results.

Methods A, B, and C with the time window width of 25 spindle revolutions show worse performance than method F with five spindle revolutions. Since 25 samples are collected in a time window, method A can detect incipient chatter in 10 dB condition. However, its average detection delay is long. Besides, method A cannot detect incipient chatter in 9.6 dB condition due to relatively heavy noise. A set of OSPR data depicts only one piece of chatter information. Therefore, the chatter feature extracted from OSPR data cannot characterize the incipient chatter dynamics reliably. The PDFs of the chatter feature extracted from the OSPR data are very close to the chatter-free and incipient chatter conditions. Although method A's false alarm rate is zero in the experimental test, it does not mean that method A avoids the false alarm rate. According to the introduction of the SPRT, method A has a specific false alarm rate since the second risk level is not set to zero in the decision-making process. It should be noted that method A may not reliably detect the incipient chatter related to the period-2 bifurcation, as shown in Fig. 1. In contrast, five other methods based on the MSPR technique can detect incipient chatter related to the period-2 bifurcation. The MSPR coupled with the data normalization in method F reduces the detection delay significantly and improves detection reliability. Additionally, the M-ACLLR in method F reduces the detection delay statistically and is more robust to variable disturbance noises, resulting in quick detection and enhanced detection reliability.

Calculating the MaxEnt feature uses a lot of the running time. Hence, real-time incipient chatter detection must take account of the time complexity of the MaxEnt principle. The time complexity of the MaxEnt principle with the first four moments was investigated by Zhao et al. [30]. In the MaxEnt principle's calculation process, the univariate Simpson's method and Newton's method are commonly used for the integrals and iteration, respectively. The univariate Simpson's method has a $O(q)$ time complexity, q being the number of subdivisions inside the integration's upper and lower bounds. The desired accuracy for Newton's method depends on the initial approximation and the applications. The time complexity of the MaxEnt principle is acceptable for implementing real-time chatter detection [30]. Mostly, the computational cost of the MaxEnt principle is relatively high when used for complex multi-variables models [53]. The computational cost of the MaxEnt principle is relatively low for models involving very few variables, peculiarly mono-variable models. Only one variable is considered for calculating the MaxEnt feature in this work. Therefore, the proposed detection method can be used for incipient chatter detection in real-time.

To highlight further the low computational cost of the proposed detection method, we have compared its computational cost to that of four existing methods proposed by Dun et al. [27], Fu et al. [17], Wang et al. [32], and Liu et al. [54]. It should be noted that the computational cost of the chatter feature extraction is only considered in the work of Dun et al. [27] since this chatter detection method is used to diagnose the chatter stability rather than detect chatter in real-time. We obtained the computational costs on one time window of five revolutions on an Intel Core i5 9400 2.9 GHz CPU desktop computer. Table 8 shows the computational costs for the different investigated detection methods.

The computational cost of the proposed method is about 0.02 s, which is less than all computational costs of the other four existing detection methods. So, the proposed detection method has a low computational cost and can detect incipient chatter in high-speed milling applications in real-time.

Few chatter frequencies appear in the chatter signal in incipient chatter condition. Nevertheless, several chatter frequencies could arise in the chatter signal with the increase in the chatter

Table 8
Computational costs of detection methods.

Detection methods	Computational costs
Method of Dun et al. [27]	437.81 s
Method of Fu et al. [17]	3.23 s
Method of Wang et al. [32]	2.23 s
Method of Liu et al. [54]	0.30 s
The proposed method	**0.02 s**

severity. The proposed method can also detect incipient chatter with several chatter frequencies. According to Refs. [31,36], Eq. (1) approximately expresses the measured milling vibration signal in a steady-state case. The chatter frequencies are different from the tooth passing frequency and the corresponding harmonics. The chatter frequencies enable leading to the fluctuation of the MSPR data. The proposed detection method can detect incipient chatter based on the fluctuation of the MSPR data.

The proposed detection method could be applicable to the processing of the rigid workpiece. No matter where the chatter comes from, the characteristic of the chatter is the same. The chatter frequencies are different from the tooth passing frequency and its harmonics. Zhao et al. [30,33] proposed two incipient chatter detection methods according to this chatter characteristic. The MSPR and OSPR techniques are used to acquire data, respectively. MSPR or OSPR data distributions are different in chatter-free and incipient chatter conditions. Their detection methods are verified by using the incipient chatter signals that appear in the processing of the rigid workpiece. Therefore, the proposed detection method can also be used to process the rigid workpiece.

Many dynamic systems, such as rotor-bearing systems and pulse energy conversion systems, are subject to the self-excited vibration phenomena, which often refer to abnormal behaviors. In rotor-bearing systems, the bearing degradation can cause self-excited vibration. Behavior change early detection is essential for diagnosing bearing fault [55]. The proposed detection method can also be used to detect early bearing degradation. According to the literature [56,57], there is an intermittent property in the self-excited vibration in the pulse energy conversion systems. In other words, stable and unstable phenomena appear alternately. Our method could also detect intermittent self-excited vibration in pulse energy conversion systems due to its information accumulation capability.

It should be noted that the proposed method needs to determine the chatter feature distributions in chatter-free and incipient chatter conditions. However, the general chatter detection methods, which use the unique threshold, require only the chatter feature distribution in chatter-free condition, which constitutes a disadvantage of the proposed detection method.

5. Conclusion

This work proposes a promising detection method of incipient chatter in high-speed milling applications based on the MSPR, two data preprocessing techniques, the M-ACLLR, and the two-risk levels-based threshold. A benchmarking milling process model has allowed assessing the effectiveness of the proposed detection method. The MSPR technique, coupled with two data preprocessing techniques, reduces the detection delay by reducing the time window width and improves the detection reliability by consolidating the chatter information in MSPR data. The M-ACLLR accumulates the chatter information from the chatter feature data for detecting incipient chatter. Compared with the ACLLR, the M-ACLLR shortens the detection delay statistically by discarding the chatter-free detection. Besides, the M-ACLLR is robust to variable disturbance noises, resulting in enhanced detection reliability. The two-risk levels-based threshold avoids the assessment risk of the unique threshold. The proposed detection method can detect the incipient chatter in real-time and has higher reliability, shorter detection delay, and lower computational cost than several existing methods.

Despite all those advantages, the proposed detection method cannot determine unique MaxEnt distributions for all machining conditions. Further work will focus on designing a chatter feature to deal with this issue.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Quintana G, Ciurana J. Chatter in machining processes: A review. Int J Mach Tools Manuf 2011;51:363–76. http://dx.doi.org/10.1016/j.ijmachtools.2011.01.001.
[2] Moradi H, Movahhedy MR, Vossoughi G. Bifurcation analysis of milling process with tool wear and process damping: Regenerative chatter with primary resonance. Nonlinear Dyn 2012;70:481–509. http://dx.doi.org/10.1007/s11071-012-0470-7.
[3] Cheng C, Wang Z, Hung W, Bukkapatnam STS, Komanduri R. Ultra-precision machining process dynamics and surface quality monitoring. Procedia Manuf 2015;1:607–18. http://dx.doi.org/10.1016/j.promfg.2015.09.044.
[4] Yuan L, Pan Z, Ding D, Sun S, Li W. A review on chatter in robotic machining process regarding both regenerative and mode coupling mechanism. IEEE/ASME Trans Mechatronics 2018;23:2240–51. http://dx.doi.org/10.1109/TMECH.2018.2864652.
[5] Schneider U, Drust M, Ansaloni M, Lehmann C, Pellicciari M, Leali F, et al. Improving robotic machining accuracy through experimental error investigation and modular compensation. Int J Adv Manuf Technol 2016;85:3–15. http://dx.doi.org/10.1007/s00170-014-6021-2.
[6] Pan Z, Zhang H, Zhu Z, Wang J. Chatter analysis of robotic machining process. J Mater Process Technol 2006;173:301–9. http://dx.doi.org/10.1016/j.jmatprotec.2005.11.033.
[7] Zhu L, Liu C. Recent progress of chatter prediction, detection and suppression in milling. Mech Syst Signal Process 2020;143:1–37. http://dx.doi.org/10.1016/j.ymssp.2020.106840.
[8] Teti R, Jemielniak K, O'Donnell G, Dornfeld D. Advanced monitoring of machining operations. CIRP Ann Technol 2010;59:717–39. http://dx.doi.org/10.1016/j.cirp.2010.05.010.
[9] Jemielniak K. Commercial tool condition monitoring systems. Int J Adv Manuf Technol 1999;15:711–21. http://dx.doi.org/10.1007/s001700050123.
[10] Yue C, Gao H, Liu X, Liang SY, Wang L. A review of chatter vibration research in milling. Chinese J Aeronaut 2019;32:215–42. http://dx.doi.org/10.1016/j.cja.2018.11.007.
[11] Yao Z, Mei D, Chen Z. On-line chatter detection and identification based on wavelet and support vector machine. J Mater Process Technol 2010;210:713–9. http://dx.doi.org/10.1016/j.jmatprotec.2009.11.007.
[12] Ding L, Sun Y, Xiong Z. Early chatter detection based on logistic regression with time and frequency domain features. In: IEEE/ASME Int conf adv intell mechatronics. AIM; 2017, p. 1052–7. http://dx.doi.org/10.1109/AIM.2017.8014158.
[13] Cao H, Lei Y, He Z. Chatter identification in end milling process using wavelet packets and Hilbert-Huang transform. Int J Mach Tools Manuf 2013;69:11–9. http://dx.doi.org/10.1016/j.ijmachtools.2013.02.007.
[14] Zhang Z, Li H, Meng G, Tu X, Cheng C. Chatter detection in milling process based on the energy entropy of VMD and WPD. Int J Mach Tools Manuf 2016;108:106–12. http://dx.doi.org/10.1016/j.ijmachtools.2016.06.002.
[15] Yang K, Wang G, Dong Y, Zhang Q, Sang L. Early chatter identification based on an optimized variational mode decomposition. Mech Syst Signal Process 2019;115:238–54. http://dx.doi.org/10.1016/j.ymssp.2018.05.052.
[16] Liu Y, Wang X, Lin J, Zhao W. Early chatter detection in gear grinding process using servo feed motor current. Int J Adv Manuf Technol 2016;83:1801–10. http://dx.doi.org/10.1007/s00170-015-7687-9.
[17] Wan S, Li X, Chen W, Hong J. Investigation on milling chatter identification at early stage with variance ratio and Hilbert–Huang transform. Int J Adv Manuf Technol 2018;95:3563–73. http://dx.doi.org/10.1007/s00170-017-1410-y.
[18] Tran MQ, Liu MK, Elsisi M. Effective multi-sensor data fusion for chatter detection in milling process. ISA Trans 2021. http://dx.doi.org/10.1016/j.isatra.2021.07.005.

[19] Cao H, Zhou K, Chen X, Zhang X. Early chatter detection in end milling based on multi-feature fusion and 3σ criterion. Int J Adv Manuf Technol 2017;92:4387–97. http://dx.doi.org/10.1007/s00170-017-0476-x.
[20] Cao H, Zhou K, Chen X. Chatter identification in end milling process based on EEMD and nonlinear dimensionless indicators. Int J Mach Tools Manuf 2015;92:52–9. http://dx.doi.org/10.1016/j.ijmachtools.2015.03.002.
[21] Du R, Elbestawi MA, Ullagaddi BC. Chatter detection in milling based on the probability distribution of cutting force signal. Mech Syst Signal Process 1992;6:345–62. http://dx.doi.org/10.1016/0888-3270(92)90036-I.
[22] Wang L, Liang M. Chatter detection based on probability distribution of wavelet modulus maxima. Robot Comput Integr Manuf 2009;25:989–98. http://dx.doi.org/10.1016/j.rcim.2009.04.011.
[23] Schmitz TL, Medicus K, Dutterer B. Exploring once-per-revolution audio signal variance as a chatter indicator. Mach Sci Technol 2002;6:215–33. http://dx.doi.org/10.1081/MST-120005957.
[24] Schmitz TL. Chatter recognition by a statistical evaluation of the synchronously sampled audio signal. J Sound Vib 2003;262:721–30. http://dx.doi.org/10.1016/S0022-460X(03)00119-6.
[25] Honeycutt A, Schmitz TL. A new metric for automated stability identification in time domain milling simulation. J Manuf Sci Eng 2016;138:1–7. http://dx.doi.org/10.1115/1.4032586.
[26] Li K, He S, Li B, Liu H, Mao X, Shi C. A novel online chatter detection method in milling process based on multiscale entropy and gradient tree boosting. Mech Syst Signal Process 2020;135:106385. http://dx.doi.org/10.1016/j.ymssp.2019.106385.
[27] Dun Y, Zhu L, Yan B, Wang S. A chatter detection method in milling of thin-walled TC4 alloy workpiece based on auto-encoding and hybrid clustering. Mech Syst Signal Process 2021;158:1–23. http://dx.doi.org/10.1016/j.ymssp.2021.107755.
[28] Wan S, Li X, Yin Y, Hong J. Milling chatter detection by multi-feature fusion and adaboost-SVM. Mech Syst Signal Process 2021;156:107671. http://dx.doi.org/10.1016/j.ymssp.2021.107671.
[29] Deng C, Miao J, Ma Y, Wei B, Feng Y. Reliability analysis of chatter stability for milling process system with uncertainties based on neural network and fourth moment method. Int J Prod Res 2019;58:2732–50. http://dx.doi.org/10.1080/00207543.2019.1636327.
[30] Zhao Y, Adjallah KH, Sava A, Wang Z. MaxEnt feature-based reliability model method for real-time detection of early chatter in high-speed milling. ISA Trans 2020;113:39–51. http://dx.doi.org/10.1016/j.isatra.2020.07.022.
[31] Mallat S, Hwang WL. Singularity detection and processing with wavelets. IEEE Trans Inform Theory 1992;38:617–43. http://dx.doi.org/10.1109/18.119727.
[32] Wang C, Zhang X, Chen X, Yan R, Wang P. Weak chatter detection in milling based on sparse dictionary. Procedia Manuf 2020;48:839–43. http://dx.doi.org/10.1016/j.promfg.2020.05.121.
[33] Zhao Y, Adjallah KH, Sava A, Wang Z. Early chatter detection using MaxEnt and SPRT. In: 2019 6th Int. conf. control. decis. inf. technol., Paris, France. 2019, p. 1550–5. http://dx.doi.org/10.1109/CoDIT.2019.8820670.
[34] Cao H, Zhang X, Chen X. The concept and progress of intelligent spindles: A review. Int J Mach Tools Manuf 2017;112:21–52. http://dx.doi.org/10.1016/j.ijmachtools.2016.10.005.
[35] Zhao Y, Adjallah KH, Sava A, Wang Z. Online incipient chatter detection based on once-per-revolution sampling and dynamic threshold variant. In: Proc. 2019 10th IEEE Int. conf. intell. data acquis. adv. comput. syst. technol. appl. IDAACS 2019, vol. 2. 2019, http://dx.doi.org/10.1109/IDAACS.2019.8924402.
[36] Chen D, Zhang X, Zhao H, Ding H. Development of a novel online chatter monitoring system for flexible milling process. Mech Syst Signal Process 2021;159:107799. http://dx.doi.org/10.1016/j.ymssp.2021.107799.
[37] Honeycutt A, Schmitz TL. Milling bifurcations: A review of literature and experiment. J Manuf Sci Eng 2018;140:1–19. http://dx.doi.org/10.1115/1.4041325.
[38] Kolokolov Y, Monovskaya A. From modifications of experimental bifurcation diagrams to operating process stability margin. Int J Bifurcation Chaos 2013;23:1–20. http://dx.doi.org/10.1142/S0218127413300243.
[39] Singh D, Singh B. Investigating the impact of data normalization on classification performance. Appl Soft Comput 2020;97:105524. http://dx.doi.org/10.1016/j.asoc.2019.105524.
[40] Singh D, Singh B. Feature wise normalization: An effective way of normalizing data. Pattern Recognit 2022;122:108307. http://dx.doi.org/10.1016/j.patcog.2021.108307.
[41] Urolagin S, Sharma N, Datta TK. A combined architecture of multivariate LSTM with mahalanobis and Z-score transformations for oil price forecasting. Energy 2021;231. http://dx.doi.org/10.1016/j.energy.2021.120963.
[42] Jaynes ET. Information theory and statistical mechanics. Inf Theory Stat Phys 1957;106:620–30. http://dx.doi.org/10.1103/PhysRev.106.620.
[43] Jaynes ET. Information entropy and statistical mechanics II. Phys Rev 1957;108:171–90. http://dx.doi.org/10.1103/PhysRev.106.620.
[44] Wald A. Sequential tests of statistical hypotheses. Ann Math Stat 1945;16:117–86. http://dx.doi.org/10.1007/978-1-4612-0919-5_18.
[45] Tartakovsky Alexander, Nikiforov I, Basseville M. Sequential analysis hypothesis testing and changepoint detection. Boca Raton: CRC Press; 2015, http://dx.doi.org/10.1201/b17279.
[46] Honeycutt A, Schmitz T. A numerical and experimental investigation of period-n bifurcations in milling. J Manuf Sci Eng 2017;139:1–11. http://dx.doi.org/10.1115/1.4034138.
[47] Smith S, Tlusty J. An overview of modeling and simulation of the milling process. J Eng Ind 1991;113:169. http://dx.doi.org/10.1115/1.2899674.
[48] Nouri M, Fussell BK, Ziniti BL, Linder E. Real-time tool wear monitoring in milling using a cutting condition independent method. Int J Mach Tools Manuf 2015;89:1–13. http://dx.doi.org/10.1016/j.ijmachtools.2014.10.011.
[49] Schmitz TL, Smith KS. Machining dynamics: frequency response to improved productivity. New York: Springer; 2009.
[50] Liu XW, Cheng K, Webb D, Longstaff AP, Widiyarto MH, Jiang XQ, et al. Investigation of the cutting force coefficients in ball-end milling. Laser Metrol Mach Perform VI 2003;44:45–54.
[51] Mann BP, Garg NK, Young KA, Helvey AM. Milling bifurcations from structural asymmetry and nonlinear regeneration. Nonlinear Dyn 2005;42:319–37. http://dx.doi.org/10.1007/s11071-005-5719-y.
[52] Mann BP, Insperger T, Bayly PV, Ste G. Stability of up-milling and down-milling, part 2 : experimental verification, 43. 2003, p. 35–40.
[53] Khudanpur S, Wu J. Maximum entropy techniques for exploiting syntactic, semantic and collocational dependencies in language modeling. Comput Speech Lang 2000;14:355–72. http://dx.doi.org/10.1006/csla.2000.0149.
[54] Liu C, Zhu L, Ni C. Chatter detection in milling process based on VMD and energy entropy. Mech Syst Signal Process 2018;105:169–82. http://dx.doi.org/10.1016/j.ymssp.2017.11.046.
[55] Zhou F, Park JH, Wen C, Hu P. Average accumulative based time variant model for early diagnosis and prognosis of slowly varying faults. Sensors 2018;18:1–27. http://dx.doi.org/10.3390/s18061804.
[56] Kolokolov Y, A. Monovskaya. Fractal approach, bifurcation poker and SUC–logic for nonlinear dynamics forecasting. Int J Bifurcation Chaos 2013;23:1–18. http://dx.doi.org/10.1142/S0218127413502015.
[57] Kolokolov Y, Monovskaya A. Estimating the uncertainty of the behavior of a PWM power converter by analyzing a set of experimental bifurcation diagrams. Int J Bifurcation Chaos 2013;23:1–14. http://dx.doi.org/10.1142/S0218127413500636.