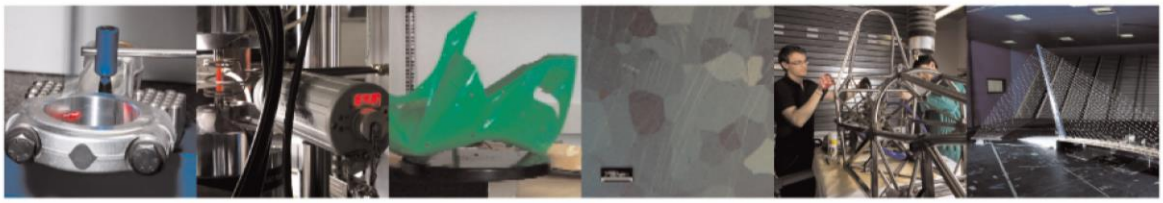
POLITECNICO
MILANO 1863

DIPARTIMENTO DI MECCANICA

mecc



An Automated Approach to Enhance Multi-Scale Signal Monitoring of Manufacturing Processes

Grasso, M.; Colosimo, B.M

This is a post-peer-review, pre-copyedit version of an article published in JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING, 138/5, on November 16, 2015. The final authenticated version is available online at: https://doi.org/10.1115/1.4031797

https://asmedigitalcollection.asme.org/manufacturingscience/article/doi/10.1115/1.4031797/376801/An-Automated-Approach-to-Enhance-Multiscale-Signal

This content is ASME © provided under CC BY-NC-ND 4.0 license

DIPARTIMENTO DI MECCANICA ■ POLITECNICO DI MILANO
via G. La Masa, 1 ■ 20156 Milano ■ EMAIL (PEC): pecmecc@cert.polimi.it
http://www.mecc.polimi.it
Rev. 0

An Automated Approach to Enhance Multi-Scale Signal Monitoring of Manufacturing Processes

Marco Grasso[1] and Bianca Maria Colosimo

Dipartimento di Meccanica, Politecnico di Milano

Via La Masa 1, 20156 Milan (Italy)

marcoluigi.grasso@polimi.it; biancamaria.colosimo@polimi.it

Abstract — Multi-scale signal decomposition represents an important step to enhance process monitoring results in many manufacturing applications. Empirical Mode Decomposition (EMD) is a data driven technique that gained an increasing interest in this framework. However, it usually yields an-over decomposition of the signal, leading to the generation of spurious and meaningless modes and the possible mixing of embedded modes. This study proposes an enhanced signal decomposition approach that synthetizes the original information content into a minimal number of relevant modes via a data-driven and automated procedure. A criterion based on the kernel estimation of density functions is proposed to estimate the dissimilarities between the intrinsic modes generated by the EMD, together with a methodology to automatically determine the optimal number of final modes. The performances of the method are demonstrated by means of simulated signals and real industrial data from a waterjet cutting application.

Keywords: Empirical Mode Decomposition, Multi-scale Analysis, Sensor Signals, Process Monitoring

1 Introduction

Several manufacturing processes are characterized by operation cycles that involve transients, non-stationary fluctuations and phenomena having different dynamics. The result is that sensor signals acquired during this kind of processes often exhibit a multi-scale nature, i.e., a superimposition of salient features belonging to different time-frequency scales. Jin and Shi [1 - 2] discussed some manufacturing applications characterized by multi-scale patterns. Those authors showed that a proper multi-resolution characterization of the natural signal pattern could enhance the process monitoring performances, because different faults may have different effects on distinct scales. However, an open issue in the application of multi-scale analysis techniques regards the achievement of a synthetic decomposition into a minimal number of physically

[1] Corresponding author; Tel.: (+39) 0523-623190, Fax: (+39) 0523-645268

relevant and interpretable oscillation modes. In some cases, the number of scales can be imposed *ex-ante*, based on some engineering knowledge [3], but the resulting modes not necessarily represent the actual intrinsic modes of the signal. More often, the signal is first decomposed into a (possibly large) number of scales, and then one selects a subset of modes of interest [4 – 6], but this yields an information loss about the process. Furthermore, mode selection may be a troublesome task in practise.

The Empirical Mode Decomposition (EMD) technique proposed by Huang *et al*. [7] is a nonparametric and adaptive method for signal decomposition that gained an increasing interest in many process monitoring and machine diagnostics applications [6, 8 – 13]. It allows decomposing any signal into a number of Intrinsic Mode Functions (IMFs) that capture the signal pattern on different scales. Thanks to its data-driven nature, it is suitable to decompose any kind of signal without a prior basis selection, and it does not require any integral transform nor convolution operation. Despite of its benefits, the number of final modes generated via EMD is not controllable, and the algorithm often yields an over-decomposition of the signal. In addition, one intrinsic mode may be split into two (or more) adjacent IMFs, or different scales may be superimposed into a single IMF: such a problem is known as "*mode mixing*" effect [14], which is influenced by the intermittence and noise properties of the signal. Different approaches have been proposed to cope with the mode mixing problem. Wu and Huang [14] proposed a method called Ensemble EMD (EEMD), which consists of defining the true IMFs as the mean of an ensemble of trials, each resulting from adding a white noise of finite amplitude to the original signal. The EEMD method was proposed also in the frame of condition monitoring applications [15 – 16]. The main limitation of EEMD is the computational cost, as it requires the computation of a sufficient number of ensemble trials. A more efficient variant of the EEMD was proposed by Zhang *et al*. [17], but the computational cost is still considerably higher than the basic EMD, and this makes EEMD-based monitoring methods poorly attractive as far as in-process applications are concerned.

The intermittency test is another method proposed to deal with mode mixing [7], which involves the prior selection of a threshold to decide whether a waveform has to be included into a specific IMF or not. However, the subjective nature of the criterion, which compromises the data-adaptivity of the EMD method, actually imposes some limits to its practical use [14].

The lack of automated and data-driven procedures to generate a synthetic signal decomposition into few modes that capture only relevant features motivates this study. The aim consists of enhancing the multi-scale characterization of sensor signals via an automated approach based on

EMD. We propose a novel approach to automatically convert the original IMFs into a minimal number of so-called Combined Mode Functions (CMFs). The method consists of combining together adjacent IMFs characterized by high similarity degree in terms of their probability density function. An increasing number of CMFs is iteratively evaluated and the final decomposition is found by maximizing an optimality index. The term "optimality" refers to the best compromise between the synthesis (minimal number of CMFs) and the physical meaning (actual separation of distinct modes) of the final decomposition. The similarity criterion based on density functions is thought to avoid the use of synthetic indexes that may capture just a portion of the IMF-related information and to better characterize the signal pattern in distinct modes. The method is suitable for a wide range of applications where the goal is to separate transient modes from noise terms and baseline oscillations. It may be applied in a fully data-driven way and hence it is appropriate for in-process monitoring applications.

This paper extends the previous study of Grasso *et al*. [18], by introducing a novel signal decomposition methodology, aimed at enhancing the separation of transient modes from the signal baseline and noise components. The proposed approach is demonstrated by means of both synthetic signals and real data from a waterjet cutting application, where the water pressure signal, which is a suitable source of information for on-line monitoring purposes, exhibits a multi-scale pattern.

Section 2 reviews the EMD and CMF methodologies; Section 3 describes the proposed approach; Section 4 presents a simulation study for the analysis of performances and the comparison with possible alternative approaches; Section 5 presents the application of the proposed method to water pressure signals acquired in a waterjet cutting application; Section 6 eventually concludes the paper.

2 EMD and CMF methodologies

The EMD is a flexible and nonparametric alternative to the multi-resolution methods that rely on the prior definition of a basis function, e.g., wavelet analysis [3]. Thanks to its data-adaptive nature, the EMD is particularly suitable for in-process use, where an automated processing of sensor signals is required. It is a data-driven and adaptive method that requires neither any integral transform nor the definition of any basis function. The IMFs are determined by the signal itself and they work as basis functions. In addition, some authors pointed out that the EMD may yield a finer time-frequency resolution, and hence a better multi-scale variability characterization than other competitor techniques [3]. Let $\boldsymbol{Y}_j = [Y_{j,1}, Y_{j,2}, \dots, Y_{j,p}]^T$ be a waveform, hereafter denoted by

the term "*profile*", which characterizes a signal pattern that repeats over time, such that $j = 1,2, \ldots$ is the index of the repeating profile and p is the number of datapoints in each profile. An example of signal profile data can be found in [1 – 2], where $\boldsymbol{Y}_j$ consists of a repeating tonnage profile acquired in stamping operations. Without loss of generality, a constant size, p, for all the acquired profiles is assumed (if required, synchronous re-sampling may be applied to achieve signal profiles of equal length). The EMD exploits the so-called "*sifting*" algorithm [7] to decompose each signal profile, $\boldsymbol{Y}_j$, into a number n_j of IMFs and a residual term:

$$\boldsymbol{Y}_j = \sum_{i=1}^{n_j} \boldsymbol{c}_{i,j} + \boldsymbol{r}_{n_j,j} \tag{1}$$

where $\boldsymbol{c}_{i,j}$ is the i[th] IMF and $\boldsymbol{r}_{n_j,j}$ is the residue obtained after extracting n_j IMFs, for $j = 1,2, \ldots$.

The sifting algorithm is briefly reviewed in Appendix A.

Due to the data-adaptive nature of the sifting algorithm, the number n_j of extracted IMFs can not be controlled or imposed a-priori. Such a number is usually much larger than the actual number of embedded signal modes. Physically meaningless IMFs are often generated in the low frequency region, because the baseline is generally split into different IMFs before the last mode is classified as a residue. In addition, the possible occurrence of the mode mixing problem further complicates the interpretation of the resulting decomposition. Different authors investigated the problem of selecting a subset of relevant IMFs [5 – 6, 19 – 20]. The proposed approaches rely on synthetic indexes, including the average value [19], the energy [5], the correlation between the IMF and the original signal [4], the peak frequency [20], and other indexes [21]. Nevertheless, the IMF selection operation is affected by some limitations: (i) the reliability of the result is strongly affected by the mode mixing effect, (ii) the choice of the most suitable synthetic index is a problem-dependent task, and different indexes may lead to different decisions, and (iii) the IMF selection operation implies a potential information loss. Problems (i) and (ii) complicate the isolation of desired time-frequency patterns, especially when the signal is processed and analysed over successive time intervals. Moreover, the IMF selection is usually applied when there is only one subset of relevant IMFs, e.g., in low-pass, high-pass or band-pass filtering applications.

A more interesting and effective approach is based on the CMF approach [6]. It allows one to cope with the over-decomposition produced by the sifting algorithm and the splitting of intrinsic modes into multiple IMFs. Instead of selecting a set of relevant IMFs, the computations of CMFs allows one to synthetize the signal decomposition into fewer final modes, which are expected to better represent the embedded information content.

The method consists of combining adjacent IMFs $\boldsymbol{c}_{i,j}, \boldsymbol{c}_{i+1,j}, \dots, \boldsymbol{c}_{i+q_k,j}$ to obtain a new CMF, $\boldsymbol{c}_{s_k,j}$, as follows:

$$\boldsymbol{c}_{s_k,j} = \boldsymbol{c}_{i,j} + \boldsymbol{c}_{i+1,j} + \cdots + \boldsymbol{c}_{i+q_k-1,j}, \quad j = 1,2,\dots; i \in [1, n_j] \tag{2}$$

where q_k is the number of IMFs combined into the k^{th} CMF, $k = 1, \dots, K$, being $1 \leq q_k \leq n_j$ and $K \leq n_j$. Such a combination of subsets of IMFs can be interpreted as a new adaptive filter bank, which is based on the intrinsic time scales of the signal, resulting in an accuracy increase of the EMD [6]. Generally speaking, the aim is to find a procedure that generates the same number K of CMFs for every realization of the random profile $\boldsymbol{Y}_j$. If the variability of K can not be reduced to zero, the symbol K_j should be used in place of K. However, for simplicity of notation, the symbol K without the subscript j will be used in this study. A discussion on the variability of K is presented in Sections 4 and 5.

Grasso *et al*. [18] showed that the CMF methodology may be exploited to design in-process monitoring techniques and, at the same time, to achieve a better characterization and interpretation of fault effects on different scales. Nevertheless, the critical issue consists of a lack of automated procedures to select the number K of CMFs and to determine which IMFs should be summed up together in each CMF. Gao *et al*. [6] proposed a criterion based on local frequency changes captured by the instantaneous frequencies of IMFs. It is aimed at identifying adjacent IMFs that share similar instantaneous frequency properties, but it relies on visual inspection of IMF patterns.

Index-based methods like the ones used for IMF selection might be used to determine the similarity between adjacent IMFs, but the resulting approach would be strongly problem-dependent. The proposed methodology is aimed at overcoming the limitations of IMF selection methods mentioned above by exploiting a more complete characterization of the IMF-related information via probability density estimation.

3 The proposed approach

The proposed approach consists of solving two inter-related problems: (i) the determination of the IMFs that should be combined together into one single CMF, and (ii) the choice of the number of CMFs that should be generated in order to achieve a synthetic representation of relevant signal modes.

The former problem requires the definition of a criterion to decide whether two (or more) adjacent IMFs share the same properties and hence they represent split components of a single mode. The

latter problem can be solved by iteratively evaluating different decompositions, corresponding to increasing numbers, K, of CMFs and to maximize (or minimize) an optimality criterion. Such a criterion consists of a compromise between the dissimilarity of modes within the same CMF and the dissimilarity between distinct CMFs. A conceptual scheme of the method is depicted in Fig. 1. The different steps of the proposed method are explained and discussed in the following sub-sections.

INSERT FIG 1 ABOUT HERE

3.1. Kernel density estimation of IMFs

In this study, we focus our analysis on the separation of transient modes from non-transient oscillations and noise terms, which is a challenging task to be dealt with in many industrial applications. To this aim, we propose the estimation of the probability density function of the IMFs, $f(\boldsymbol{c}_{i,j})$, $i = 1, \dots, n_j$ and $j = 1,2, \dots$, to determine to what extent adjacent IMFs are similar to each other. Differently from a few previous studies in the EMD literature [22], our aim is to use the IMF density functions to estimate the similarity between adjacent IMFs rather than the similarity between each IMF and the original signal for filtering purposes.

Let $x(t)$ be a sine wave of amplitude X and frequency $f = 1/T$. The sine wave can be considered a random process such that its probability density function is [23]:

$$\begin{aligned} f(x) &= \left(\pi\sqrt{2\sigma^2 - x^2}\right)^{-1}, |x| < X \\ f(x) &= 0, \qquad\qquad\quad |x| \geq X \end{aligned} \tag{3}$$

where $\sigma = X/\sqrt{2}$ is the standard deviation of the sine wave. Fig. 2 shows a sine wave with amplitude $X = 1$ and the corresponding probability density function.

INSERT FIG 2 ABOUT HERE

The density function $f(x)$ reaches a minimum at the mean value of the wave, whereas it reaches its maximum value at the peak values of the wave. Such a pattern characterizes any oscillation with constant amplitude over time, and it is considerably different from the probability density function of a narrow-band noise. In addition, it is also considerably different from the density of transient oscillations that are localized in time. Let $x(t)$ be a sine wave of amplitude X that is localized in time, i.e., $x(t) > 0$ over a narrow portion of the overall time window and $x(t) = 0$

otherwise. The corresponding density function, characterized by a high "peakedness", is shown in Fig. 3.

INSERT FIG 3 ABOUT HERE

Therefore, when the dissimilarity between adjacent IMFs is defined as the difference between their probability density functions, the capability of separating transient modes from non-transient oscillations and noise components is expected to be maximized. The density function allows detecting differences between IMFs that are related with various synthetic statistics, including the amplitude of the IMFs, their variance, kurtosis and skewness. Because of this, it allows defining an information-richer criterion than the ones based on single indexes. However, the expression of $f(x)$ in Eq. 3 does not depend on the frequency of $x(t)$, and hence the density function is not suitable to separate oscillations with equal amplitude and different frequency. Thus, if the goal consists of separating intrinsic modes that differ only in terms of their frequency, other criteria should be preferred, e.g., the one proposed by Gao *et al*. [6] or one based on the peak frequency of IMFs [20]. The possible coupling of the proposed approach with those criteria may be investigated in future studies.

Since the form of the distribution of different IMFs is not known, a nonparametric density estimation is required [24], which exploits the kernel fitting technique. The basic idea is to estimate the density function $f(x)$ of a random variable x at a point x_i by using the neighboring observations, such that the influence of x_i on the estimate at any x vanishes asymptotically. The methodology is widely used in practice, and there is an extensive literature in this field [25 – 27]. Two relevant issues consist of the choice of the kernel function, $Ker(x)$, and the selection of an optimal kernel bandwidth, h. The latter issue is the most critical one, and several methods have been proposed thus far. One simple approach is to use rule-of-thumb estimates, which are known to approximate the optimal choice in the presence of normal data [27]. However, when strong departures from normality are observed, other methods should be preferred, which are aimed at estimating the bandwidth, h, in a data-driven way. Being unknown the true density function, the most common approach consists of applying selection criteria based on cross-validation [25]. Among them, the one based on unbiased cross-validation (UCV) is probably the most popular and studied one [28 – 29]. The kernel estimator of $f(\boldsymbol{c}_{i,j})$, denoted by $\hat{f}(\boldsymbol{c}_{i,j})$, is given by [29]:

$$\hat{f}(\boldsymbol{c}_{i,j}) = \hat{f}(x) = p^{-1}\sum_{l=1}^{p} h_{i,j}^{-1} Ker\left((x - c_{i,l,j})/h_{i,j}\right), \qquad i = 1, \dots, n_j;\ j = 1,2, \dots \tag{4}$$

where p is the number of datapoints in each IMF and $h_{i,j}$ is the kernel bandwidth of the i^{th} IMF extracted from the signal profile $\boldsymbol{Y}_j$. Notice that, since the probability distribution may considerably change from IMF to IMF, different optimal choices of $h_{i,j}$ can be made, for $i = 1, \dots, n_j$. The essential idea of UCV is to use the bandwidth, $h_{i,j} = \hat{h}$, that minimizes the function:

$$UCV_{i,j}(\hat{h}) = \int \hat{f}_{\hat{h}}(x)^2 dx - 2p^{-1} \sum\nolimits_l \hat{f}_{\hat{h},l}(c_{i,l,j}), \quad i = 1, \dots, n_j;\ j = 1,2, \dots \tag{5}$$

where $\hat{f}_{\hat{h}}$ denotes the kernel estimator based on the choice $h_{i,j} = \hat{h}$, and $\hat{f}_{\hat{h},l}$ denotes the leave-one-out kernel estimator, defined as follows:

$$\hat{f}_{\hat{h},l}(x) = p^{-1} \sum_{\substack{u=1 \\ u \neq l}}^{p} h_{i,j}^{-1} Ker\big((x - c_{i,u,j}) / \hat{h}\big), \qquad i = 1, \dots, n_j;\ j = 1,2, \dots \tag{6}$$

Thus, the optimal choice of $h_{i,j}$ is defined by:

$$h_{i,j} = \hat{h} = arg \min_{h>0} UCV_{i,j}(h), \quad i = 1, \dots, n_j;\ j = 1,2, \dots \tag{7}$$

3.2. Separation of IMFs into distinct CMFs

Let $\boldsymbol{c}_{i,j}$ and $\boldsymbol{c}_{i+1,j}$ be two adjacent IMFs of the signal profile $\boldsymbol{Y}_j$, $j = 1,2, \dots$. Then their dissimilarity can be expressed as follows:

$$D_{i,i+1,j} = \left\| \hat{f}(\boldsymbol{c}_{i,j}) - \hat{f}(\boldsymbol{c}_{i+1,j}) \right\|, \qquad j = 1,2, \dots; i = 1, \dots, n_j - 1 \tag{8}$$

where $\|\cdot\|$ is the Euclidean norm. A decomposition into $K = 2$ CMFs can be achieved by computing the successive distances, $D_{i,i+1,j}$, for $i = 1, \dots, n_j - 1$, and by separating the IMFs into $\boldsymbol{c}_{S_1,j}$ and $\boldsymbol{c}_{S_2,j}$ as follows:

$$\begin{aligned} \boldsymbol{c}_{S_1,j} &= \boldsymbol{c}_{1,j} + \boldsymbol{c}_{2,j} + \cdots + \boldsymbol{c}_{q_1,j} \\ \boldsymbol{c}_{S_2,j} &= \boldsymbol{c}_{q_1+1,j} + \boldsymbol{c}_{q_1+2,j} + \cdots + \boldsymbol{c}_{n_j,j} \end{aligned} \qquad j = 1,2, \dots \tag{9}$$

where $q_1 = arg \max_i D_{i,i+1,j}$, i.e., when the largest distance between adjacent IMFs is found, a separation between two CMFs is introduced. This approach can be iteratively repeated by imposing a further splitting at the argument of the $(K-1)^{\text{th}}$ largest distance $D_{i,i+1,j}$, where $K \leq n_j$ is the number of CMFs. When $K = n_j$ the CMF decomposition coincides with the IMF decomposition.

In most applications, better performances could be achieved by introducing a weight into Eq. 4 to penalize the generation of a CMF that captures meaningless low amplitude modes in the low frequency region. Due to the nature of the sifting algorithm, one (or more) last IMFs capture the

baseline of the signal. However, when such a baseline is absent, few last IMFs may provide irrelevant information, i.e., a quasi-constant level (see example A in Section 4). To avoid the generation of poorly informative CMFs corresponding to these IMFs, we advocate the normalization of the distance $D_{i,i+1,j}$ by a weight $w_{i,j} = i/supp\left(\hat{f}(\boldsymbol{c}_{i+1,j})\right)$, where i is the IMF index and $supp(\cdot)$ denotes the support of the density function:

$$supp\left(\hat{f}(x)\right) = \{x \in \mathbb{R} | f(x) > tol\} \tag{10}$$

being tol a tolerance threshold, e.g., $tol = 1e-6$. The weighted distance is defined as follows:

$$D^*_{i,i+1,j} = \frac{\left\|\hat{f}(\boldsymbol{c}_{i,j}) - \hat{f}(\boldsymbol{c}_{i+1,j})\right\|}{w_{i,j}}, \quad j = 1,2,\dots; i = 1,\dots,n_j - 1 \tag{11}$$

The smaller is the support of the density function and the larger is i, the greater is the penalization of the distance in Eq. 7, and hence the priority of CMFs that include irrelevant higher-order IMFs is reduced.

3.3. Selection of the optimum number of CMFs

The proposed approach for the selection of the optimum number of CMFs is inspired by the cluster validity criteria used in unsupervised learning [24, 30], which allow determining the "correct" number of clusters in a dataset. Analogously to clustering problems, we want to determine a CMF decomposition such that both the similarity of IMFs within each CMF and the dissimilarity between distinct CMFs are as large as possible. The use of a similarity measure based on IMF densities allows applying cluster validity criteria to the EMD framework. Several indexes have been proposed in the clustering literature so far [30], but to the best of the authors' knowledge, no attempt was made to extend those criteria to the EMD methodology. A category of validity criteria rely on the measure of the within-group variance and the between-group variance: they include the *Ball & Hall* index [31], the *Calinski & Harabasz* index [32], the *Hartigan* index [33], and many others [34 – 35]. See Appendix B for a brief review of these indexes and their use. Although the general idea of analyzing the within group and between group variances is applicable to the CMF decomposition problem, the cluster validity criteria proposed in the literature may need to be adapted to the present application. The following approach is proposed: let $SSW_{f,j}(K) = \sum_{k=1}^{K} \sum_{i \in c_{s_k,j}} \left\| \hat{f}(\boldsymbol{c}_{i,j})_k - \overline{\hat{f}_{k,j}}(k) \right\|$ be the sum-of-squares within the K CMFs, where $\overline{\hat{f}_{k,j}}(k)$ is the average density function within the k^{th} cluster, and let $SSB_{f,j}(K) = \sum_{k=1}^{K} n_{k,j}(k) \left\| \overline{\hat{f}_{k,j}}(k) - \overline{\hat{f}_j} \right\|$ be the sum-of-squares between the K CMFs, where $n_{k,j}(k)$ is the number of IMFs in the k^{th} cluster and $\overline{\hat{f}_j}$ is the average density function of the original IMF

decomposition. The statistic $SSW_{f,j}(K)$ is a monotone decreasing function of K, whereas $SSB_{f,j}(K)$ is a monotone increasing function of K. The proposed criterion to determine the best number of CMFs is based on the following inequality:

$$SSW_{f,j}(K) < K * SSB_{f,j}(K), \quad j = 1,2,\dots \tag{12}$$

When $SSW_{f,j}(K) > K * SSB_{f,j}(K)$, the variability within the CMFs (in terms of distances between density functions) is larger than the variability between the CMFs: this means that a further decomposition into $K+1$ CMFs allows separating distinct modes that are currently included into the same CMF. When $SSW_{f,j}(K) < K * SSB_{f,j}(K)$ the desired condition is achieved, i.e., the IMFs in each CMF are close to each other, and different intrinsic modes are separated into distinct CMFs. From a parsimony viewpoint, the multiplication by K is used to penalize the selection of larger number of CMFs. Thus, the following procedure can be designed:

1. Initialize a counter, $c = 1$;
2. Set $K = c$ and decompose the IMFs into K CMFs by using the density-based criterion presented in Section 3.2. Notice that when $K = 1$, all the IMFs are summed up into a single CMF;
3. Compute the values of statistics $SSW_{f,j}(K)$ and $SSB_{f,j}(K)$: if $SSW_{f,j}(K) < K * SSB_{f,j}(K)$ stop the procedure (the final configuration is the one with K CMFs), otherwise set $c = c + 1$ and repeat steps 2 and 3.

4 Examples with synthetic signals

4.1. Application of the proposed approach

The methodology is first demonstrated by means of two examples based on synthetic signals. The two examples are representative of two different multi-scale patterns characterized by a noise term, one or more transient patterns and a low frequency baseline. The first example, hereafter called "Example A", consists of a signal profile $Y_j(t)$ that results from the superimposition of a white noise term, $Y_{n_j}(t)$, a high frequency transient that is localized in time, $Y_{T_j}(t)$, and a stationary sine wave, $Y_{S_j}(t)$. The following model was used:

$$Y_j(t) = Y_{n_j}(t) + Y_{T_j}(t) + Y_{S_j}(t), \qquad j = 1,2,\dots; t \in [0,1]s$$

$$Y_{n_j}(t) \sim N(0, \sigma_n^2) \tag{13}$$

$$Y_{T_j}(t) = A_T f(t; \mu_T, \sigma_T^2) \sin(2\pi F_T t)$$

$$Y_{S_j}(t) = A_S \sin(2\pi F_S t)$$

where $f(t; \mu_T, \sigma_T^2)$ is the normal probability density with parameters μ_T and σ_T^2. The following model parameters were used: $\sigma_n^2 = 25, A_T = 8, \mu_T = 0.5, \sigma_1^2 = 0.01, F_T = 50\ Hz, A_S = 25$, and $F_S = 12\ Hz$. N random realizations of the profiles were generated via Monte-Carlo simulations. Sampled profiles $\boldsymbol{Y}_j, j = 1,2, \dots, N$, were generated by sampling the signal at $f_s = 10\ kHz$.

One realization of the random profile $\boldsymbol{Y}_j$ is shown in Fig. 4.

INSERT FIG 4 ABOUT HERE

Fig. 5 shows the corresponding multi-scale decomposition and the density functions estimated for each IMF.

INSERT FIG 5 ABOUT HERE

The EMD in Fig. 5 yields a multi-scale decomposition into ten IMFs, which are much more than the modes required to describe the relevant intrinsic scales of the signal. The noise term is decomposed into the first four IMFs, $\boldsymbol{c}_1, \dots, \boldsymbol{c}_4$. The transient pattern is mainly captured by IMFs $\boldsymbol{c}_5$ and $\boldsymbol{c}_6$[2], whereas the sine wave is captured by IMF $\boldsymbol{c}_7$. The last three IMFs, $\boldsymbol{c}_8$, $\boldsymbol{c}_9$ and $\boldsymbol{c}_{10}$, represent a baseline that has no physical meaning and results from splitting the sine wave into more than one mode.

Fig. 6 (left panel) shows the distance values $D_{i,i+1}^*$ computed for $i = 1{:}\,9$, whereas Fig. 6 (right panel) shows the values of $SSW_f(K)$ and $K * SSB_f(K)$ as a function of the number of extracted CMFs, K. In this case, the proposed approach yields the selection of $K = 3$ CMFs that are shown in Fig. 7.

INSERT FIG 6 ABOUT HERE

The automatically selected CMF decomposition consists of $\boldsymbol{c}_{S_1} = \boldsymbol{c}_1$, $\boldsymbol{c}_{S_2} = \sum_{i=2}^{6} \boldsymbol{c}_i$ and $\boldsymbol{c}_{S_3} = \sum_{i=7}^{10} \boldsymbol{c}_i$, since the two largest values of $D_{i,i+1}^*$ correspond to the first and the sixth IMF (Fig. 6, left panel).

Fig. 7 shows that the multi-scale characterization of the signal is considerably enhanced thanks to a synthetic representation of actually relevant modes and a clear separation of the transient pattern

[2] Notice the anomalous splitting of the transient mode, maybe due to a mode mixing effect.

from the baseline sine wave. This is achieved without loss of information and in a fully data-driven way. In spite of a synthetic characterization of intrinsic modes, however, the noise term separation might be further enhanced. Fig. 5 (right panel) shows that the first IMF exhibits a bimodal density function that differs from the density of other noise modes. From Wu and Huang [35] it is known that the EMD of white noise signals yields a first IMF with different spectral properties from the higher order IMFs, and this reflects into a different probability distribution. By isolating the first IMF into the first CMF, $\boldsymbol{c}_{S_1}$, part of the noise is removed from other CMFs, but a residual noise component still remains in the transient mode, captured by $\boldsymbol{c}_{S_2}$. A further improvement of the methodology should be focused on the incorporation of denoising techniques into the CMF decomposition procedure.

INSERT FIG 7 ABOUT HERE

A second synthetic signal, "Example B", is shown in Fig. 8. In this case, the baseline is not a stationary sine wave, but a non-cyclical pattern that originates from the sum of three sine waves. In this case, the signal $Y_j(t)$ results from the superimposition of a white noise term, $Y_{n_j}(t)$, two high frequency transients localized in time having different frequencies, $Y_{T_j}(t)$, and a baseline, $Y_{B_j}(t)$.

INSERT FIG 8 ABOUT HERE

The generating model is:

$$
\begin{gathered}
Y_j(t) = Y_{n_j}(t) + Y_{T_j}(t) + Y_{S_j}(t), \qquad j = 1,2,\dots; t \in [0,1]s \\
Y_{n_j}(t) \sim N(0, \sigma_n^2) \\
Y_{T_j}(t) = A_{T,1} f\left(t; \mu_{T_1}, \sigma_{T_2}^2\right) \sin\left(2\pi F_{T,1} t\right) + A_{T,2} f\left(t; \mu_{T_2}, \sigma_{T_2}^2\right) \sin\left(2\pi F_{T,2} t\right) \\
Y_{S_j}(t) = A_{S,1} \sin\left(2\pi F_{S,1} t\right) + A_{S,2} \sin\left(2\pi F_{S,2} t\right) + A_{S,3} \sin\left(2\pi F_{S,3} t\right)
\end{gathered}
\tag{14}
$$

The following model parameters were used: $\sigma_n^2 = 25$, $A_{T,1} = 3$, $\mu_{T_1} = 0.2$, $\sigma_{T_1}^2 = 1e-4$, $A_{T,2} = 2$, $\mu_{T_2} = 0.8$, $\sigma_{T_2}^2 = 1e-4$, $F_{T,1} = 100\ Hz$, $F_{T,2} = 200\ Hz$, $A_{S,1} = 50$, $A_{S,2} = 25$, $A_{S,3} = 20$, $F_{S,1} = 1.5\ Hz$, $F_{S,2} = 2\ Hz$ and $F_{S,3} = 2.5\ Hz$. The sampling frequency was $f_s = 10\ kHz$. The

multi-scale decomposition of the signal profile in Fig. 8 is shown in Fig. 9, together with the density functions estimated for each IMF.

INSERT FIG 9 ABOUT HERE

In this case, the EMD yields a multi-scale decomposition into nine IMFs. The transient pattern at $F_{T,1} = 100\ Hz$ is split into IMFs $\boldsymbol{c}_4, \boldsymbol{c}_5$ and $\boldsymbol{c}_6$, whereas the transient pattern at $F_{T,2} = 200\ Hz$ is mainly captured by IMF $\boldsymbol{c}_4$. The baseline is the last IMF, $\boldsymbol{c}_9$, whereas IMFs $\boldsymbol{c}_7$ and $\boldsymbol{c}_8$ capture an intermediate range of embedded frequencies. Fig. 10 (left panel) shows the distance values $D^*_{i,i+1,j}$ computed for $i = 1{:}8$, whereas Fig. 11 (right panel) shows the values of $SSW_f(K)$ and $K * SSB_f(K)$ as a function of the number of extracted CMFs, K. The CMF selection methodology yields $K = 3$ CMFs, which are shown in Fig. 11.

INSERT FIG 10 ABOUT HERE

INSERT FIG 11 ABOUT HERE

The automatically selected CMF decomposition consists of $\boldsymbol{c}_{S_1} = \sum_{i=1}^{3} \boldsymbol{c}_i$, $\boldsymbol{c}_{S_2} = \sum_{i=4}^{8} \boldsymbol{c}_i$ and $\boldsymbol{c}_{S_3} = \boldsymbol{c}_9$. In this case, the separation between the noise term and the transient mode is better than the one achieved in Example A, because the three IMFs that capture most of the signal noise are summed up in the first CMF, $\boldsymbol{c}_{S_1}$, whereas in Example A some mixing of noise and transient terms occurred. The proper isolation of the noise term from other modes depends on many factors, including the properties of the noise term itself and the number of spurious modes generated by the sifting algorithm in the high frequency range.

The clear separation between the transient patterns and the baseline into a minimal number of modes is expected to enhance the multi-scale decomposition and its interpretability with respect to the original IMF decomposition. It is worth to notice that the proposed approach is able to deal with a larger number of embedded modes too. As an example, Fig. 12 shows the final decomposition into four CMFs when a variant of the synthetic signal in Example A is generated by adding a Gaussian baseline with parameters $\mu = 0.5$ and $\sigma = 0.2$. In this case, the proposed approach is able to generate a further CMF that separates the Gaussian baseline from other modes. Nevertheless, the use of a clustering-like approach makes the methodology more effective when few modes are of interest, i.e., when basic EMD and competitor multi-scale decomposition techniques exhibit their major limitations.

INSERT FIG 12 ABOUT HERE

4.2. Competitor methods

The literature lacks automated approaches for the decomposition of a multi-scale signal into a minimal number of CMFs via EMD. Nevertheless, it is possible to consider alternative ways to perform the two major steps of the proposed methodology, i.e., (i) the IMFs combination into a reduced number of CMFs and (ii) the selection of the best value of the parameter K. With regard to the selection of which IMFs should be grouped together, different synthetic indexes proposed in the literature may be used in place of the density function. The following indexes, which associate one single value to each IMF, are considered in this study: (1) energy, $E_{i,j}$, (2) correlation with the original signal, $C_{i,j}$, (3) range, $R_{i,j}$, (4) kurtosis, $K_{i,j}$ and (5) skewness, $S_{i,j}$. They are defined as follows:

$$
\begin{aligned}
E_{i,j} &= \sum_{u=1}^{p} c_{i,u,j}^2\,, \qquad i = 1,\dots,n_j; j = 1,2,\dots \\
C_{i,j} &= \frac{cov(\boldsymbol{c}_{i,j}, \boldsymbol{Y}_j)}{\sigma_{\boldsymbol{c}_{i,j}}\sigma_{\boldsymbol{Y}_j}}, \qquad i = 1,\dots,n_j; j = 1,2,\dots \\
R_{i,j} &= \max(\boldsymbol{c}_{i,j}) - \min(\boldsymbol{c}_{i,j})\,, \qquad i = 1,\dots,n_j; j = 1,2,\dots \\
K_{i,j} &= \frac{\frac{1}{p}\sum_{u=1}^{p}(c_{i,u,j} - \bar{c}_{i,j})^4}{\left[\frac{1}{p}\sum_{u=1}^{p}(c_{i,u,j} - \bar{c}_{i,j})^2\right]^2}, \qquad i = 1,\dots,n_j; j = 1,2,\dots \\
S_{i,j} &= \frac{\frac{1}{p}\sum_{u=1}^{p}(c_{i,u,j} - \bar{c}_{i,j})^3}{\left[\frac{1}{p}\sum_{u=1}^{p}(c_{i,u,j} - \bar{c}_{i,j})^2\right]^{3/2}}, \qquad i = 1,\dots,n_j; j = 1,2,\dots
\end{aligned}
\tag{15}
$$

where $cov(\boldsymbol{c}_{i,j}, \boldsymbol{Y}_j)$ is the covariance between the i[th] IMF and the original signal, $\sigma_{\boldsymbol{c}_{i,j}}$ is the sample standard deviation of the i[th] IMF, and $\sigma_{\boldsymbol{Y}_j}$ is the sample standard deviation of the original signal, $\boldsymbol{Y}_j$. Two index-based solutions were investigated, i.e., one based on using one single index at a time (hereafter denoted by "univariate" approach), and one based on merging the information of every index into a 5-variate vector $\boldsymbol{V}_{i,j} = [E_{i,j}, C_{i,j}, R_{i,j}, K_{i,j}, S_{i,j}]^T$ (hereafter denoted by "multivariate approach"). Being $I_{i,j}$ any of synthetic indexes mentioned above, the distance $D^*_{i,i+1,j}$ in Eq. 7 can be computed as follows (univariate approach):

$$D^*_{i,i+1,j} = \frac{\left\| I_{i,j} - I_{i+1,j} \right\|}{w_{i,j}} \tag{16}$$

In the multivariate case, the distance $D^*_{i,i+1,j}$ can be alternatively computed as follows:

$$D^*_{i,i+1,j} = \frac{\left\| \boldsymbol{V}_{i,j} - \boldsymbol{V}_{i+1,j} \right\|}{w_{i,j}} \tag{17}$$

With regard to the selection of the optimal number K of CMFs, different criteria might be considered, inherited from cluster validity applications [30]. As discussed in Appendix B, some criteria are not suitable for the present application, because the minimum number of modes they can find is $K = 3$. Two possible competitors consist of the so-called *Ball & Hall* criterion and the *Zhao et al*. criterion, described in Appendix B. A third competitor may be considered too, which consists of an unpenalized version of the *Ball & Hall* criterion: the corresponding index is $SSW_f(K)$, instead of $SSW_f(K)/K$, and the optimum K corresponds to the elbow point in the $SSW_f(K)$ curve.

A comparison study was carried out to study the advantages of the proposed approach with respect to alternative methods. Each competitor approach was tested on $N = 1000$ random realizations of the signal profiles $\boldsymbol{Y}_j$, $j = 1,..,N$, in Example A and Example B. The frequency of selecting a given number K of CMFs is a performance index that shows to what extent each decomposition method yields the same number of CMFs for patterns that repeat over time (under the same process conditions). Because of this, it is used as a "stability index" of the algorithm: the more stable is the result in terms of the number K, the more suitable is the algorithm for in-process monitoring applications, when repeating patterns must be automatically decomposed and analysed. Table 1 compares the results based on different ways to compute the distance $D^*_{i,i+1,j}$ and to group the IMFs into the CMFs. In this case, the criterion in Eq. 8 is used to find the optimal number K. Table 2 compares the results based on different criteria for the selection of the number K of CMFs, by using the distance $D^*_{i,i+1,j}$ defined in Eq. 7. In both the examples, the proposed method outperforms competing approaches in terms of stability in selecting the appropriate number K, with the only exception of the kurtosis-based and the multivariate methods. However, these two last methods yield a worse decomposition of the signal, as shown by Fig. 13 (Example A). As a matter of fact both these approaches fail in appropriately distinguishing the transient and the stationary modes. As a final remark, index-based methods exploit only a portion of the information content of the IMFs to determine their dissimilarity, and hence the choice of the most suitable index is strongly problem-dependent. The use of multivariate statistics that merge multiple indexes do not guarantee adequate performances too, and final performances still depend

on the choice of the indexes. The proposed method, which is based on a density-based distance, takes advantage of all the available information content and avoids the troublesome and arbitrary index selection problem.

INSERT TABLE 1 ABOUT HERE

INSERT TABLE 2 ABOUT HERE

INSERT FIG 13 ABOUT HERE

5 A real case study in waterjet cutting

An example of a manufacturing process characterized by multi-scale signals is the waterjet cutting process [37], where multi-scale pressure fluctuations are caused by the natural working regime of the ultra high pressure (UHP) pump. [18] showed that the EMD technique can be used to enhance the characterization of the water pressure signal, which represents a suitable source of information for machine health monitoring purposes. The interested reader is referred to [38] for a review of machine health assessment methods.

INSERT FIG 14 ABOUT HERE

Fig. 14 (left panel) shows the UHP pump of the waterjet plant considered in this study, which is a positive displacement pump comprised of three parallel single-acting piston/cylinder groups. The installation of the pressure transducer downstream of the three discharge check valves is shown in Fig. 14 (left panel) and schematically depicted in Fig. 14 (right panel).

The pressure signal was acquired at $2\ kHz$ and segmented into repeating "*profiles*", such that each pressure profile refers to a complete pumping cycle that includes the three consecutive acting strokes (for additional details about the plant and the process, the interested reader may refer to [18, 39]). Signal data were collected during repeated cuts on an aluminium laminate by using a $45\ kW$ UHP with a nominal working point characterized by a water pressure of $350\ MPa$, a water flow rate of $5\ l/min$ and a $0.25\ mm$ orifice. Fig. 15 (top panel) shows the pattern of the dynamic pressure signal (i.e., the pressure fluctuation about the static level) for one pumping cycle. The

pressure signal exhibits a superimposition of a noise component, some high-frequency transients that are localized in time (Fig. 15, bottom panels), and a low frequency oscillation.

INSERT FIG 15 ABOUT HERE

The low frequency pressure ripples are due to the plunger kinematics, whereas the high frequency transients correspond to transitions between the end of a plunger active stroke and the beginning of the next one [39]. When a piston reaches the top dead centre, a valve commutation occurs, with a consequent oil flow-rate reduction in the next cylinder, resulting in a dynamic pressure discontinuity. Some minor transients correspond to the time instants when each plunger completes its suction step: they are caused by a flow rate modification when the piston reaches the bottom dead centre.

As demonstrated in [18], different faults of the most critical components (e.g., components of the UHP pump or the cutting head) may differently affect the scale-dependent features of the signal. Because of this, a proper decomposition of those features is expected to provide multiple benefits: (i) a better characterization of the natural pattern variability, (ii) enhanced signal monitoring performances provided by a faster detection of small shifts affecting a single scale or a sub-set of scales, and (iii) an enhanced diagnostic capability, thanks to a better isolation of fault effects.
The waterjet cutting process is used as a real case study to discuss the performances of the proposed method in a real industrial application. In this study, $N = 100$ profiles acquired under stable process conditions and healthy machine state are considered.
An example of EMD decomposition and corresponding density function for one pressure profile is shown in Fig. 16.

INSERT FIG 16 ABOUT HERE

Fig. 17 (left panel) shows the distance values $D^*_{i,i+1,j}$ computed for $i = 1{:}8$, whereas Fig. 17 (right panel) shows the values of $SSW_f(K)$ and $K * SSB_f(K)$ as a function of the number of extracted CMFs, K. The $K = 3$ CMFs extracted from the signal are shown in Fig. 18.

INSERT FIG 17 ABOUT HERE

The automatically selected CMF decomposition consists of $\boldsymbol{c}_{S_1} = \boldsymbol{c}_1$, $\boldsymbol{c}_{S_2} = \sum_{i=2}^{6} \boldsymbol{c}_i$ and $\boldsymbol{c}_{S_3} = \sum_{i=7}^{9} \boldsymbol{c}_i$. The first CMF, $\boldsymbol{c}_{S_1}$, captures most of the signal noise and the third CMF, $\boldsymbol{c}_{S_3}$, captures the low frequency ripples, whereas the middle CMF, $\boldsymbol{c}_{S_2}$, sums up all the IMFs that describe the transients.

INSERT FIG 18 ABOUT HERE

Fig. 18 shows that the final CMF decomposition is more synthetic than the original IMF decomposition, and it yields a clear separation of transient modes from low frequency ripples. The noise term, whose amplitude is modulated by the kinematics of the plungers, is well separated from the other two scales too.

INSERT TABLE 3 ABOUT HERE

Since the literature lacks methods for the automatic extraction of a minimal number of embedded components, a fair comparison with other time-frequency analysis techniques is not applicable. Nevertheless, it is possible to evaluate what kind of decomposition may be achieved by using the wavelet transformation as an alternative to the EMD. As an example, the wavelet decomposition of one pressure profile is shown in Fig. 19, based on a mother wavelet that is widely used in practice, i.e., a fourth order Daubechies function. In this case, nine decomposition levels were required to separate the baseline from transient and noise modes into distinct scales. In this case, the number of decomposition levels is a controllable factor, but at least nine levels are required to achieve a suitable isolation of embedded modes, and analogous results were achieved by using different wavelet bases. The approximation at level 9, $\boldsymbol{a}_9$, captures the equivalent of the residual term in the EMD, whereas the detail components, $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_9$, captures high frequency and low frequency oscillations. Both the basic EMD and the wavelet transform yield a multi-scale description of the signal, but further processing steps are required to isolate and extract only the relevant oscillations modes.

INSERT FIGURE 19 ABOUT HERE

As already stated in Section 3, from a process monitoring viewpoint, both the correctness of the intrinsic mode separation and the stability of the automated CMF decomposition are of great importance. Table 3 and Table 4 show the stability over time of the number K of extracted CMFs

when the proposed approach is applied to repeating pressure profiles, $\boldsymbol{Y}_j$, for $j = 1, \ldots, N = 100$. The two tables compare the performances of the proposed approach with the performances achieved by using alternative ways either to group the IMFs into a reduced number of CMFs (Table 3) or to determine the best number K of CMF (Table 4).

INSERT TABLE 4 ABOUT HERE

Table 3 shows that different choices of synthetic dissimilarity indexes yield different numbers of CMFs. The energy index, $E_{i,j}$, the correlation index, $C_{i,j}$, and the range index, $R_{i,j}$, yield $K = 2$ for most pressure profiles, because they are not able to separate the noise term from the transient modes. The kurtosis index, $K_{i,j}$, the skewness index, $S_{i,j}$, and the multivariate index, $\boldsymbol{V}_{i,j}$, instead, yield $K = 3$ CMFs, being the kurtosis index the one that leads to the most stable results. Nevertheless, the kurtosis-based decomposition is worse than the one provided by our proposed approach (Fig. 20, left panel), because some high-frequency oscillation is added to the low frequency ripples, creating a mixing of the two modes. In this case, the multivariate index-based method (Fig. 20, right panel) provides the same decomposition shown in Fig. 18, but the simulation study showed that this approach is less effective than the one based on density functions.

Table 4 shows that the *Ball & Hall* criterion and the *Zhao et al*. criterion for the optimal choice of K lead to a decomposition into $K = 2$ CMFs (i.e., no separation between low frequency ripples and the high-frequency transients is achieved). Only the unpenalized *Ball & Hall* criterion yields $K = 3$, but only with a lower frequency than the one reached by the density-based approach.

INSERT FIG 20 ABOUT HERE

6 Conclusions

Sensor signals acquired in manufacturing processes are often characterized by cyclically repeating patterns that involve transients, non-stationary fluctuations and phenomena having different dynamics. It is known that a proper multi-scale characterization of the natural signal pattern could enhance process monitoring performances, because different faults may have different effects on distinct scales. Nevertheless, the achievement of a synthetic decomposition into a minimal number of physically relevant and interpretable oscillation modes is still an open issue. The lack of automated and data-driven procedures to generate a synthetic signal decomposition into few

relevant modes motivated this study. The proposed method is based on the EMD technique and it is aimed at decomposing the signal into a minimal number of CMFs. The algorithm works by solving two major problems: (i) the determination of which IMFs should be summed up together into a single CMF (i.e., those IMFs that represent split components of a single intrinsic mode), and (ii) the determination of the optimal number K of CMFs that is required to characterize the signal pattern. A density function-based criterion was proposed to solve the former problem, and a sum-of-squares based index was proposed to solve the latter one. The proposed method was tested on both synthetic and real data, showing that it is able to separate transient modes that are localized in time from noise terms and low frequency baselines. The method outperformed the competitor criteria both in the simulation study and in a real case study that involved water pressure signals in a waterjet cutting application. The fully data-driven nature of the method makes it suitable for in-process monitoring applications, where one is interested in automatically decomposing profile patterns that repeat over time, in order to detect fault conditions affecting one (or more) features belonging to different scales.

There are different known issues that might be investigated in the future to further enhance the proposed methodology. They include the following: (i) density function-based dissimilarity criteria do not allow separating non-transient waveforms having the same amplitude but different frequency: the proposed method might be coupled with other criteria to extended the application field; (ii) the separation between the noise term and other modes might be enhanced and investigated in more detail; (iii) in the presence of many (i.e., more than four) distinct scales, the criterion for the choice of K might be too conservative, yielding less CMFs than those actually required: further criteria might be considered in the future.

Eventually, it is worth to notice that the CMF methodology may deal with one of most critical issues in EMD, i.e., the mode mixing effect, at least when single oscillation modes are split into multiple IMFs. Nevertheless, the development of computationally efficient EMD-based algorithms that allow minimizing the occurrence of such an effect is expected to improve the performances of CMF-based methods like the one proposed in this study.

Nomenclature

$\boldsymbol{a}_k$	k^{th} approximation level in wavelet analysis
$BH(K)$	*Ball & Hall* index for K CMFs
$\boldsymbol{c}_{i,j}$	i^{th} IMF extracted from the j^{th} signal profile $\boldsymbol{Y}_j$, $i = 1, \dots, n_j,\ j = 1,2, \dots$
$\boldsymbol{c}_{s_k,j}$	k^{th} CMF extracted from the j^{th} signal profile $\boldsymbol{Y}_j$, $k = 1, \dots, K, j = 1,2, \dots$

$C_{i,j}$	cross-correlation between i[th] IMF and $\boldsymbol{Y}_j$, $i = 1, \dots, n_j$, $j = 1,2, \dots$
$CH(K)$	*Calinski & Harabasz* index for K CMFs
CMF	Combined Mode Function
$\boldsymbol{d}_k$	k[th] detail level in wavelet analysis
$D_{i,i+1,j}$	distance (dissimilarity) between the i[th] and the (i+1)[th] IMFs from $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$D^*_{i,i+1,j}$	penalized distance between the i[th] and the (i+1)[th] IMFs from $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$E_{i,j}$	energy of i[th] IMF from $\boldsymbol{Y}_j$, $i = 1, \dots, n_j$, $j = 1,2, \dots$
EMD	Empirical Mode Decomposition
n_j	Number of IMFs extracted from the j[th] signal profile $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$f(x)$	probability density function of random process $x(t)$
$\hat{f}(x)$	kernel estimator of the probability density function of random process $x(t)$
f	frequency of a generic sine wave $x(t)$
f_s	sampling frequency
h	bandwidth of the kernel function
$\hat{h}$	optimal bandwidth of the kernel function
$\boldsymbol{h}_{u,j}$	difference between the signal $\boldsymbol{Y}_j$ and $\boldsymbol{m}_{u,j}$, at u[th] step of the sifting algorithm
$H(K)$	*Hartigan* index for K CMFs
K	final number of CMFs extracted from the j[th] signal profile $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$K_{i,j}$	kurtosis of i[th] IMF from $\boldsymbol{Y}_j$, $i = 1, \dots, n_j$, $j = 1,2, \dots$
$Ker(x)$	kernel function
IMF	Intrinsic Mode Function
$\boldsymbol{m}_{u,j}$	mean of envelopes at the u[th] step of the sifting algorith
p	Number of datapoints in each signal profile
q_k	number of IMFs included into the k[th] CMF, $k = 1, \dots, K$, $j = 1,2, \dots$
$\boldsymbol{r}_{n_j,j}$	residue of the EMD for the j[th] signal profile $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$R_{i,j}$	range of i[th] IMF from $\boldsymbol{Y}_j$, $i = 1, \dots, n_j$, $j = 1,2, \dots$
$S_{i,j}$	skewness of i[th] IMF from $\boldsymbol{Y}_j$, $i = 1, \dots, n_j$, $j = 1,2, \dots$
$SSW_{f,j}(K)$	sum-of-squares within the K CMFs from $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$SSB_{f,j}(K)$	sum-of-squares between the K CMFs from $\boldsymbol{Y}_j$, $j = 1,2, \dots$
$supp\left(\hat{f}(x)\right)$	support of estimated density function $\hat{f}(x)$
T	period of a generic sine wave $x(t)$
UCV	Unbiased Cross Validation

$UCV_{i,j}(\hat{h})$ unbiased cross-validation statistic for i[th] IMF from $\boldsymbol{Y}_j$, with bandwidth $\hat{h}$, $j = 1,2, \ldots$

$\boldsymbol{V}_{i,j}$ 5-variate vector of synthetic indexes for the i[th] IMF from $\boldsymbol{Y}_j$, $i = 1, \ldots, n_j, j = 1,2, \ldots$

$w_{i,j}$ weight used in the expression of $D^*_{i,i+1,j}$, $i = 1, \ldots, n_j$, $j = 1,2, \ldots$

X amplitude of a generic sine wave $x(t)$

$\boldsymbol{Y}_j$ j[th] realization of a random signal profile, $j = 1,2, \ldots$

$Y_j(t)$ j[th] signal profile in the simulation study

$Y_{n_j}(t)$ noise term of synthetic signals (parameter: σ_n^2)

$Y_{T_j}(t)$ transient term of synthetic signals (parameters: A_T, μ_T, σ_T^2, F_T)

$Y_{S_j}(t)$ baseline term of synthetic signals (parameters: A_S, F_S)

$Z(K)$ *Zhao et al*. index for K CMFs

Acknowledgments

The authors want to thank Mr. Andrea Carroccia and Mr. Matteo Maggioni for their valuable contribution in the implementation of relevant parts of the algorithm presented in this study.

References

[1] Jin J., Shi, J., 1999, "Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets", Technometrics, 41(4), pp. 327 – 339

[2] Jin, J., Shi, J., 2001, "Automatic Feature Extraction of Waveform Signals for In-process Diagnostic Performance Improvement", Journal of Intelligent Manufacturing, 12, pp. 257 – 268

[3] Feng, Z., Liang, M., Chu, F., 2013, "Recent Advances in Time-Frequency Analysis Methods for Machinery Fault Diagnosis: A Review with Application Examples", Mechanical Systems and Signal Processing, 38, pp. 165-205

[4] Ayenu-Prah, A., Attoh-Okine, N., 2010, "A Criterion for Selecting Relevant Intrinsic Mode Functions in Empirical Mode Decomposition", Advances in Adaptive Data Analysis, 2(1), pp. 1-24

[5] Flandrin, P., Goncalves, P., Rilling, G., 2004, "Detrending and Denoising with Empirical Mode Decomposition", XII European Signal Processing Conference (EUSIPCO), September 6-10, 2004, Vienna, Austria

[6] Gao, Q., Duan, C., Fan, H., Meng, Q., 2008, "Rotating Machine Fault Diagnosis using Empirical Mode Decomposition", Mechanical Systems and Signal Processing, 22(5), pp. 1072-1081.

[7] Huang, N.E., Shen, Z., Long, S.R., Wu, M.L., Shih, H.H., et al., 1998, "The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear And Nonstationary Time Series Analysis", Proc. R. Soc. London Ser. A, 454, pp. 903–95

[8] Jia, Z., Zhang, L., Wang, F., Liu, W., 2010, "A new method for discharge state prediction of micro-EDM using empirical mode decomposition", Journal of manufacturing science and engineering, 132(1), pp. 014501-014501-6

[9] Peng, Y., 2006, "Empirical model decomposition based time-frequency analysis for the effective detection of tool breakage", Journal of manufacturing science and engineering, 128(1), pp. 154-166.

[10] Rai, V. K., Mohanty, A. R., 2007, "Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform", Mechanical Systems and Signal Processing, 21(6), pp. 2607-2615.

[11] Guo, W., Tse, P.W., 2013, "A novel signal compression method based on optimal ensemble empirical mode decomposition for bearing vibration signals", Journal of sound and vibration, 332(2), pp. 423-441.

[12] Yan, R., Gao, R. X., 2008, "Rotary machine health diagnosis based on empirical mode decomposition", Journal of Vibration and Acoustics, 130(2), pp. 021007-021007-12.

[13] Roth, J. T., Djurdjanovic, D., Yang, X., Mears, L., Kurfess, T., 2010, "Quality and inspection of machining operations: tool condition monitoring", Journal of Manufacturing Science and Engineering, 132(4), pp. 041015-041015-16.

[14] Wu, Z., Huang, N.E., 2009, "Ensemble Empirical Mode Decomposition: a Noise Assisted Data Analysis Method", Advances in Adaptive Data Analysis, 1(1), pp. 1-41

[15] Zvokelj, M., Zupan, S., Prebil, I., 2010, "Multivariate and Multi-scale Monitoring of Large-size Low Speed Bearings using Ensemble Empirical Mode Decomposition Method Combined with Principal Component Analysis", Mechanical Systems and Signal Processing, 24, pp. 1049-1067

[16] Zvokelj, M., Zupan, S., Prebil, I., 2011, "Non-Linear Multivariate and Multi-scale Monitoring and Signal Denoising Strategy using Kernel Principal Component Analysis Combined with Ensemble Empirical Mode Decomposition Method", Mechanical Systems and Signal Processing, 25(7), pp. 2631-2653

[17] Zhang, J., Yan, R., Gao, R.X., Feng, Z., 2010, "Performance Enhancement of Ensemble Empirical Mode Decomposition", Mechanical Systems and Signal Processing, 24, pp. 2104-2123

[18] Grasso, M., Pennacchi, P., Colosimo, B. M., 2014, "Empirical mode decomposition of pressure signal for health condition monitoring in waterjet cutting", International Journal of Advanced Manufacturing Technology, 72(1-4), pp. 347-364

[19] Liang, H., Lin, Q-H., Chen, J.D.Z., 2005, "Application of the Empirical Mode Decomposition to the Analysis of the Esophageal Manometric Data in Gastroesophageal Reflux Disease", IEEE Transactions on Biomedical Engineering, 52(10), pp. 1692-1701

[20] Bu, N., Ueno, N., Fukuda, O., 2007, "Monitoring of Respiration and Heartbeat during Sleep using a Flexible Piezoelectric Film Sensor and Empirical Mode Decomposition", Proceedings of the 29th Annual International Conference of the IEEE EMBS, 2007, Lyon (France)

[21] Ricci, R., Pennacchi, P., 2011, "Diagnostics of Gear Faults Based on EMD and Automatic Selection of Intrinsic Mode Functions", Mechanical Systems and Signal Processing, 25(3), pp. 821-838

[22] Komaty, A., Boudraa, A. O., Augier, B., Daré-Emzivat, D., 2014, "EMD-based filtering using similarity measure between probability density functions of IMFs", Instrumentation and Measurement, IEEE Transactions on, 63(1), pp. 27-34.

[23] Cheremisinoff, N. P., Ferrante, L., 1987, "Practical statistics for engineers and scientists", CRC Press, UK

[24] Izenman, A. J., 1991, "Review papers: recent developments in nonparametric density estimation", Journal of the American Statistical Association, 86(413), pp. 205-224.

[25] Zambon A.Z., Dias, R., 2012, "A review of Kernel Density Estimation with Applications to Econometrics", arXiv:1212.2812v1 [stat.ME], pp. 1-35

[26] Venables, W. N., Ripley, B. D., 2002, "Modern Applied Statistics with S", Springer, New York.

[27] Bowman, A. W., Azzalini, A., 1997, "Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations", Oxford University Press, Oxford.

[28] Bowman, A. W., 1984, "An alternative method of cross-validation for the smoothing of kernel density estimates", Biometrika, 71, pp. 353–360.

[29] Hardle, W., Marron, J. S., Wand, M. P., 1990, "Bandwidth choice for density derivatives", Journal of the Royal Statistical Society. Series B (Methodological), pp. 223-232.

[30] Vendramin, L., Campello, R. J., Hruschka, E. R., 2010, "Relative clustering validity criteria: A comparative overview", Statistical Analysis and Data Mining: The ASA Data Science Journal, 3(4), pp. 209-235

[31] Ball, G.H., Hall, D.J., 1965, "ISODATA, A novel method of data analysis and pattern classification", Tech. Rep. NTIS No. AD 699616. Standford Research Institute, Menlo Park

[32] Calinski, T., Harabasz, J., 1974, "A dendrite method for cluster analysis", Communication in statistics, 3, pp. 1–27

[33] Hartigan, J.A., 1975, "Clustering algorithms", Wiley Ed., New York

[34] Zhao, Q., Xu, M., Fränti, P., 2009, "Sum-of-squares based cluster validity index and significance analysis. Adaptive and Natural Computing Algorithms", Lecture Notes in Computer Science, 5495, pp. 313-322

[35] Xu, L., 1997, "Bayesian Ying-Yang machine, clustering and number of clusters", Pattern Recognition Letters, 18, pp. 1167–1178

[36] Wu, Z., Huang, N. E., 2004, "A study of the characteristics of white noise using the empirical mode decomposition method", Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 460(2046), pp. 1597-1611.

[37] Kovacevic R., Hashish M., R. Mohan, Ramulu M., Kim T.J., Geskin E.S., 1997, "State Of The Art Of Research And Development In Abrasive Waterjet Machining", Journal of Manufacturing Science and Engineering, 119(4), pp. 776 – 785

[38] Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., Siegel, D., 2014, "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications", Mechanical Systems and Signal Processing, 42(1), pp. 314-334.

[39] Grasso M., Goletti M., Annoni M., Colosimo B.M., 2013, "A New Approach for Online Health Assessment of Abrasive Waterjet Cutting Systems", International Journal of Abrasive Technology, 6(2), pp. 158-181

[40] Rilling, G., Flandrin, P., Goncalves, P., 2003, "On empirical mode decomposition and its algorithms", IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, 2003, Grado (Italy)

[41] Carroccia, A., Grasso, M., Maggioni, M., Colosimo, B.M., 2015, "Improved Signal Characterization via Empirical Mode Decomposition to Enhance In-Line Quality Monitoring", 48th CIRP Conference on Manufacturing Systems, CIRP CMS 2015, 24-26 June 2015, Ischia, Italy

Appendix A: the sifting algorithm

Let $\boldsymbol{Y}_j = [Y_{j,1}, Y_{j,2}, \dots, Y_{j,p}]^T$ be a signal profile, where p is the number of datapoints. Then, the IMFs that capture intrinsic oscillation modes can be extracted by means of the "sifting" process, which consists of the following steps [7]:

1. All the local minima and maxima of the profile $\boldsymbol{Y}_j$, $j = 1,2, \dots$, are identified and they are interpolated respectively by an upper and a lower envelope expressed on a cubic spline basis;
2. the mean of the two envelopes is computed and designated as $\boldsymbol{m}_{1,j}$; then, the difference between the signal $\boldsymbol{Y}_j$ and $\boldsymbol{m}_{1,j}$ is computed and designated as $\boldsymbol{h}_{1,j}$:

$$\boldsymbol{h}_{1,j} = \boldsymbol{Y}_j - \boldsymbol{m}_{1,j}, \quad j = 1,2, \dots \tag{A1}$$

If $\boldsymbol{h}_{1,j}$ is an IMF, i.e., if $\boldsymbol{h}_{1,j}$ satisfies the following conditions:

a) in the entire dataset, the number of extremes and the number of zero crossings must be either equal or different at most by one;

b) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero;

then, $\boldsymbol{h}_{1,j}$ is taken as the first IMF of the signal and designated as $\boldsymbol{c}_{1,j}$. If $\boldsymbol{h}_{1,j}$ is not an IMF, $\boldsymbol{h}_{1,j}$ replaces the original signal and the above steps are repeated until an IMF is obtained.

3. The first IMF, $\boldsymbol{c}_{1,j}$, is separated from the signal $\boldsymbol{Y}_j$ by:

$$\boldsymbol{r}_{1,j} = \boldsymbol{Y}_j - \boldsymbol{c}_{1,j}, \quad j = 1,2,\ldots \tag{A2}$$

The residue $\boldsymbol{r}_{1,j}$ is treated as the original signal and the above steps are repeated, leading to the extraction of the following IMFs $\boldsymbol{c}_{2,j}, \ldots, \boldsymbol{c}_{n_j,j}$ such that:

$$\begin{aligned} \boldsymbol{r}_{1,j} - \boldsymbol{c}_{2,j} &= \boldsymbol{r}_{2,j} \\ \boldsymbol{r}_{n_j-1,j} - \boldsymbol{c}_{n_j,j} &= \boldsymbol{r}_{n_j,j}, \quad j = 1,2,\ldots \end{aligned} \tag{A3}$$

At the end of the process, the signal is decomposed into n_j IMFs and a residue $\boldsymbol{r}_{n_j,j}$:

$$\boldsymbol{Y}_j = \sum_{i=1}^{n_j} \boldsymbol{c}_{i,j} + \boldsymbol{r}_{n_j,j}, \quad j = 1,2,\ldots \tag{A4}$$

The residue is a signal such that no further decomposition is possible. In this study, the Amplitude Ratio criterion proposed by Rilling *et al*. [40] was applied.

Appendix B: cluster validity criteria

The procedure of evaluating and comparing the results of clustering algorithms is known as "*cluster validity*" [30]. The aim is not only to select the best algorithm, but also to determine the "correct" number of clusters in the dataset. Some of these validity criteria measure the similarity of data within the group and the dissimilarity of data between the groups, by means of sum of squared indexes. The same underlying concept might be exploited in the EMD framework, to determine the number K of CMFs corresponding to the best compromise between the similarity of IMFs in the same CMF and the dissimilarity among IMFs belonging to distinct CMFs.

Let $SSW_{f,j}(K) = \sum_{k=1}^{K} \sum_{i \in \boldsymbol{c}_{s_k,j}} \left\| \hat{f}\left(\boldsymbol{c}_{i,j}\right)_k - \overline{\hat{f}_{k,J}}(k) \right\|$ be the sum-of-squares within the K CMFs, and $SSB_{f,j}(K) = \sum_{k=1}^{K} n_{k,j}(k) \left\| \overline{\hat{f}_{k,J}}(k) - \overline{\hat{f}_J} \right\|$ be the sum-of-squares between the K CMFs, defined in Section 3.2. Table B.1 summarizes the sum-of-squares based criteria, named after the authors who proposed them, which were evaluated in this study.

INSERT TABLE B.1 ABOUT HERE

All those indexes, apart from the *Zhao et al*. index, require the estimation of a knee point, based on successive differences computation. The optimal number K corresponds to a knee of the target function, i.e., the maximum value of the absolute successive differences. However, this implies a constraint on the minimum number K that can be determined by the methods that involves the computation of $SSB_{f,j}(K)$. Since $SSB_{f,j}(1) = 0$, the minimum number K that can be determined by using successive differences method is $K = 3$. Because of this, only the *Ball & Hall* index and the *Zhao et al*. index are applicable. The *Ball & Hall* index requires successive differences computation, but it relies only on the $SSW_{f,j}(K)$ statistic, which can be computed for any $K < n_j$. The *Zhao et al*. index relies on both the $SSW_{f,j}(K)$ and $SSB_{f,j}(K)$ statistics but it requires no elbow detection: the optimal choice of K corresponds to the minimal value of the index $Z(K)$.

A variant of the *Ball & Hall* index is considered too, which avoids the penalization, i.e., the division by K in the formulation of $BH(K)$. Such a variant will be referred to as "*unpenalized Ball & Hall*" index. This variant was considered in a previous study [41] and it was empirically observed by the authors that it might yield better decomposition results than the penalized version. Nevertheless, the major limitation of the *Ball & Hall* approach in the EMD framework consists of ignoring the information captured by the $SSB_{f,j}(K)$ statistic.

Figure captions list

Fig. 1 Conceptual scheme of the proposed approach

Fig. 2 Sine wave of amplitude $X = 1$ (left panel) and corresponding probability density function (right panel)

Fig. 3 Transient sine wave of amplitude $X = 1$ (left panel) and corresponding probability density function (right panel)

Fig. 4 One realization of multi-scale random profiles in Example A

Fig. 5 Example of EMD of one profile in Example A (left panel) and corresponding probability density functions (right panel)

Fig. 6 Plot of the distance statistics $D^*_{i,i+1}$ (left panel) and plot of statistics $SSW_f(K)$ and $K * SSB_f(K)$ for different choices of K (right panel) – one profile realization in Example A

Fig. 7 Final CMF decomposition provided by the proposed approach for one profile realization in Example A

Fig. 8 One realization of multi-scale random profiles in Example B

Fig. 9 Example of EMD of one profile in Example B (left panel) and corresponding probability density functions (right panel)

Fig. 10 Plot of the distance statistics $D^*_{i,i+1}$ (left panel) and plot of statistics $SSW_f(K)$ and $K * SSB_f(K)$ for different choices of K (right panel) – one profile realization in Example B

Fig. 11 Final CMF decomposition provided by the proposed approach for one profile realization in Example B

Fig. 12 Final CMF decomposition provided by the proposed approach for a variant of Example A where a Gaussian baseline was added to the source signal

Fig. 13 Examples of CMF decompositions achieved by using the kurtosis-based method (left panel) and the multivariate index-based method (right panel) for IMF dissimilarity estimation - Example A

Fig. 14 Installation of the pressure transducer in the UHP pump (left panel) and scheme of the waterjet plant (right panel)

Fig. 15 Example of one pressure signal profile corresponding to a complete pumping cycle (top panel) and details of transients corresponding to the end of active strokes of plunger 1 and plunger 2 (bottom panels)

Fig. 16 EMD of one water pressure profile (left panel) and corresponding density functions (right panel)

Fig. 17 Plot of the distance statistics $D^*_{i,i+1}$ (left panel) and plot of statistics $SSW_f(K)$ and $K * SSB_f(K)$ for different choices of K (right panel) – one pressure profile in the waterjet case study

Fig. 18 CMF decomposition provided by the proposed approach for one pressure profile in the waterjet case study

Fig. 19 Example of pressure signal wavelet decomposition – Daubechies mother wavelet of fourth order, nine decomposition levels

Fig. 20 Examples of CMF decompositions achieved by using the kurtosis-based method (left panel) and multivariate index-based method (right panel) for one pressure profile in the waterjet case study

Table captions list

Table 1 Frequency of selection of K CMFs by using different IMF dissimilarity criteria (optimal number K is selected by using Eq. 8)

Table 2 Frequency of selection of K CMFs by using different criteria for the selection of the optimal number K (IMF dissimilarity estimated by using Eq. 7)

Table 3 Frequency of selection of K CMFs by using different IMF dissimilarity criteria (optimal number K is selected by using Eq. 8)

Table 4 Frequency of selection of K CMFs by using different criteria for the selection of the optimal number K (IMF dissimilarity estimated by using Eq. 7)

Table B.1 Cluster validity indexes

Fig 1

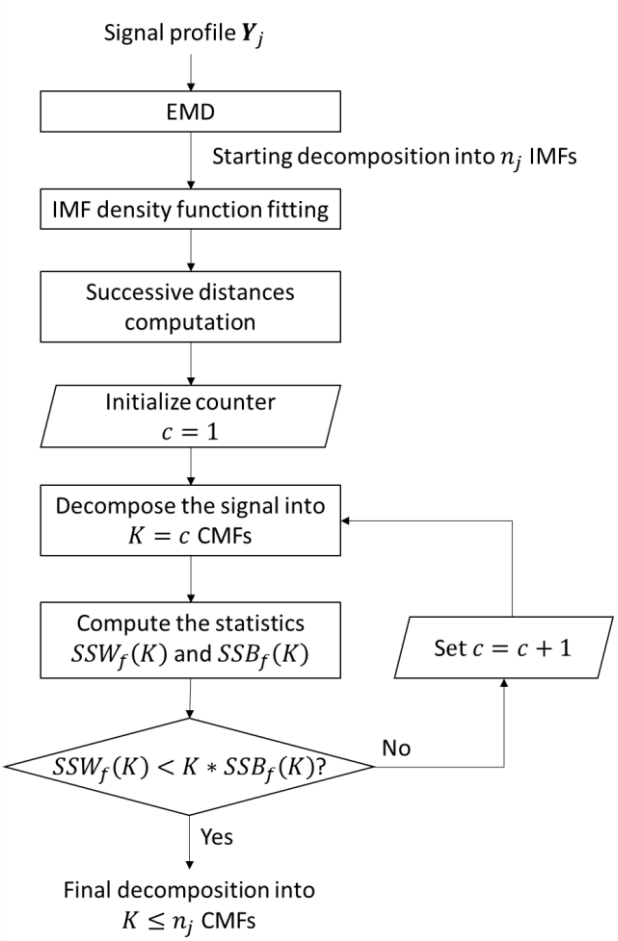

Signal profile $\boldsymbol{Y}_j$
EMD
Starting decomposition into n_j IMFs
IMF density function fitting
Successive distances computation
Initialize counter $c = 1$
Decompose the signal into $K = c$ CMFs
Compute the statistics $SSW_f(K)$ and $SSB_f(K)$
Set $c = c + 1$
$SSW_f(K) < K * SSB_f(K)$?
No
Yes
Final decomposition into $K \leq n_j$ CMFs

Fig 2

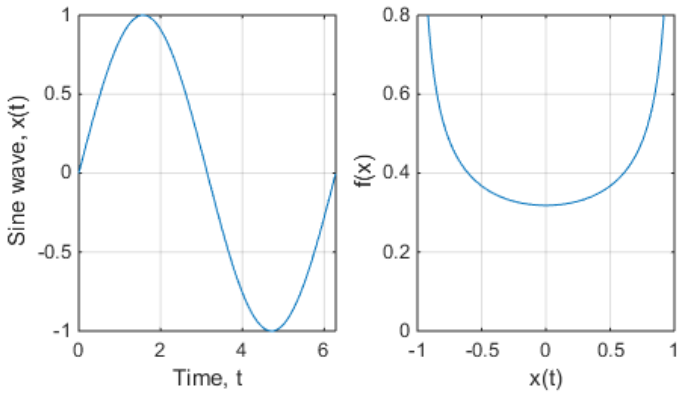


Fig 3

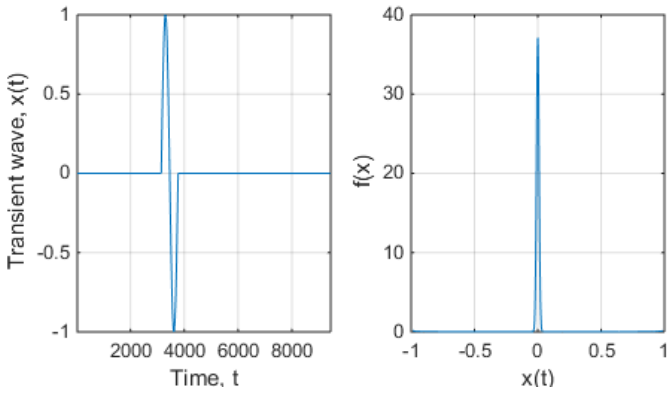

Transient wave, x(t)
1
0.5
0
-0.5
-1
2000 4000 6000 8000
Time, t
f(x)
40
30
20
10
0
-1 -0.5 0 0.5 1
x(t)

Fig 4

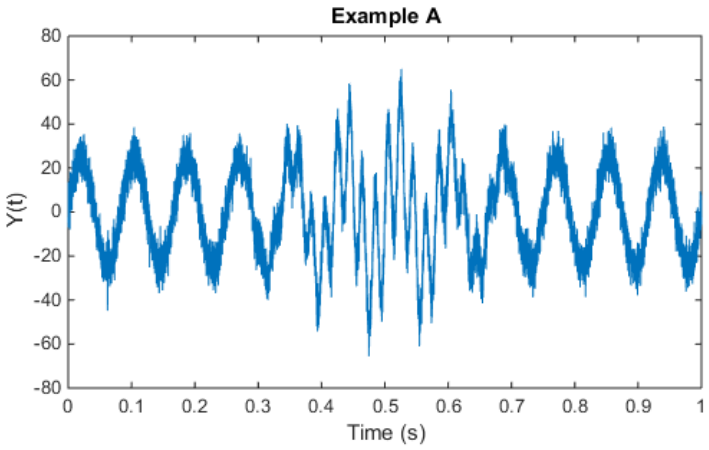

Example A
80
60
40
20
0
-20
-40
-60
-80
Y(t)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time (s)

Fig 5

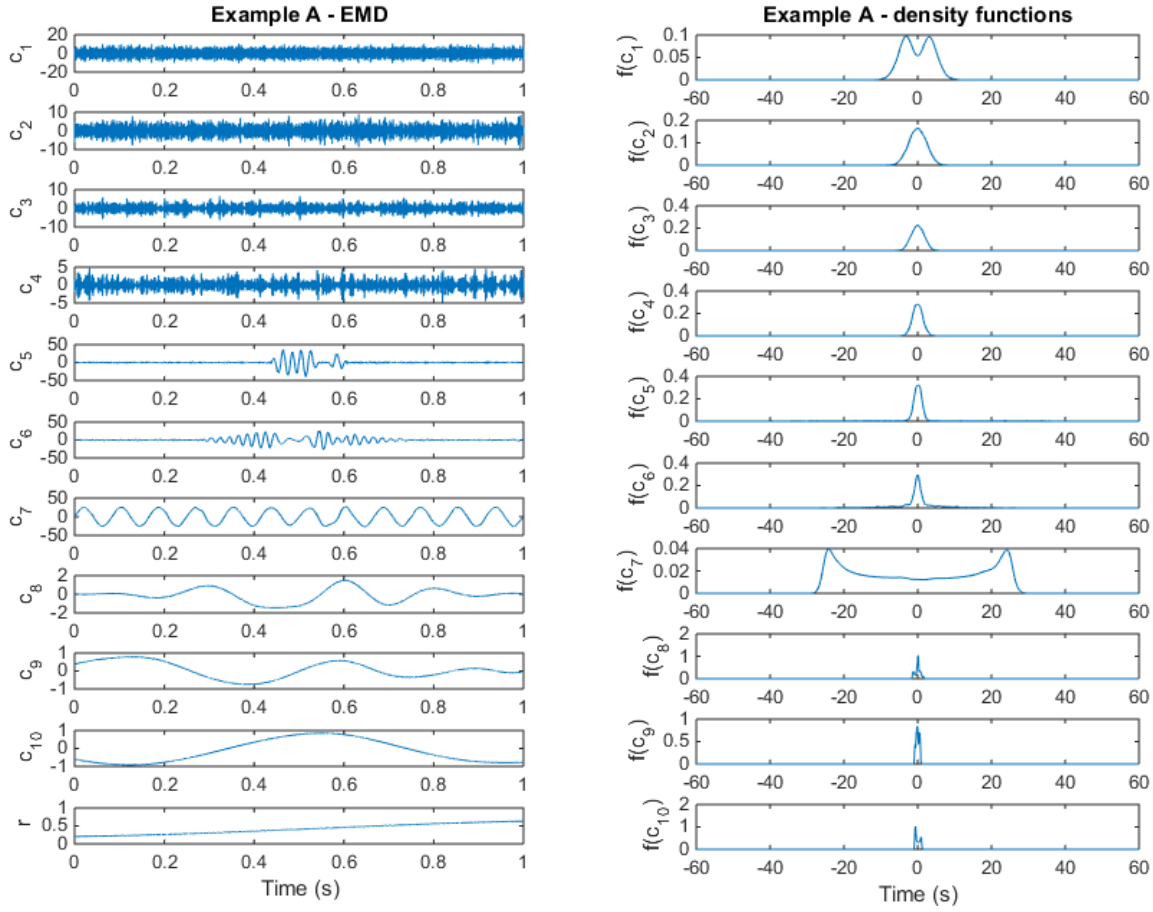


Fig 6

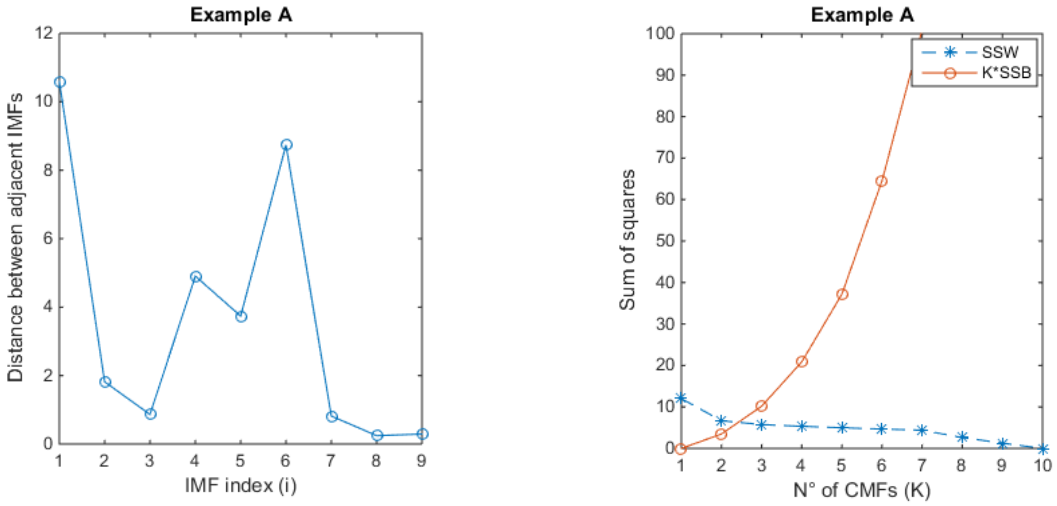

Example A
Distance between adjacent IMFs
IMF index (i)
Example A
Sum of squares
N° of CMFs (K)
SSW
K*SSB

Fig 7

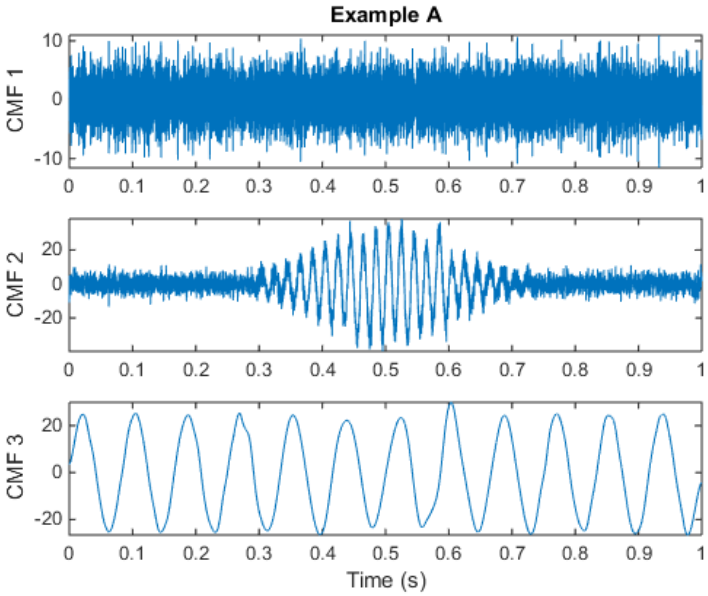


Fig 8

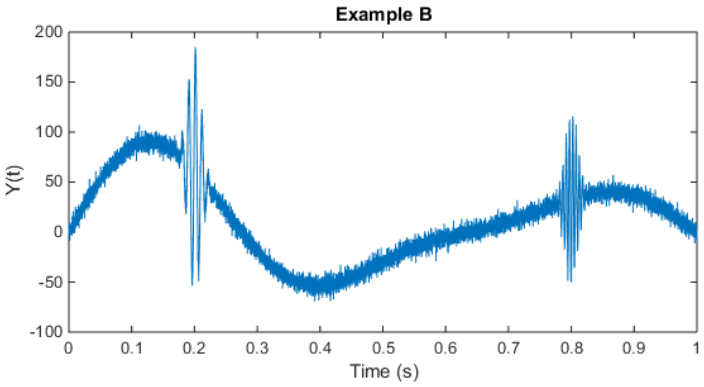

Example B
200
150
100
50
0
-50
-100
Y(t)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Time (s)

Fig 9

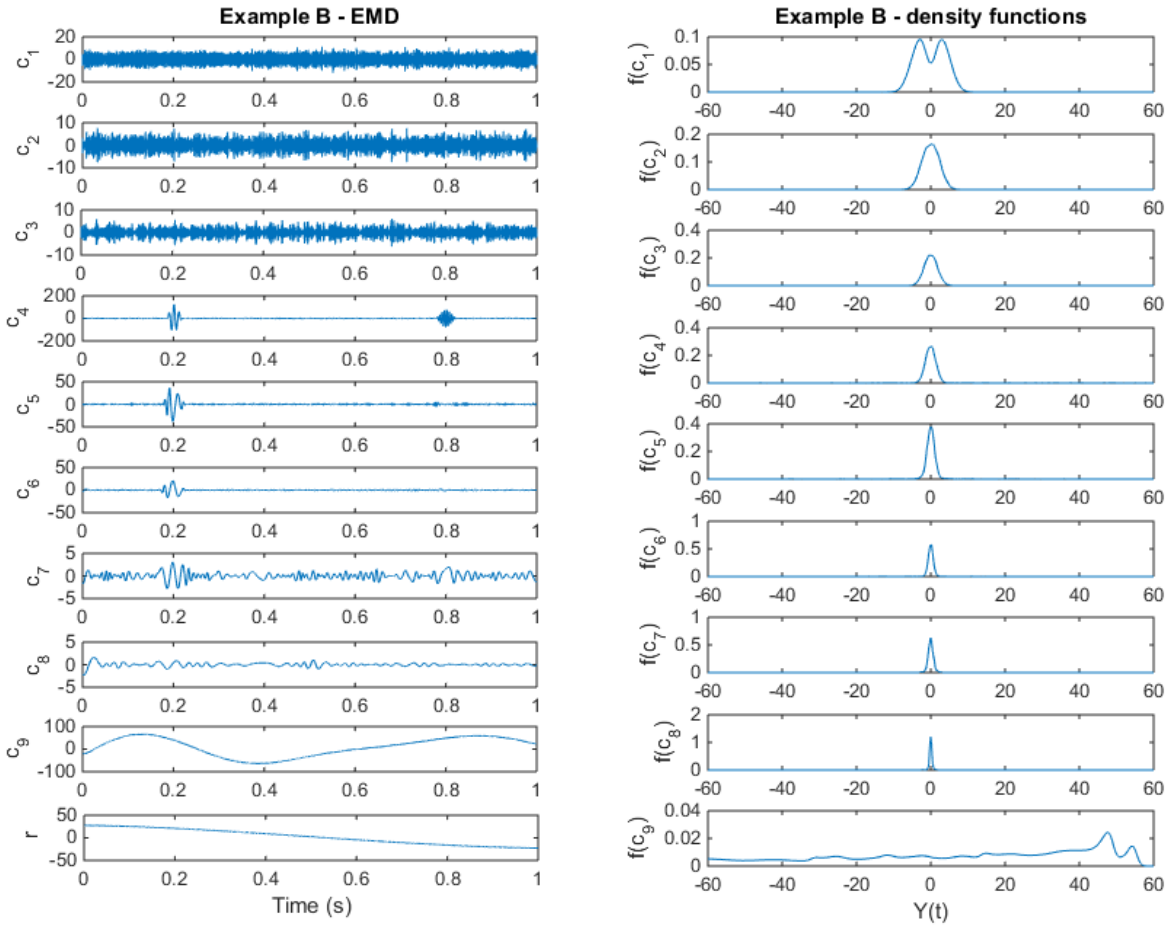


Fig 10

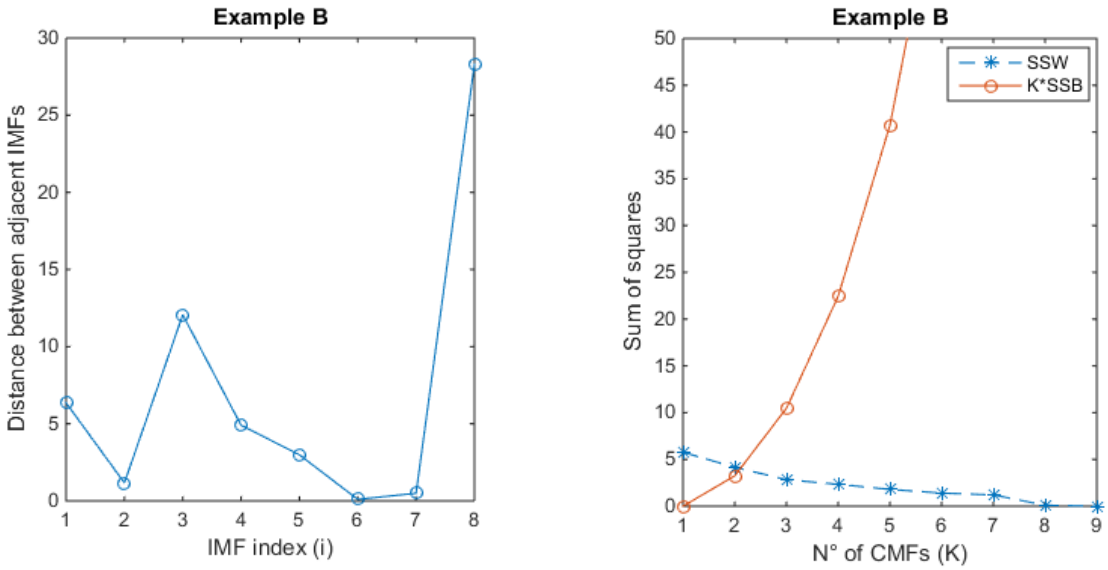

Example B
Distance between adjacent IMFs
IMF index (i)
Example B
Sum of squares
N° of CMFs (K)
SSW
K*SSB

Fig 11

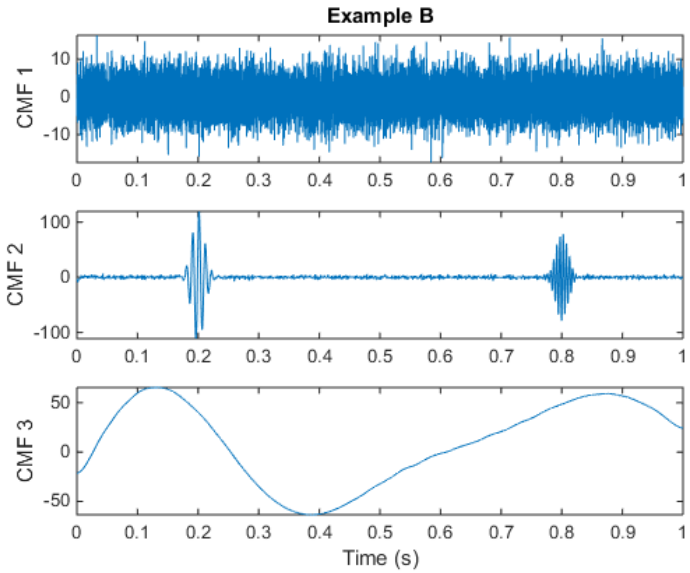

Example B
CMF 1
10
0
-10
CMF 2
100
0
-100
CMF 3
50
0
-50
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Time (s)

Fig 12

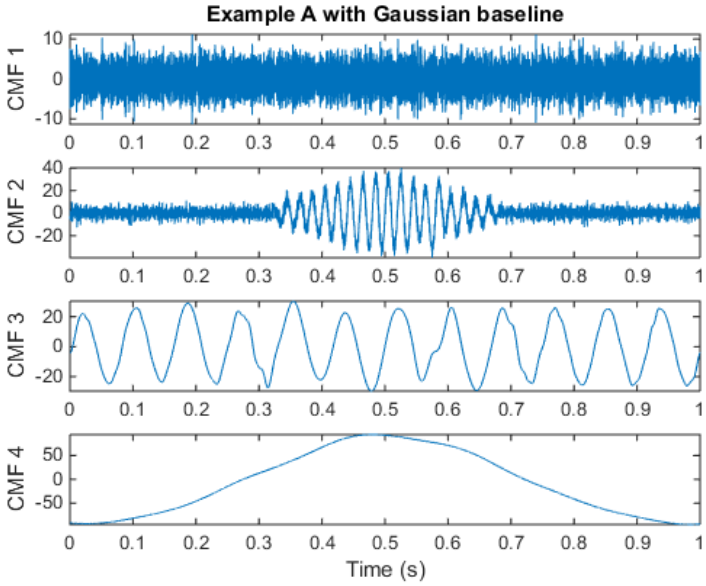


Fig 13

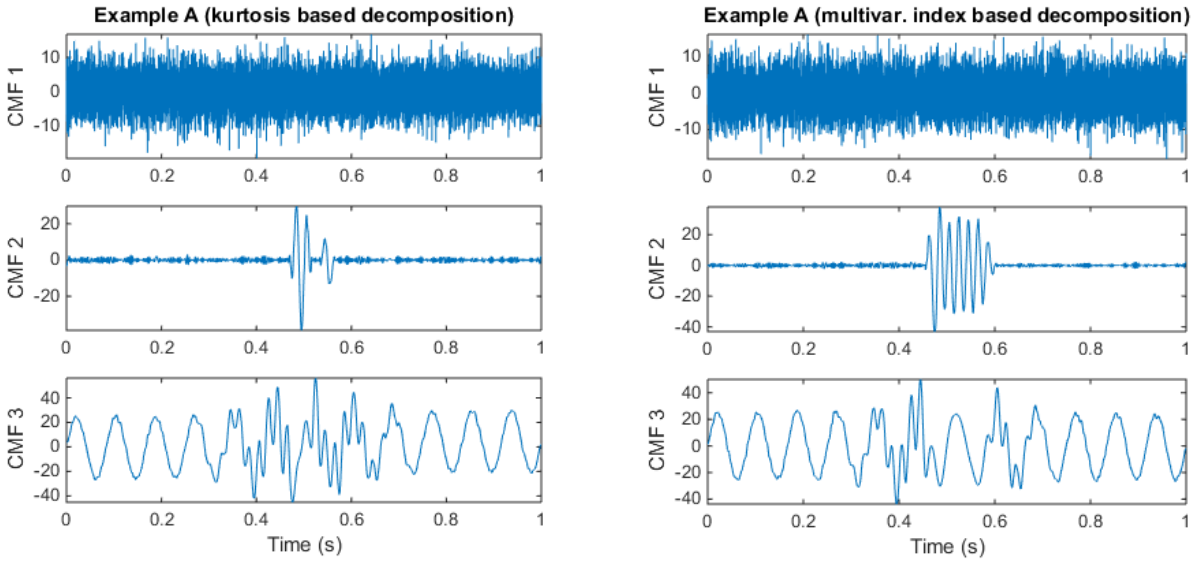
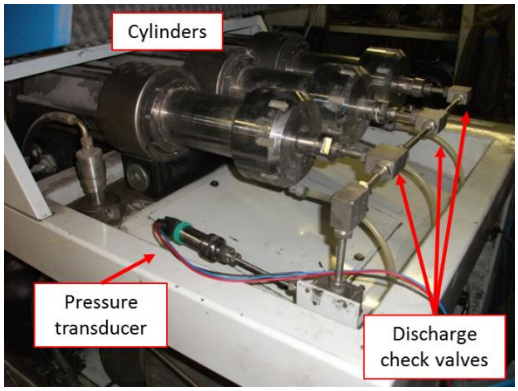

Example A (kurtosis based decomposition)
CMF 1
CMF 2
CMF 3
Time (s)
Example A (multivar. index based decomposition)
CMF 1
CMF 2
CMF 3
Time (s)

Fig 14


Cylinders
Pressure transducer
Discharge check valves

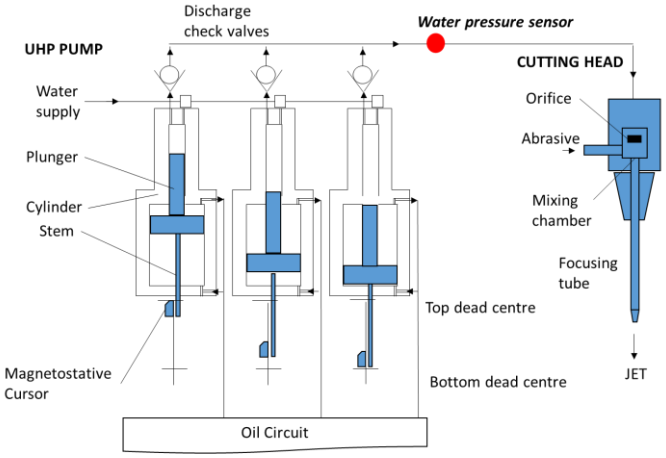

Discharge check valves
Water pressure sensor
UHP PUMP
CUTTING HEAD
Water supply
Orifice
Abrasive
Plunger
Cylinder
Stem
Mixing chamber
Focusing tube
Top dead centre
Magnetostative Cursor
Bottom dead centre
JET
Oil Circuit

Fig 15

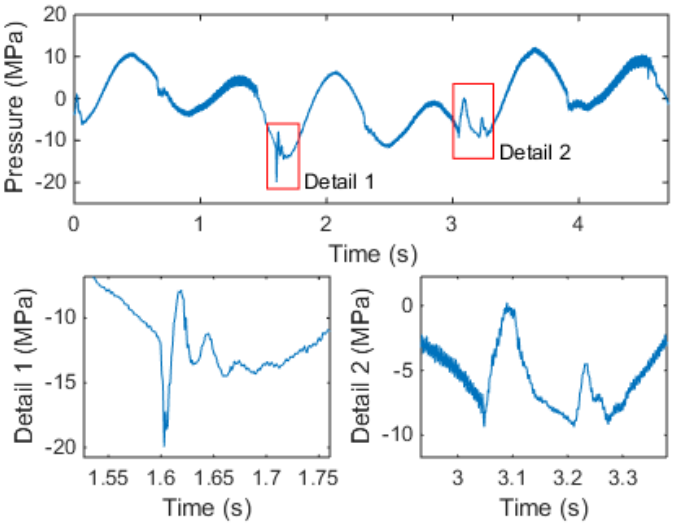

Pressure (MPa)
20
10
0
-10
-20
0
1
2
3
4
Time (s)
Detail 1
Detail 2
Detail 1 (MPa)
-10
-15
-20
1.55
1.6
1.65
1.7
1.75
Time (s)
Detail 2 (MPa)
0
-5
-10
3
3.1
3.2
3.3
Time (s)

Fig 16

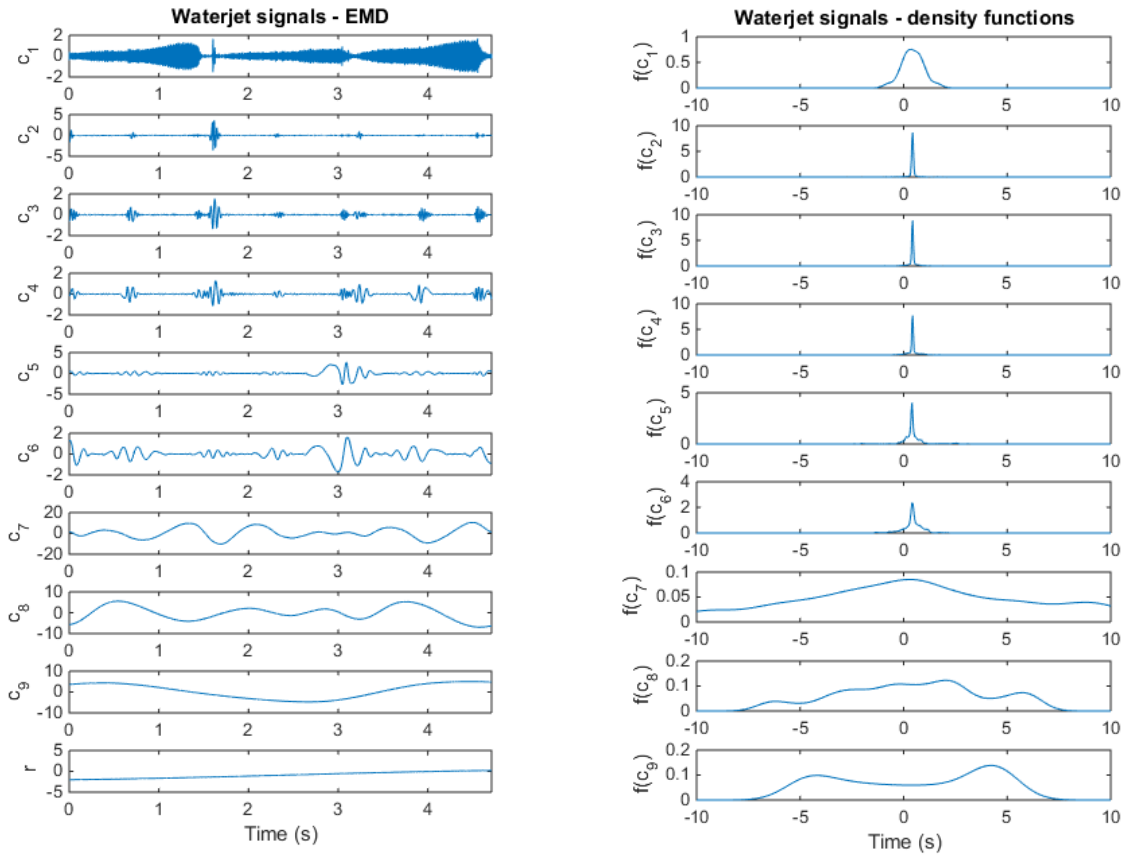
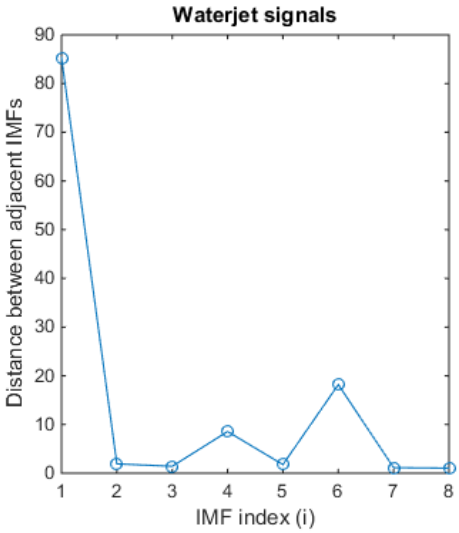

Waterjet signals - EMD
Time (s)
Waterjet signals - density functions
Time (s)

Fig 17


Waterjet signals
Distance between adjacent IMFs
IMF index (i)
0
10
20
30
40
50
60
70
80
90
1
2
3
4
5
6
7
8

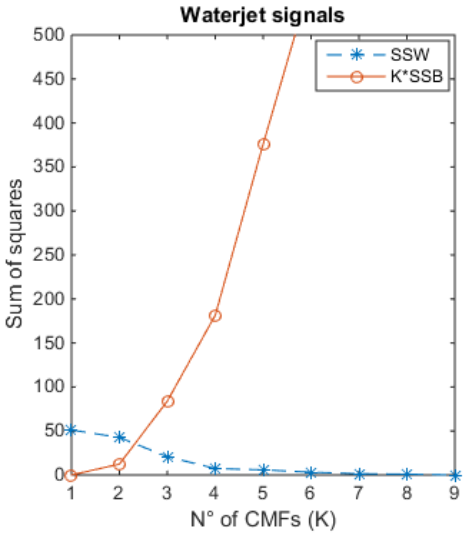

Waterjet signals
SSW
K*SSB
Sum of squares
N° of CMFs (K)
0
50
100
150
200
250
300
350
400
450
500
1
2
3
4
5
6
7
8
9

Fig 18

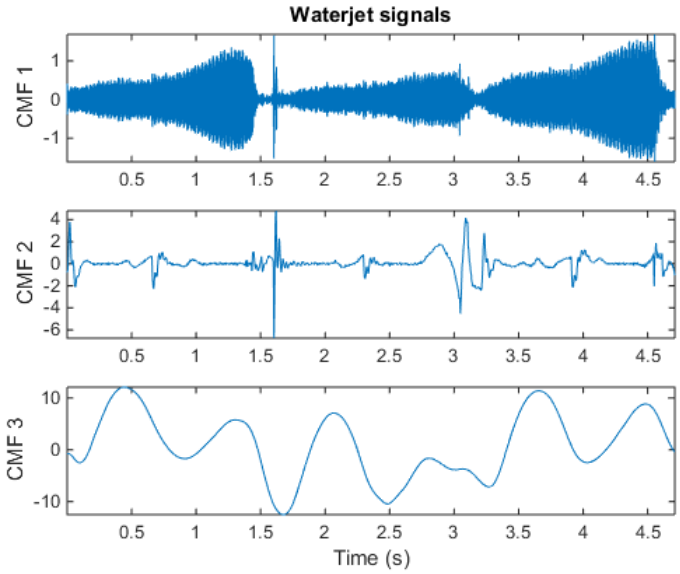


Fig 19

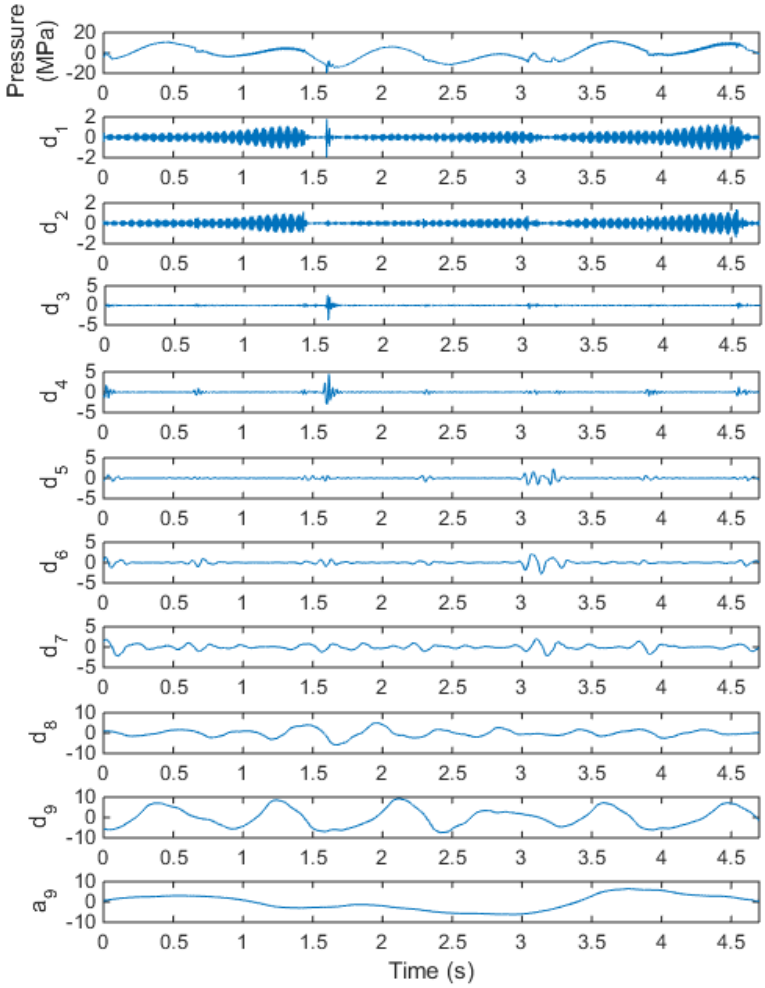


Fig 20

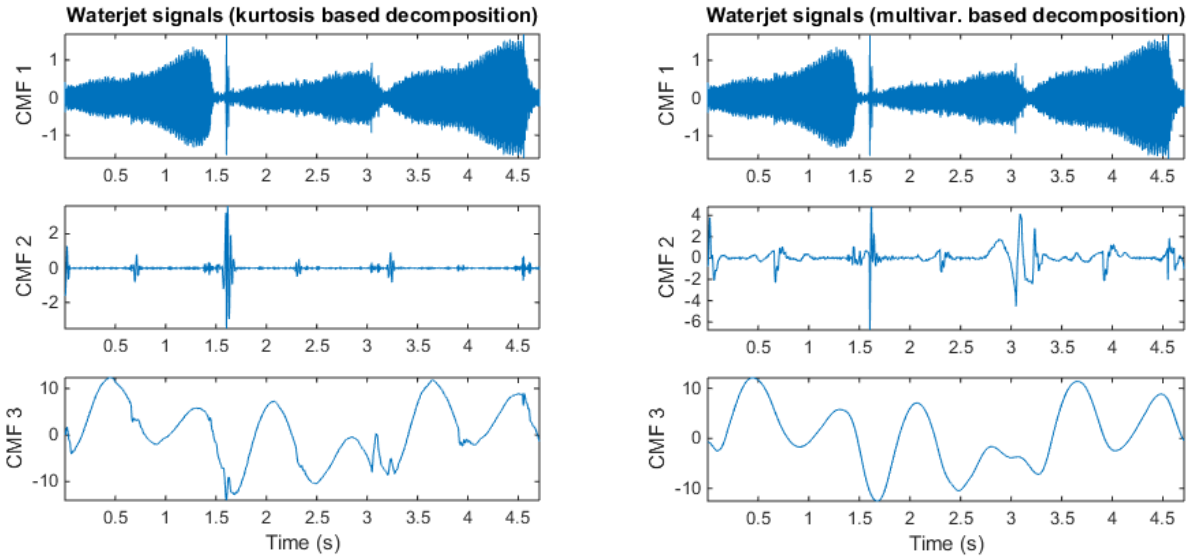

Waterjet signals (kurtosis based decomposition)
CMF 1
CMF 2
CMF 3
Time (s)
Waterjet signals (multivar. based decomposition)
CMF 1
CMF 2
CMF 3
Time (s)

Table 1

	Methods for IMF sub-grouping		Frequency of selection of *K* CMFs (%)			
			K=2	*K=3*	*K=4*	*K>4*
Example A	**Proposed approach** (density-based distance)		**1.9**	**92.7**	**5.4**	**0.0**
	Index-based distance (univariate)	Energy, $E_{i,j}$	0.7	85.4	13.9	0.0
		Corr. with signal, $C_{i,j}$	0.4	71.6	28.0	0.0
		Range, $R_{i,j}$	0.4	86.8	12.8	0.0
		Kurtosis, $K_{i,j}$	0.5	99.4	0.1	0.0
		Skewness, $S_{i,j}$	3.0	82.4	13.6	1.0
	Index-based distance (multivariate)		0.3	99.7	0.0	0.0
Example B	**Proposed approach** (density-based distance)		**0.1**	**99.7**	**0.2**	**0.0**
	Index-based distance (univariate)	Energy, $E_{i,j}$	3.6	96.1	0.3	0.0
		Corr. with signal, $C_{i,j}$	55.3	31.7	13.0	0.0
		Range, $R_{i,j}$	0.0	97.4	2.6	0.0
		Kurtosis, $K_{i,j}$	0.0	100	0.0	0.0
		Skewness, $S_{i,j}$	6.8	78.1	14.7	0.4
	Index-based distance (multivariate)		0.0	100	0.0	0.0

Table 2

	Methods for the selection of K	Frequency of selection of *K* CMFs (%)			
		K=2	*K=3*	*K=4*	*K>4*
Example A	**Proposed criterion** ($SSW_f(K) > K * SSB_f(K)$)	**1.9**	**92.7**	**5.4**	**0.0**
	Ball & Hall criterion	2.2	93.4	4.4	0.0
	Ball & Hall criterion (unpenalized)	1.2	94.7	4.1	0.0
	Zhao et al. criterion	1.8	94.0	4.2	0.0
Example B	**Proposed criterion** ($SSW_f(K) > K * SSB_f(K)$)	**0.1**	**99.7**	**0.2**	**0.0**
	Ball & Hall criterion	0.2	99.5	0.3	0.0
	Ball & Hall criterion (unpenalized)	0.6	99.2	0.2	0.0
	Zhao et al. criterion	0.4	99.3	0.3	0.0

Table 3

Methods for IMF sub-grouping		Frequency of selection of *K* CMFs (%)			
		K=2	*K=3*	*K=4*	*K>4*
Proposed approach (density-based distance)		**2**	**98**	**0**	**0**
Index-based distance (univariate)	Energy, $E_{i,j}$	96	4	0	0
	Corr. with signal, $C_{i,j}$	83	17	0	0
	Range, $R_{i,j}$	73	25	2	0
	Kurtosis, $K_{i,j}$	0	99	1	0
	Skewness, $S_{i,j}$	2	85	13	0
Index-based distance (multivariate)		5	94	1	0

Table 4

Methods for the selection of K	Frequency of selection of *K* CMFs (%)			
	K=2	*K=3*	*K=4*	*K>4*
Proposed criterion ($SSW_f(K) > K * SSB_f(K)$)	**2**	**98**	**0**	**0**
Ball & Hall criterion	100	0	0	0
Ball & Hall criterion (unpenalized)	5	80	15	0
Zhao et al. criterion	100	0	0	0

Table B.1

Index name	Formula
Ball & Hall index	$BH(K) = SSW_f(K)/K$
Calinski & Harabasz index	$CH(K) = \frac{SSB_f(K)/(K-1)}{SSW_f(K)/(n_j - K)}$
Hartigan index	$H(K) = -\log\left(SSW_f(K)/SSB_f(K)\right)$
Zhao et al. index	$Z(K) = K\ \left(SSW_f(K)/SSB_f(K)\right)$