Mechanical Systems and Signal Processing 33 (2012) 237–255

Contents lists available at SciVerse ScienceDirect

Mechanical Systems and Signal Processing

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

Maximum correlated Kurtosis deconvolution and application on gear tooth chip fault detection ☆

Geoff L. McDonald [a], Qing Zhao [a,*], Ming J. Zuo [b]

[a] Advanced Control Systems Laboratory, Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada
[b] Reliability Research Lab, Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada

ARTICLE INFO

Article history:
Received 11 May 2011
Received in revised form
26 April 2012
Accepted 5 June 2012
Available online 15 July 2012

Keywords:
Gear tooth fault diagnosis
Gear tooth chip
Minimum entropy deconvolution
Maximum correlated Kurtosis deconvolution
Correlated Kurtosis
Autoregressive

ABSTRACT

In this paper a new deconvolution method is presented for the detection of gear and bearing faults from vibration data. The proposed maximum correlated Kurtosis deconvolution method takes advantage of the periodic nature of the faults as well as the impulse-like vibration behaviour associated with most types of faults. The results are compared to the standard minimum entropy deconvolution method on both simulated and experimental data. The experimental data is from a gearbox with gear chip fault, and the results are compared between healthy and faulty vibrations. The results indicate that the proposed maximum correlated Kurtosis deconvolution method performs considerably better than the traditional minimum entropy deconvolution method, and often performs several times better at fault detection. In addition to this improved performance, deconvolution of separate fault periods is possible; allowing for concurrent fault detection. Finally, an online implementation is proposed and shown to perform well and be computationally achievable on a personal computer.

1. Introduction

Detecting gear faults has applications in rotating machinery fields such as wind turbines [1] and helicopter transmissions [2]. Detecting and diagnosing gear faults are important to maintenance planning, preventing equipment damage, and preventing failure. In some applications, such as helicopter transmissions, a gear fault can potentially result in a life-threatening situation [3].

The main focus of this paper is gear tooth chip fault detection from accelerometer data, however the presented methodology is expected to be applicable to other impact-faults such as rotor rubbing, rolling element bearing inner and outer-race faults, and other gear tooth faults. These faults manifest in similar ways on the machine vibration data, and the detection methodologies typically work for all of these classes of faults. Research on rotating machine fault detection has remained an active topic over the decades, and existing methods span a large range including model-based methods [4–8], filtering methods [9–12], spectral analysis methods [13–15], and time–frequency analysis methods [16–20]. A few methods particularly of note include wavelet transform-based techniques [16–20], filter selected by spectral Kurtosis [10–12], cyclostationary analysis [14,21,22], and minimum entropy deconvolution [9,10,23]. This paper presents a novel deconvolution approach which takes advantage of the periodicity of the faults.

☆ This work is supported by NSERC and Syncrude Canada Ltd.
* Corresponding author. Tel.: +1 780 492 5792; fax: +1 780 492 1811.
E-mail addresses: glmcdona@gmail.com (G.L. McDonald), qingzhao@ece.ualberta.ca (Q. Zhao), ming.zuo@ualberta.ca (M.J. Zuo).

0888-3270/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ymssp.2012.06.010

For detecting gear and rolling-element bearing faults there has been a growing trend towards autoregressive (AR) model prediction residual fault detection [4,24,9,6,10], which has been shown to be effective in detecting the impulse-like behaviour associated with gear faults. Although this method is shown to be effective on some datasets, the expected residual signal is the fault signal plus noise, and therefore requiring the fault amplitude to be significantly larger than the noise level to be detectable by using the AR model alone. In addition, the AR method requires a priori knowledge of the vibration under no-fault conditions for data-fitting.

Minimum Entropy Deconvolution (MED), originally proposed by Wiggins for application on seismic recordings in 1978 [25], iteratively selects a finite impulse response (FIR) filter to minimize the entropy of the filtered signal and has had widespread applications across many fields. Unlike the AR method, the MED technique aims to extract the fault impulses while minimizing the noise and therefore resulting in clear detection results even under high noise. Endo and Randall [9] proposed applying the AR method followed by MED, forming the method called ARMED and demonstrated the method to be very effective in detecting spalls and tooth cracks. Sawalhi et al. [10] demonstrated the effectiveness of the ARMED process in detecting faults in ball bearing elements. A limitation of the ARMED method is the preference of the MED algorithm to deconvolve only a single impulse or a selection of impulses, as opposed to the desired periodic impulses repeating at the period of the fault.

Inspired by the MED deconvolution technique, this paper proposes an improved novel deconvolution norm, Correlated Kurtosis (*CK*), which takes advantage of the periodicity of the faults and requires no AR model stage prior to deconvolution. The deconvolution technique, maximum correlated Kurtosis deconvolution (MCKD), is proposed to select a FIR filter to maximize the *CK* of the resulting signal which emphasizes high Kurtosis while encouraging periodicity about a specific period. An iterative selection technique to the deconvolution is derived for first and M-shift MCKD, and the results are compared using simulation and experimental data from a controlled gear tooth chip experiment. Despite the faulty gear vibration data showing no visible indication of fault in the original or AR residual data, the deconvolution methods are able to successfully extract the fault clearly, with the proposed MCKD method performing the best. An online threshold alarm implementation of the MCKD method is presented, shown to be computationally achievable, and shown effective on looped experimental data.

In the rest of the paper, review of ARMED for fault detection is given in Section 2. The novel deconvolution norm *CK* is presented in Section 3, along with the *CK* values for some sample input signals. Next, an iterative selection process for the MCKD problem is derived for first and M-shift in Section 4. Simulated deconvolution results on an impulse train plus noise signal are presented for the standard MED and the proposed MCKD method in Section 5.1, from which the advantage of the MCKD method is clearly demonstrated. A simulated concurrent fault case is analysed in Section 5.2, and the results demonstrate the ability to deconvolve faults with different periods separately. Experimental validation, Section 6, is then performed on a controlled gear chip gearbox test, and results are compared among the AR, ARMED, MED and MCDK methods. The proposed MCKD method most clearly identifies the repeating fault in the time domain, and is able to indicate a fault significantly better when comparing fault indicators between fault and no-fault data. Finally a computationally simple online concurrent fault detection implementation of the MCKD fault detection method is presented in Section 7. Validation of the online implementation is performed on looped experimental data, and is shown to have strong fault detection results while being computationally achievable for online application.

2. Review of minimum entropy deconvolution-based fault detection

2.1. Autoregressive model

Autoregressive (AR) models have been a growing trend in rotating machine vibration fault detection and have been shown to be effective in extracting gear faults with little a priori knowledge [4,24,9]. The AR system model with no input has structure

$$y_n = -a_1 y_{n-1} - a_2 y_{n-2} - \cdots - a_N y_{n-N} + e_n$$

where a_i are scalar model coefficients, N is the order of the AR model, e_n is the white noise, and y_n is the signal being modelled. That is, the current sample is a linear combination of the N previous samples plus additive white noise. To solve for the scalar model coefficients, a_i, there are several approaches. For this paper Burg's lattice-based method [26] is applied due to the robustness of estimation. This method selects the parameters by minimizing the least-squares of both the forward and backward prediction errors.

Fault detection by AR model is approached through the following steps:

Step 1: Select AR model order N. This order is often selected by Akaike information criterion.
Step 2: Fit the AR model to the no-fault data by calculating a_is by Burg's lattice-based method.
Step 3: Perform 1-step ahead prediction on the potentially faulty vibration data and calculate the prediction error, Fig. 1.

The prediction error consists of white noise, disturbances, and potentially some trended data from system dynamics changes. The impulse-like faults associated with gear cracks are expected to be more prominent in this residual.

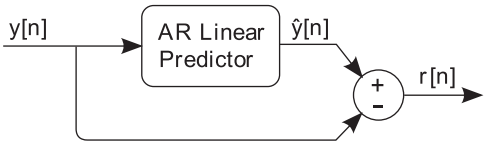

Fig. 1. AR prediction residual method. $y[n]$, $\hat{y}[n]$, and $r[n]$ are the input signal, predicted signal, and prediction residual respectively.

Several major drawbacks exist for this proposed method. First of all and most importantly, the expected residual is white noise plus fault impulses; so the fault signal must be significantly larger in amplitude than the noise e_n for detection. And secondly, it requires knowledge of the system under no-fault conditions. The application of MED in the next section helps to resolve these issues.

2.2. Minimum entropy deconvolution

MED was originally proposed for application on seismic recordings by Wiggins in 1978 [25] and recently applied to gear fault detection by Endo and Randall [9] in 2007. MED poses a deconvolution problem where a FIR filter is selected to minimize the entropy of the filtered signal.

Starting from a general linear time-invariant machine acceleration signal model:

$$x_n = (-a_1 x_{n-1} - a_2 x_{n-2} - \cdots - a_K x_{n-K}) + (b_1 u_{n-1} + b_2 u_{n-2} + \cdots + b_L u_{n-L}) + (c_1 d_{n-1} + c_2 d_{n-2} + \cdots + c_M d_{n-M}) + e_n$$

where x_n is the sampled acceleration signal, u_n is an unknown input sequence, d_n is the repeating impulse-like gear fault input sequence, and e_n is the noise. a_k, b_k, and c_k are scalars representing the dependence of x_n on previous x_n, u_n, and d_n respectively. By taking the z-transform and solving for the system acceleration, we have

$$X = \frac{B(z^{-1})}{A(z^{-1})}U + \frac{C(z^{-1})}{A(z^{-1})}D + \frac{1}{A(z^{-1})}E$$

$$A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_K z^{-K}$$

$$B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \cdots + b_L z^{-L}$$

$$C(z^{-1}) = c_1 z^{-1} + c_2 z^{-2} + \cdots + c_M z^{-M}$$

where X, U, and E are the z-transform of x_n, u_n, and e_n respectively.

Any stable transfer function $P(z^{-1})/Q(z^{-1})$ can be approximated as a FIR filter. Since these transfer functions are clearly stable or marginally stable in this case (otherwise the machine would explode with infinitely growing vibration), the resulting time domain approximation form in terms of convolution is

$$\vec{x} = \vec{h}_u * \vec{u} + \vec{h}_d * \vec{d} + \vec{h}_e * \vec{e}$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}, \quad \vec{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix}$$

where $\vec{h}$'s are the FIR filter approximations to their respective transfer functions under stable assumption and N is the number of measured samples of vibration signal x_n. The deconvolution problem aims to reconstruct the fault signal $\vec{d}$ by applying FIR filter $\vec{f}$ with L samples to measured machine acceleration $\vec{x}$:

$$\vec{y} = \vec{f} * \vec{x} = \vec{f} * (\vec{h}_u * \vec{u}) + \vec{f} * (\vec{h}_d * \vec{d}) + \vec{f} * (\vec{h}_e * \vec{e})$$

$$\vec{f} = [f_1 \ f_2 \ \cdots \ f_L]^T$$

It is desired that the resulting filtered signal $\vec{y}$ approximates fault signal $\vec{d}$ and this is approached by selecting filter $\vec{f}$ to minimize the noise effect $\vec{f} * (\vec{h}_e * \vec{e}) \to \vec{0}$, while closely cancelling the system $\vec{f} * (\vec{h}_u * \vec{u}) \to \vec{0}$, and extracting a shifted approximation to the fault impulse train signal $\vec{f} * (\vec{h}_d * \vec{d}) \approx \vec{d}$. Selection of the filter $\vec{f}$ given only acceleration $\vec{x}$ measurements may seem to be a difficult problem, but the fault signal $\vec{d}$ is expected to be impulse-like (a signal of very high Kurtosis) while competing signals $\vec{u}$ and $\vec{e}$ are of very low Kurtosis. As a result of this Kurtosis difference between the signals, the filter can be selected to reach a maximum in Kurtosis. To achieve this, Wiggins [25] proposed maximizing of a norm function called the Varimax Norm, which in the case of one-dimensional MED is equivalent to maximizing

Kurtosis with assumed zero-mean:

$$\text{Kurtosis} = \frac{\frac{1}{n}\sum_{n=1}^{N}(y_n-\mu_y)^4}{\left(\frac{1}{n}\sum_{n=1}^{N}(y_n-\mu_y)^2\right)^2}-3$$

$$\max_{\vec{f}} \text{Kurtosis} = \max_{\vec{f}} \frac{\sum_{n=1}^{N}(y_n-\mu_y)^4}{(\sum_{n=1}^{N}(y_n-\mu_y)^2)^2}$$

Describing the deconvolution method in terms of Kurtosis is chosen because it is commonly used to quantify the impulse-like fault level of a vibration signal [4,9,11]. Assuming y_n is zero-mean, $\mu_y = 0$:

$$\max_{\vec{f}} \text{Kurtosis} = \max_{\vec{f}} \frac{\sum_{n=1}^{N} y_n^4}{(\sum_{n=1}^{N} y_n^2)^2} \tag{1}$$

By taking the derivatives of Eq. (1) with respect to filter coefficients $\vec{f}$ and solving it equal to zero, an iteratively converging local-maximum solution can be derived as

$$\vec{f} = \frac{\sum_{n=1}^{N} y_n^2}{\sum_{n=1}^{N} y_n^4}(X_0 X_0^T)^{-1} X_0 [y_1^3 \;\; y_2^3 \;\; \dots \;\; y_N^3]^T \tag{2}$$

$$X_0 = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_N \\ 0 & x_1 & x_2 & \dots & x_{N-1} \\ 0 & 0 & x_1 & \dots & x_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & x_{N-L+1} \end{bmatrix}_{L \text{ by } N}$$

where $\vec{f}$ is iteratively selected. The iterative procedure is implemented with MATLAB and available in the External Resources Section. The general procedure is as follows:

Step 1: Assume initial filter as a centred impulse, $\vec{f} = [0\ 0\ \dots\ 1 \dots\ 0\ 0]^T$.
Step 2: Calculate X_0 and $(X_0 X_0^T)^{-1}$ from input signal $\vec{x}$.
Step 3: Calculate $\vec{y}$ as $\vec{y} = X_0^T \vec{f}$.
Step 4: Determine new filter coefficients by solving for $\vec{f}$ in Eq. (2).
Step 5: Repeat from Step 3 for a specified number of iterations or until the change in Kurtosis between iterations is below a specified small value.
Step 6: The final deconvolved signal is calculated as $\vec{y} = X_0^T \vec{f}$.

Endo et al. [9] proposed the fault detection method ARMED, which applies the AR fault detection method followed by MED on the resulting prediction residual. The expected results for MED are approximately the high Kurtosis shifted fault signal $\vec{d}$, unlike the AR method which in the best-case scenario expects the fault signal plus white noise. As a result, the ARMED results in improved performance over the traditional AR method. The AR residual processing stage is commonly applied, mainly to remove the predictable time-invariant components of the vibration as a processing step, e.g. the regular gear meshing signals, hence the visibility of the hard-to-predict local fault signal is improved [4,24,9,6]. However, this AR stage may not be necessary in the case of MED since the maximization problem itself aims to filter out low-Kurtosis components; therefore both ARMED and direct MED techniques are applied for comparison purposes. Also important to note is that in the case of direct MED fault extraction, no a priori knowledge of no-fault machine vibrations is required for fault deconvolution.

One major drawback of the MED technique is that given a finite-length white noise signal x_n, MED is able to successfully deconvolve a single impulse when the filter size L is sufficiently large. This is a commonly seen problem since in the ideal no-fault case, the expected AR model residual is white noise. By then using Kurtosis as the fault comparison, the resulting single impulse may then be of higher Kurtosis than the train of impulses deconvolved in the fault case, improperly indicating the fault-state. Additionally, the MED algorithm by maximizing Kurtosis prefers a solution with the fewest number of impulses. This can sometimes result in deconvolution solutions of fewer-than-desired impulses. These are major drawbacks for the reliability of automated fault detection by MED and ARMED.

3. Correlated Kurtosis

To improve upon the MED deconvolution technique, the periodicity of the fault can be taken advantage of through the definition of a new deconvolution norm. This proposed norm

$$\text{Correlated Kurtosis of first} - \text{shift} = CK_1(T) = \frac{\sum_{n=1}^{N}(y_n y_{n-T})^2}{(\sum_{n=1}^{N} y_n^2)^2}$$

$$\text{Correlated Kurtosis of M} - \text{shift} = CK_M(T) = \frac{\sum_{n=1}^{N}(\prod_{m=0}^{M} y_{n-mT})^2}{(\sum_{n=1}^{N} y_n^2)^{M+1}} \tag{3}$$

$$y_n = \sum_{k=1}^{L} f_k x_{n-k+1}, \quad x_n = 0 \text{ and } y_n = 0 \text{ for } n \neq 1,2,\ldots,N$$

encourages filter output periodicity about a period T and high-Kurtosis, hence the name correlated Kurtosis. N is the number of samples in the input signal $\vec{x}$, L is the length of FIR filter $\vec{f}$, and T is the period of interest. It should be noted that if $T=0$ and $M=1$, then CK is the Kurtosis norm used by MED in Eq. (1). Fig. 2 illustrates the CK_1 versus Kurtosis for several simple signals. It can be seen that the proposed CK_1 approaches a maximum for a periodic impulse about the specified period as opposed to the Kurtosis which tends to a maximum with a single impulse. Higher shift CK emphasizes larger sequences of impulses in a row.

To illustrate the usage of CK towards extracting fault signals, we compare the first-shift CK values for a simple simulated vibration example where we have a sinusoidal vibration with an impact once per rotation:

Signal 1: $y_n = \sin(2\pi n/100) + 0.5\sin(4\pi n/100) + \text{noise}$,
Signal 2: $y_n = 0.3\sum_{k=0}^{\inf}\delta_{n-k100} + 1[\sin(2\pi n/100) + 0.5 * \sin(4\pi n/100) + \text{noise}]$,
Signal 3: $y_n = 0.3\sum_{k=0}^{\inf}\delta_{n-k100} + 0.5[(\sin(2\pi n/100) + 0.5\sin(4\pi n/100)) + \text{noise}]$,
Signal 4: $y_n = 0.3\sum_{k=0}^{\inf}\delta_{n-k100} + 0.2[(\sin(2\pi n/100) + 0.5\sin(4\pi n/100)) + \text{noise}]$,
Signal 5: $y_n = 0.3\sum_{k=0}^{\inf}\delta_{n-k100}$,

where $\delta_k = 1$ for $k=0$ and $\delta_k = 0$ otherwise, and $n=1,2,\ldots,2000$. The noise is zero-mean white noise with variance of 0.22. Fig. 3 illustrates $CK_1(100)$ for these signals and it is clear that the maximum of these signals is the fault signal by itself,

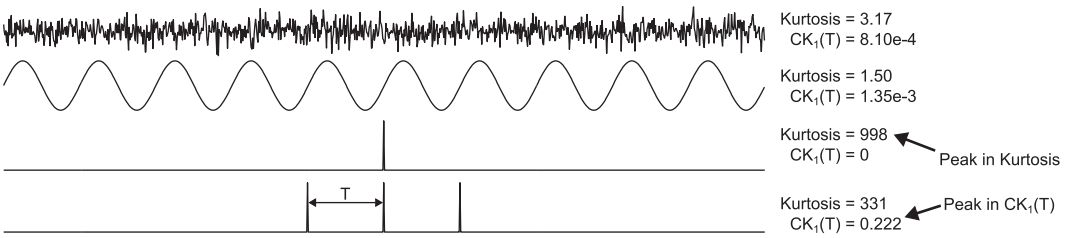

Fig. 2. Kurtosis and CK values for several signals. The Kurtosis reaches a maximum with a single impulse, the CK_1 reaches a maximum with three impulses. The first signal is white noise.

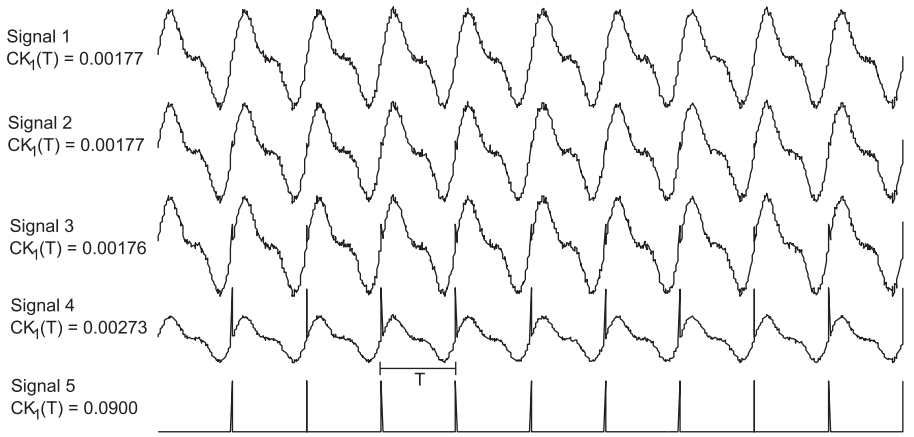

Fig. 3. CK_1 values for a simple simulated repetitive fault with various amplitudes of the base harmonic vibration. All signals are normalized by their peak value for better illustration.

Signal 5. The goal is to extract the fault impulses through maximizing *CK*, and this problem is approached through a deconvolution method presented in the following section. Special consideration is taken in the next section by initial conditions to prevent the local maximum solution achieved by Signal 1.

4. Maximum correlated Kurtosis deconvolution

4.1. First-shift maximum correlated Kurtosis deconvolution

The MCKD technique aims to maximize *CK* for input signal, $\overrightarrow{x}$, about the period, *T*, by selecting a FIR filter $\overrightarrow{f}$. First we only consider the first-shift MCKD algorithm, and then expand the method to M-shift in the next section. Starting from the maximization problem

$$MCKD_1(T) = \max_{\overrightarrow{f}} CK_1(T) = \max_{\overrightarrow{f}} \frac{\sum_{n=1}^{N}(y_n y_{n-T})^2}{(\sum_{n=1}^{N} y_n^2)^2} \tag{4}$$

$$\overrightarrow{f} = [f_1 \ f_2 \ \cdots \ f_L]^T$$

Towards solving for the filter coefficients corresponding to the maximum, we solve

$$\frac{d}{df_k} CK_1(T) = 0, \quad k = 1, 2, \ldots, L \tag{5}$$

First, the derivatives of the numerator and denominator portions are solved separately. Starting with the numerator

$$\frac{d}{df_k} CK_1 \text{ Numerator} = \frac{d}{df_k} \sum_{n=1}^{N} (y_n y_{n-T})^2 = \sum_{n=1}^{N} 2 y_n y_{n-T} \frac{d}{df_k} y_n y_{n-T} = \left(\sum_{n=1}^{N} 2 y_n y_{n-T}^2 \frac{d}{df_k} y_n \right) + \left(\sum_{n=1}^{N} 2 y_n^2 y_{n-T} \frac{d}{df_k} y_{n-T} \right)$$

and since

$$\frac{d}{df_k} y_n = x_{n-k+1}$$

resulting in

$$\frac{d}{df_k} CK_1 \text{ Numerator} = \sum_{n=1}^{N} 2 x_{n-k+1} y_n y_{n-T}^2 + \sum_{n=1}^{N} 2 x_{n-T-k+1} y_{n-T} y_n^2 \tag{6}$$

Similarly for the denominator

$$\frac{d}{df_k} CK_1 \text{ Denominator} = \frac{d}{df_k} \left(\sum_{n=1}^{N} y_n^2 \right)^2 = 2 \left(\sum_{n=1}^{N} y_n^2 \right) \frac{d}{df_k} \sum_{n=1}^{N} y_n^2 = 4 \left(\sum_{n=1}^{N} y_n^2 \right) \sum_{n=1}^{N} y_n x_{n-k+1} \tag{7}$$

From Eqs. (5) to (7) it follows:

$$\frac{d}{df_k} CK_1(T) = 2 \| \overrightarrow{y} \|^{-4} \left(\sum_{n=1}^{N} x_{n-k+1} y_n y_{n-T}^2 + \sum_{n=1}^{N} x_{n-T-k+1} y_{n-T} y_n^2 \right) - 4 \| \overrightarrow{y} \|^{-6} \sum_{n=1}^{N} (y_n y_{n-T})^2 \sum_{n=1}^{N} y_n x_{n-k+1}$$

And rewriting in matrix form:

$$\frac{d}{d\overrightarrow{f}} CK_1(T) = \overrightarrow{0} = 2 \| \overrightarrow{y} \|^{-4} (X_0 \overrightarrow{\alpha}_0 + X_T \overrightarrow{\alpha}_1) - 4 \| \overrightarrow{y} \|^{-6} \| \overrightarrow{\beta} \|^2 X_0 \overrightarrow{y} \tag{8}$$

$$X_r = \begin{bmatrix} x_{1-r} & x_{2-r} & x_{3-r} & \cdots & x_{N-r} \\ 0 & x_{1-r} & x_{2-r} & \cdots & x_{N-1-r} \\ 0 & 0 & x_{1-r} & \cdots & x_{N-2-r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x_{N-L-r+1} \end{bmatrix}_{L \text{ by } N}$$

$$\overrightarrow{\alpha}_0 = [y_1 y_{1-T}^2 \ y_2 y_{2-T}^2 \ \cdots \ y_N y_{N-T}^2]^T, \quad \overrightarrow{\alpha}_1 = [y_{1-T} y_1^2 \ y_{2-T} y_2^2 \ \cdots \ y_{N-T} y_N^2]^T$$

$$\overrightarrow{\beta} = [y_1 y_{1-T} \ y_2 y_{2-T} \ \cdots \ y_N y_{N-T}]^T$$

Rearranging Eq. (8)

$$2 \| \overrightarrow{\beta} \|^2 X_0 \overrightarrow{y} = \| \overrightarrow{y} \|^2 (X_0 \overrightarrow{\alpha}_0 + X_T \overrightarrow{\alpha}_1)$$

and substituting the relation

$$\vec{y} = X_0^T \vec{f} \tag{9}$$

results in

$$X_0 X_0^T \vec{f} = \frac{\|\vec{y}\|^2}{2\|\vec{\beta}\|^2}(X_0 \vec{\alpha}_0 + X_T \vec{\alpha}_1)$$

The matrix $X_0 X_0^T$ is the Toeplitz autocorrelation matrix of $\vec{x}$ and the inverse $(X_0 X_0^T)^{-1}$ is assumed to exist

$$\vec{f} = \frac{\|\vec{y}\|^2}{2\|\vec{\beta}\|^2}(X_0 X_0^T)^{-1}(X_0 \vec{\alpha}_0 + X_T \vec{\alpha}_1) \tag{10}$$

This resulting equation is nonlinear, but a local maximum solution for $\vec{f}$ can be solved iteratively. All datasets processed so far have been found to be monotonically convergent to a local maximum solution. The procedure for solving for $\vec{f}$ iteratively is as follows, and a link to the MATLAB implementation can be found in the External Resources Section:

Step 1: Select period of interest, T. Ensure T is within the range of approximately 20–300, and if T is in excess of this range the dataset should firstly be downsampled.
Step 2: Calculate X_T, X_0^T, and $(X_0 X_0^T)^{-1}$ from your input signal $\vec{x}$.
Step 3: Select filter size, L, and assume an initial filter of $\vec{f} = [0\ 0\ \cdots\ 1\ -1\ \cdots\ 0\ 0]^T$. This is selected as a difference filter to prevent the algorithm from converging to the local solution of Signal 1 in Fig. 3 and the difference is in the centre because performance can be improved through not assuming a minimum phase filter.
Step 4: Calculate the filtered output, $\vec{y}$, from Eq. (9).
Step 5: Calculate $\vec{\alpha}_0$, $\vec{\alpha}_1$, and $\vec{\beta}$ from $\vec{y}$.
Step 6: Calculate the new filter coefficient, $\vec{f}$, from Eq. (10).
Step 7: Is $\Delta CK_1(T) > \epsilon$? Loop from Step 4 while true. ϵ is a small positive number controlling when the algorithm terminates and ΔCK_1 denotes the change for the iteration.
Step 8: The final first-shift MCKD-filtered signal about period T is calculated from Eq. (9).

4.2. M-shift maximum correlated Kurtosis deconvolution

Results generally improve significantly on experimental data by using a higher shift MCKD method because increasing the shift increases the number of sequential impulses the algorithm is looking to deconvolve. This performance improvement in fault versus no-fault comparison is mostly from the reduced likelihood of erroneously deconvolving periodic impulses in the no-fault datasets. However, higher order shifts require better estimates of the fault period T and increase the complexity of the calculation.

Similar to the first-shift derivation, we start from the maximization problem

$$MCKD_M(T) = \max_{\vec{f}} CK_M(T) = \max_{\vec{f}} \frac{\sum_{n=1}^{N}(\prod_{m=0}^{M} y_{n-mT})^2}{(\sum_{n=1}^{N} y_n^2)^{M+1}} \tag{11}$$

and by solving the derivative of the numerator and denominator of $CK_M(T)$ with respect to filter coefficients f_k:

$$\frac{d}{df_k} CK_M \text{ Numerator} = 2\sum_{n=1}^{N}\left[\left(\prod_{m=0}^{M} y_{n-mT}\right)^2 \left(\sum_{m=0}^{M} \frac{x_{n-mT-k+1}}{y_{n-mT}}\right)\right] \tag{12}$$

$$\frac{d}{df_k} CK_M \text{ Denominator} = 2(M+1)\|\vec{y}\|^{2M} \sum_{n=1}^{N} y_n x_{n-k+1} \tag{13}$$

Combining Eqs. (12) and (13) results in

$$\frac{d}{df_k} CK_M(T) = 0 = 2\|\vec{y}\|^{-2M-2} \sum_{n=1}^{N}\left[\left(\prod_{m=0}^{M} y_{n-mT}\right)^2 \left(\sum_{m=0}^{M} \frac{x_{n-mT-k+1}}{y_{n-mT}}\right)\right] - 2(M+1)\|\vec{y}\|^{-2M-4} \sum_{n=1}^{N}\left(\left(\prod_{m=0}^{M} y_{n-mT}\right)^2\right) \sum_{n=1}^{N} y_n x_{n-k+1}$$

and converting to matrix form with $k = 1, 2, \ldots, L$ and rearranging results in the iterative solution

$$\vec{f} = \frac{\|\vec{y}\|^2}{2\|\vec{\beta}\|^2}(X_0 X_0^T)^{-1} \sum_{m=0}^{M} X_{mT} \vec{\alpha}_m \tag{14}$$

$$\overrightarrow{\alpha}_m = \begin{bmatrix} y_{1-mT}^{-1}(y_1^2 y_{1-T}^2 \cdots y_{1-MT}^2) \\ y_{2-mT}^{-1}(y_2^2 y_{2-T}^2 \cdots y_{2-MT}^2) \\ \vdots \\ y_{N-mT}^{-1}(y_N^2 y_{N-T}^2 \cdots y_{N-MT}^2) \end{bmatrix}, \quad \overrightarrow{\beta} = \begin{bmatrix} y_1 y_{1-T} \cdots y_{1-MT} \\ y_2 y_{2-T} \cdots y_{2-MT} \\ \vdots \\ y_N y_{N-T} \cdots y_{N-MT} \end{bmatrix}$$

Eq. (14) can be processed iteratively by following a similar procedure as to the first-shift MCKD presented above. Unlike the first-shift MCDM, the solution does not necessarily converge monotonically, and as a result the filter $\overrightarrow{f}$ is chosen as the filter which results in the maximum in CK_M during the iterative process. For large M, from experience around eight or more, the iterative method can result in loss of numerical precision because of exceeding the range of the floating point exponent.

Because the higher-shift MCKD requires better estimates of the fault period T and in application the period is a fractional number (e.g. 170.21 samples per gear revolution), an additional resampling stage is introduced as a preprocessing step. This additional step resamples the input signal $\overrightarrow{x}$ at a ratio of

$$\lfloor 20T + 0.5 \rfloor : \lfloor 20\lceil T \rceil + 0.5 \rfloor$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denotes the floor and ceiling operations respectively and the factor of 20 is chosen as a good balance between computation time and precision. This resamples the data so that the samples per revolution are approximately the nearest larger integer; for example 170.21 samples per gear revolution resamples at a ratio of 3404:3600, resulting in 180.01 samples per gear revolution. The resampling is performed by polyphase filter implementation due to the computational efficiency of the method. This choice of resampling method should be investigated as future work since it may affect computational or fault-detection performance.

The implementation of this method in MATLAB is available in the External Resources Section.

5. Simulation results

5.1. Impulse train with noise

To more clearly explain the difference between the MED and the effect of the MCKD shift-order, a simple deconvolution of impulses from white noise is analysed. The signal is formed as

$$x_n = e_n + f\left(\sum_{k=0}^{\inf} \delta_{n-k100}\right), \quad n = 0, 1, 3, \ldots, 999$$

where e_n is zero-mean Gaussian white noise of 1 standard deviation and the fault amplitude, f, is varied between 0 and 5 at a step of 0.01. This signal is not a realistic example of rotating machine fault detection signals since each impulse is extracted by a nearly independent deconvolution path, while in practical cases the deconvolution paths for the impulses are similar. This simulation is structured in this manner to clearly illustrate the conceptual difference between MED and M-shift MCKD. Fig. 4a illustrates the fault detection versus f for i.i.d. e_n for each f, while Fig. 4b illustrates the resulting deconvolved signals when $f=3$. A deconvolution filter size of 200 and iteration count of 100 is used for each method. From these results the conceptual difference in increasing MCKD shift level is illustrated; where MCKD deconvolves more and more impulses with not-necessarily the same deconvolution path as the shift level increases. This clearly illustrates one of the problems with the standard MED algorithm, which prefers to deconvolve a single impulse as the result.

5.2. Concurrent fault detection

Consider the simple vibration model

$$x_n = \sin(2\pi n/30) + 0.7\sin(3\pi n/30) + f_n + e_n$$

where f_n denotes the periodic fault signal caused by the faulty gear, e_n is the zero-mean additive white Gaussian noise with a standard deviation of 0.1, and $n = 0, 2, \ldots, 2999$. Then consider two faulty elements in the system on with fault periods of 30 and 100 samples. These faulty elements result in impulse-like vibration signals convolved with two separate transmission paths

$$f_n^a = h_n^a * \left(\sum_{k=0}^{\infty} \delta_{n-k30}\right)$$

$$f_n^b = h_n^b * \left(\sum_{k=0}^{\infty} \delta_{n-k100}\right)$$

Impulse responses h_n^a and h_n^b refer to two different vibration transmission paths from the faulty gears to vibration sensor, combined with slightly different fault characteristics. We select these two transmission paths as

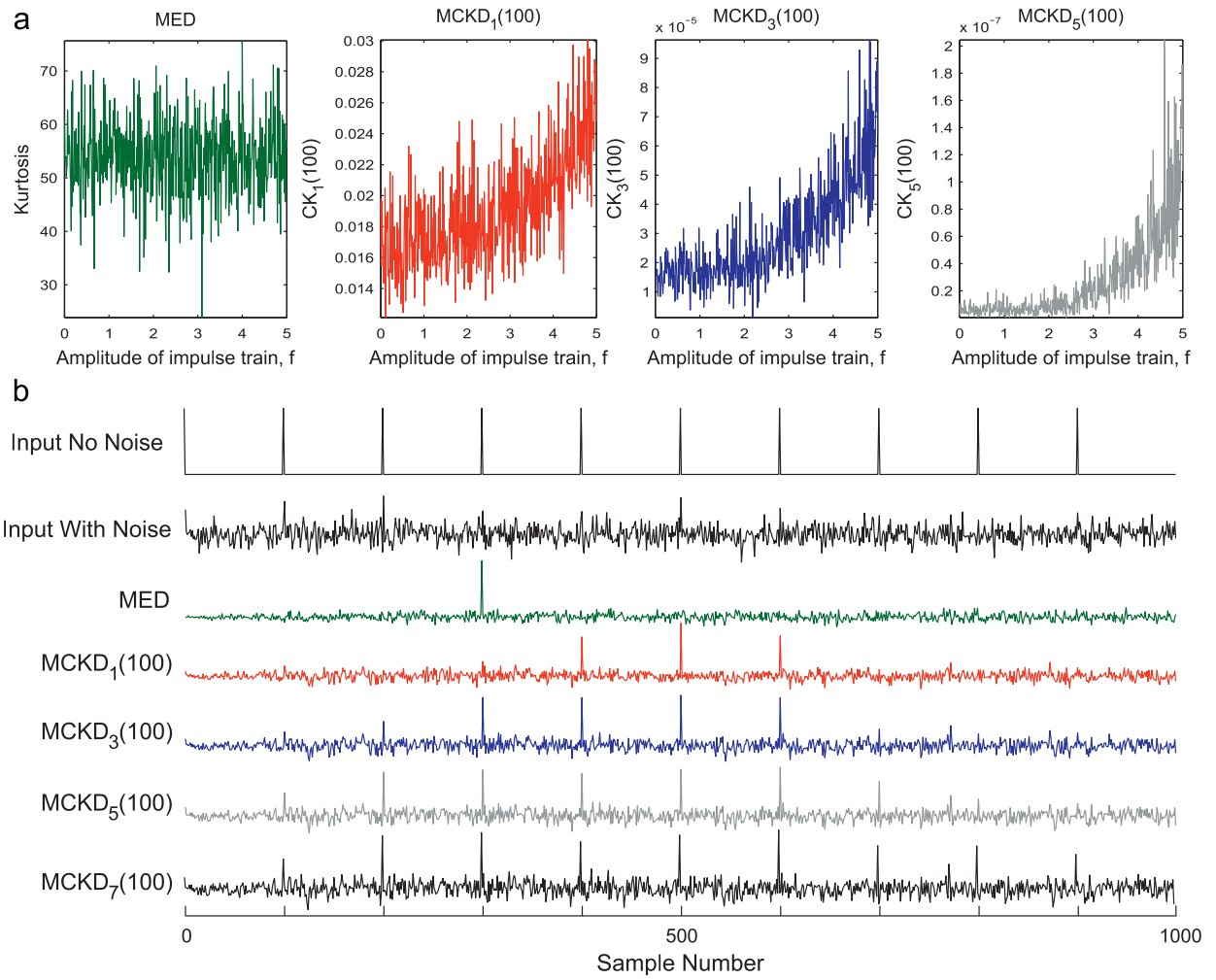

Fig. 4. Deconvolution of a noisy periodic impulse train of period 100 with (a) illustrating the fault detection versus impulse train amplitude with noise, and (b) normalized plot of the resulting deconvolved signals with impulse train amplitude of $f=3$, some signals are polarity inverted for illustration purposes. (Kurtosis and correlated Kurtosis are independent of the amplitude or polarity of the signal, so the filter designed by the deconvolution process can be of any amplitude or polarity. Better illustration is achieved by normalizing the resulting signals by peak value and ensuring the impulses are of the same polarity.) A different white noise seed is used for each value of f.

$$\vec{h}^{a} = [0.9 \ \ -0.3 \ \ 0.1 \ \ -0.05]^T$$

$$\vec{h}^{b} = [0.1 \ \ 0.8 \ \ -0.7 \ \ -0.6 \ \ 0.4 \ \ -0.2 \ \ 0.1 \ \ -0.05]^T$$

and form four cases:

$F1$: no fault: $f_n=0$,
$F2$: fault on period 30: $f_n = 0.4f_n^a$,
$F3$: fault on period 100: $f_n = 0.4f_n^b$,
$F4$: concurrent faults on periods 30 and 100: $f_n = 0.4f_n^a + 0.4f_n^b$.

The desired output signals include only the impulse fault signals before the transmission path. The resulting fault indicators, $CK_M(T)$ for MCKD and the Kurtosis for MED-based methods, are tabulated in Table 1. All algorithms were with filter size of 200, and the MED and MCKD stages were iterated exactly 100 times each. Fifth-shift MCKD was chosen as a good balance between a higher shift method and lower computation time. AR model orders of 20 were chosen because it is a fairly high model order with respect to the system dynamics order. Higher numbers indicate a larger indicated fault. From the table, it is clear to see that only the MCKD algorithm detects the faults successfully. The MED technique performs

Table 1
Fault indicator values by MCKD, MED, and ARMED. The fault indicators are final *CK* values for MCKD and the final Kurtosis values for MED and ARMED. All values are no-fault normalized.

Method	*F*1, no fault	*F*2, fault on period 30	*F*3, fault on period 100	*F*4, concurrent faults on periods 30 and 100
$MCKD_5(30)$	1	9.75	0.929	8.55
$MCKD_5(100)$	1	2.62	79.8	77.4
MED	1	0.139	0.111	0.137
ARMED	1	1.01	1.09	0.995

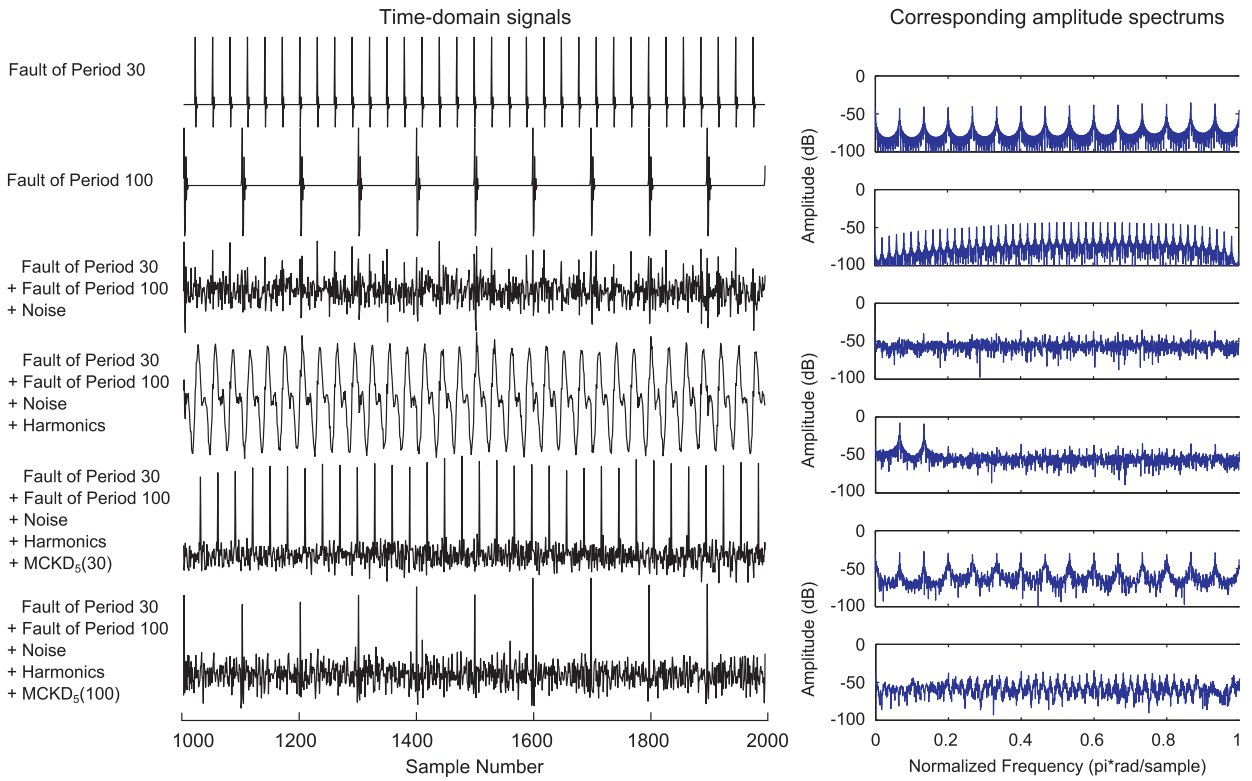

Fig. 5. Simulated resulting signals for concurrent fault detection after processing with indicated method. All signals are normalized and some signals are polarity flipped for better illustration.

poorly as a result of incorrectly deconvolving a single impulse from the no-fault case, and hence a high fault indicator for the no-fault case. The MCKD method is not only able to strongly detect the faults, but is also able to indicate exactly which fault periods are present in each case.

Fig. 5 illustrates the resulting signals for the MCKD approach under concurrent fault condition, *F*4. It is clear that the faults with a period of 30 and 100 are successfully isolated separately during the deconvolution process, despite the fact that the noise level seems to obfuscate the fault signals entirely.

Additionally, the simple fault cases *F*1 (no fault) and *F*3 (fault of period 100) are compared using the complex shifted Morlet wavelet transform [17]. The complex shifted Morlet wavelet transform is a Morlet wavelet-based narrow-band filter approach with centre frequency varied across the frequency range. The bandwidth is fixed across all centre frequencies. Results for these two simple fault cases are illustrated in Fig. 6 for the discrete Fourier transform and complex shifted Morlet wavelet transform. A slight difference can be observed in the Fourier spectrum, the repeating fault can be slightly observed in the wide-bandwidth complex shifted Morlet wavelet transform, and the fault can barely be seen in the narrow-bandwidth complex shifted Morlet wavelet transform. In the wide-bandwidth case, Fig. 6b, the narrow-band envelope for centre frequency of 0.75π rad/sample is compared for no fault and fault cases. This centre frequency is chosen for investigation since the scalogram for the fault signal appears to indicate the fault best around this centre frequency. The Kurtosis of this envelope indicates the fault correctly, but not nearly as strong as the difference observed for MCKD in Table 1 and illustrated in Fig. 5. Unlike standard time–frequency and filtering methods, the deconvolution methods aim to actively cancel the noise—as opposed to extracting a subset of the noise.

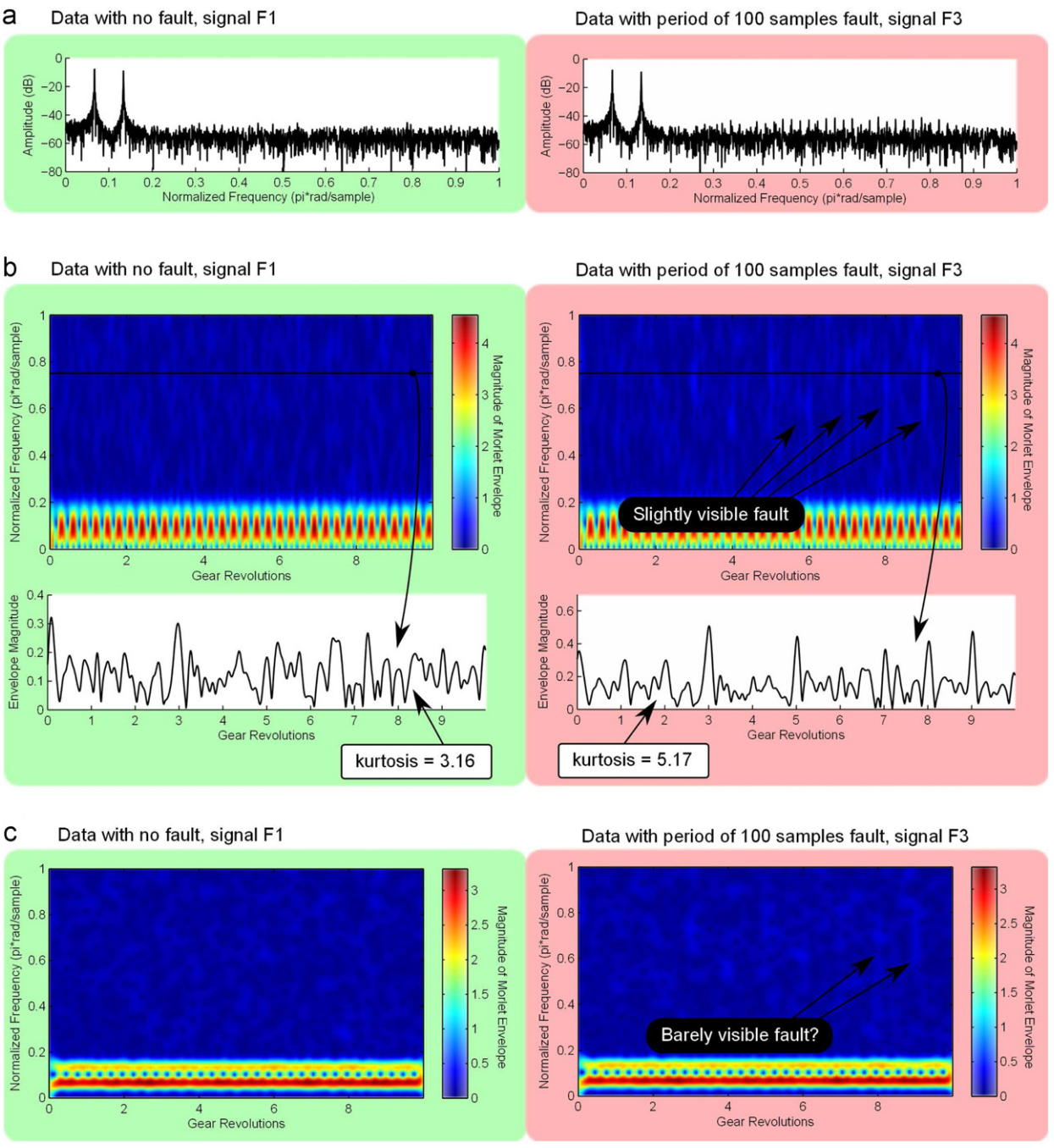

Fig. 6. Simulated signals with no fault (signal *F*1) and fault (signal *F*3) compared using (a) Discrete Fourier Transform spectrum of simulated signals, (b) Complex Shifted Morlet Wavelet Transform with wide bandwidth and envelope at 0.75 pi*rad/sample and (c) Complex Shifted Morlet Wavelet Transform with narrow bandwidth.

6. Experimental results

For validation, vibration data is collected and compared from a gearbox with and without a gear chip. The machine configuration, Fig. 7, is composed of a motor, gearbox, and brake. Two gears are used for the Gear 1 position; one in healthy condition and one with a gear crack, Fig. 8. The machine is operated at 10% load and the rotational frequency of Shaft 1 is varied between 10, 15, 20, 25, 30, 35, and 40 Hz [27]. The accelerometer is of model PCB 352C67 and the data is acquired through a DSP Siglab 20-42 Signal Analyzer to a laptop [28]. The vibration accelerometer sampling frequency is varied according to the rotational speed between 1280 Hz and 5120 Hz, and 8192 samples are acquired for each measurement.

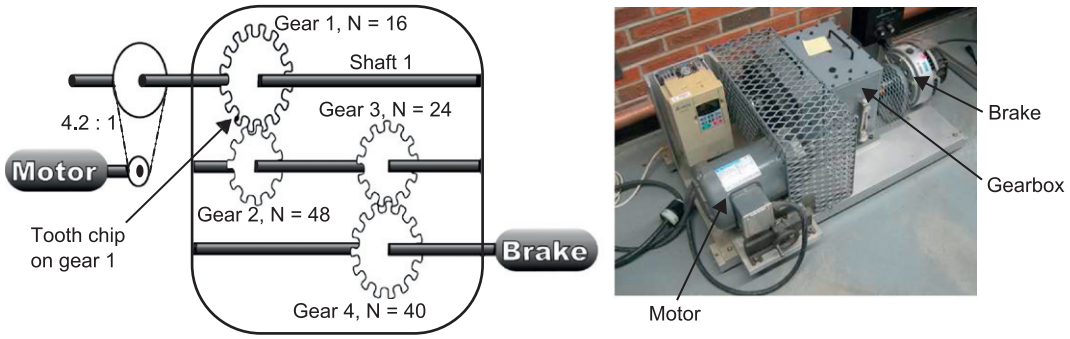

Fig. 7. Gearbox and experimental equipment layout.

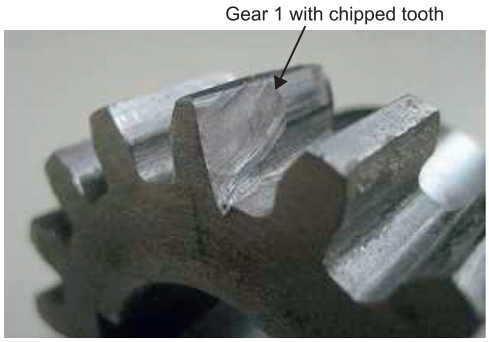

Fig. 8. Faulty gear 1 with seeded tooth chip.

At each frequency, two non-fault vibration measurements and a single fault vibration measurement are acquired and processed.

In the ARMED method proposed by Endo and Randall, the authors suggest that the MED stage should just be performed for only a few iterations to prevent erroneously extracting impulses [9]. Fig. 9 illustrates the results of MCKD, MED, and ARMED versus the iteration in the deconvolution stage for the experimental gear chip data. It can be seen that for both the MED and ARMED approach, the deconvolved signals indicate the fault clearest with a high iteration count, although the methods fail to extract a fault impulse for each revolution of the faulty gear. Reducing the iteration count does not appear to resolve this problem in either case. As a result, for this paper the MED algorithm is executed to convergence instead of being limited to just a few iterations. Fig. 10 illustrates the deconvolution results for the same fault vibration signal versus filter size. By close inspection, it is clear that for the MCKD method a higher filter size results in improved results. With the MED approach the repeating fault is roughly extracted at around a filter size of 50, and larger filter size cleans the signal up but at a cost of missing many of the fault impulses.

The general procedure for the data processing is as follows:

Step 1: Shaft 1 speeds for each data measurement are assumed to be constant and close to the indicated frequency. To calculate the speed, the discrete Fourier transform is applied and peak detection is performed within the vicinity of the expected shaft speed. The period of Shaft 1 in number of samples is referred to as T, and can be a fractional number of samples (such as 172.32 samples per gear revolution). Direct speed or tachometer measurement would result in better performance of the MCDK method.

Step 2: Generate AR models at each operational frequency from the second no-fault data measurement of order $\lfloor 0.4T+0.5 \rfloor$, i.e. 40% the number of data samples corresponding to a single Shaft 1 revolution and rounded to the nearest integer. The Akaike information criterion is not used to select the AR model orders because it erroneously overfits the data by suggesting model orders almost equal to the dataset sizes of around 8000 samples.

Step 3: Calculate AR model prediction residuals for the first no-fault data and the fault data measurements by 1-step ahead prediction.

Step 4: Deconvolution filter sizes for both MED and MCKD are selected as $\lfloor 0.8T+0.5 \rfloor$, i.e. 80% the number of data samples in a single Shaft 1 revolution and rounded to the nearest integer.

Step 5: MED is performed on each AR residual, each first no-fault dataset, and each fault dataset; resulting in the signals for MED and ARMED methods. The Kurtosis of each resulting output signal is calculated as a fault indicator.

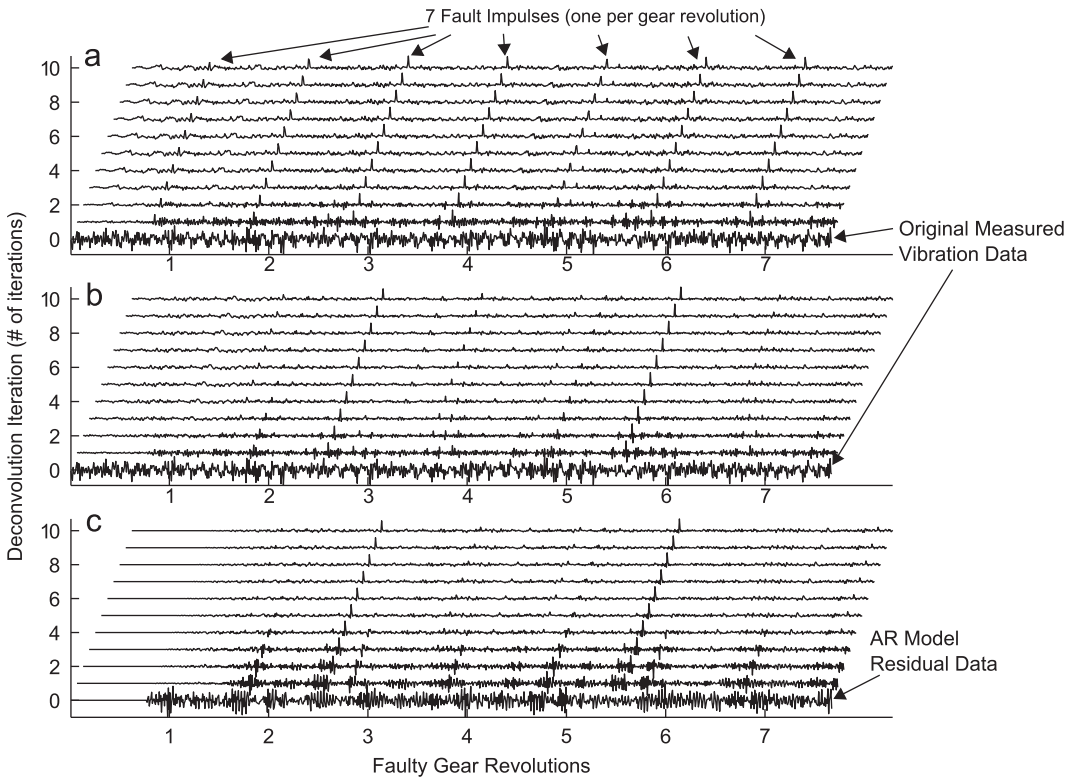

Fig. 9. Deconvolved signals versus deconvolution iteration for an AR model order of 100 and filter sizes of 200 samples. The vibration data is from the gearbox under gear chip fault at 40 Hz shaft speed and 10% load and resulting signals are normalized by peak values with some signals being polarity swapped for better illustration. (a) the $MCKD_1$ Method, (b) the MED Method and (c) the ARMED Method.

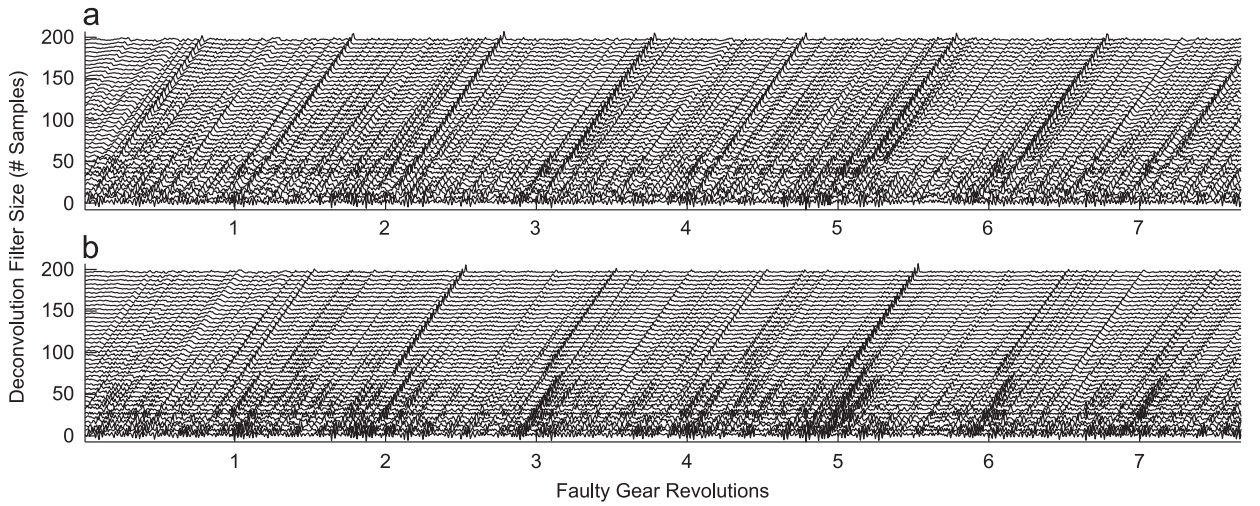

Fig. 10. (a) MCKD and (b) MED methods versus deconvolution filter size for the machine under gear chip fault at 40 Hz shaft speed and 10% load. Each deconvolution method is iterated 50 times and the signals are normalized by peak values polarity swapped for clearer illustration.

Step 6: First, third, and fifth-shift MCKD is applied about period T for each first no-fault dataset, and each fault dataset. $CK_M(T)$ is calculated from the resulting deconvolved signals as a fault indicator where M corresponds to the shift of the MCKD algorithm (e.g. $CK_3(T)$ is used as the fault indicator for third-shift MCKD).

Time synchronous averaging is not considered as a preprocessing step due the significant loss of information through this operation, and the lack of tachometer or detailed speedometer information in this experimental setup.

The deconvolution filter sizes and AR model orders for the fault datasets and first no-fault datasets are tabulated in Table 2. These high-length deconvolution filters are selected because the time-domain plots more clearly illustrate the faults.

The resulting processed signals for no-fault and fault conditions at 35 Hz are compared in Fig. 11. All processed signals under gear chip fault are plotted in Fig. 12 for all frequencies and all methods. From these results it is clear that the MCKD method extracts the expected single impulse per gear revolution, while the MED-based methods tend to miss most of the fault impulses. The detected fault impulses by different methods do not necessarily line up as a result of the different phase lags of the resulting filters. The exact placement of the faults in the time domain can be recovered by simply plotting the FIR filter responses and typically the lag is clear in these plots.

Fault detection based on deconvolution filter sizes of 2 through 150 samples are analysed and compared, Fig. 13. The filter size is chosen as the investigated parameter here because Fig. 9 illustrates that both methods indicate the fault better at a high iteration count, while from Fig. 10 it is unclear at which filter size each method performs best at. From the plots it is clear that MCKD method significantly outperforms the MED method, often performing over 10 times better. The performance gain by increasing the shift from first to third and fifth shift is very significant. In Fig. 13f, the fifth-shift MCKD performs worse than all other methods for a brief filter size range, this could be as a result of poor machine speed estimation or deconvolution convergence problems. Also important to note is that the MED approach performs similarly to the ARMED approach, indicating that the AR stage may not be required for some datasets.

Table 2
Estimated Shaft 1 period and corresponding deconvolution filter sizes for no-fault and fault cases.

Shaft 1 frequency (Hz)	No-fault dataset			Fault dataset		
	Period, T (# Samples)	Deconv. filter size L (# Samples)	AR model order	Period, T (# Samples)	Deconv. filter size L (# Samples)	AR model order
10	260.1	208	104	277.0	222	111
15	173.7	139	69	177.2	142	71
20	259.2	207	104	263.6	211	105
25	207.0	166	83	209.6	168	84
30	172.4	138	69	174.2	139	70
35	147.6	118	59	149.0	119	60
40	129.2	103	52	130.3	104	52

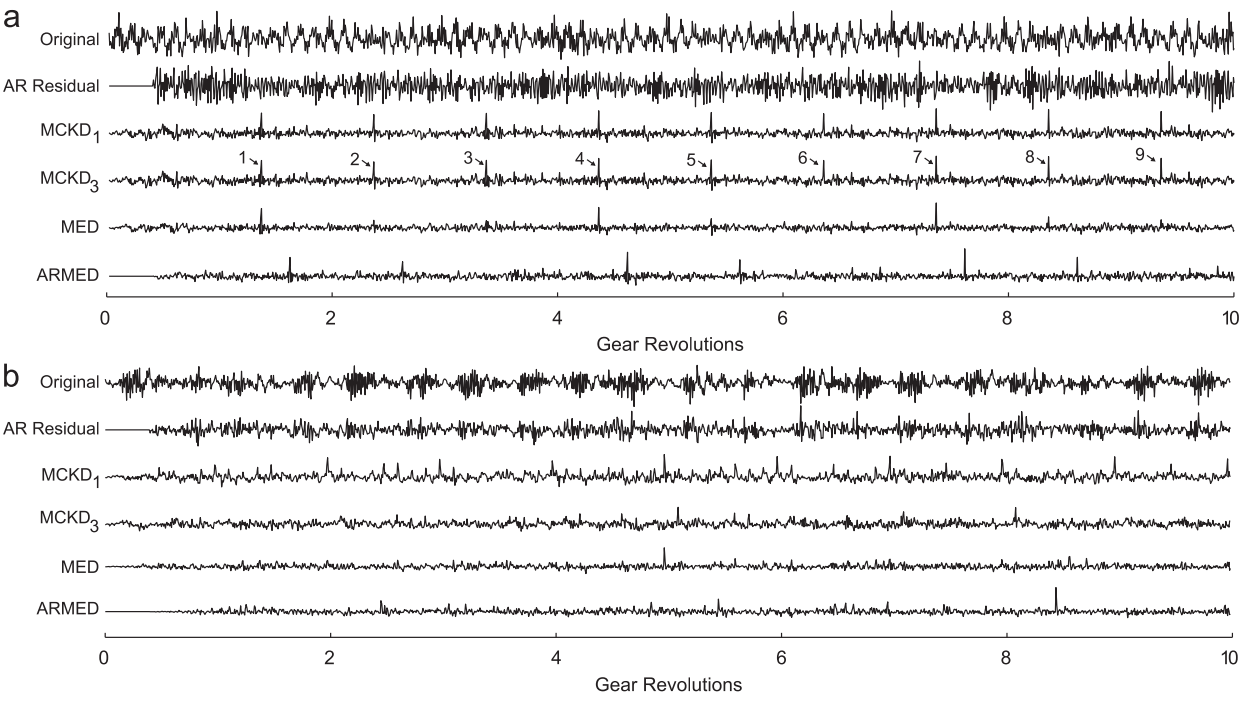

Fig. 11. Experimental acceleration data with fault detection processing with (a) Faulty gear chip and Shaft 1 at a speed of 35 Hz, and (b) Healthy normal gear and Shaft 1 at a speed of 35 Hz. All signals are normalized by peak values and some signals are reversed in polarity for better illustration.

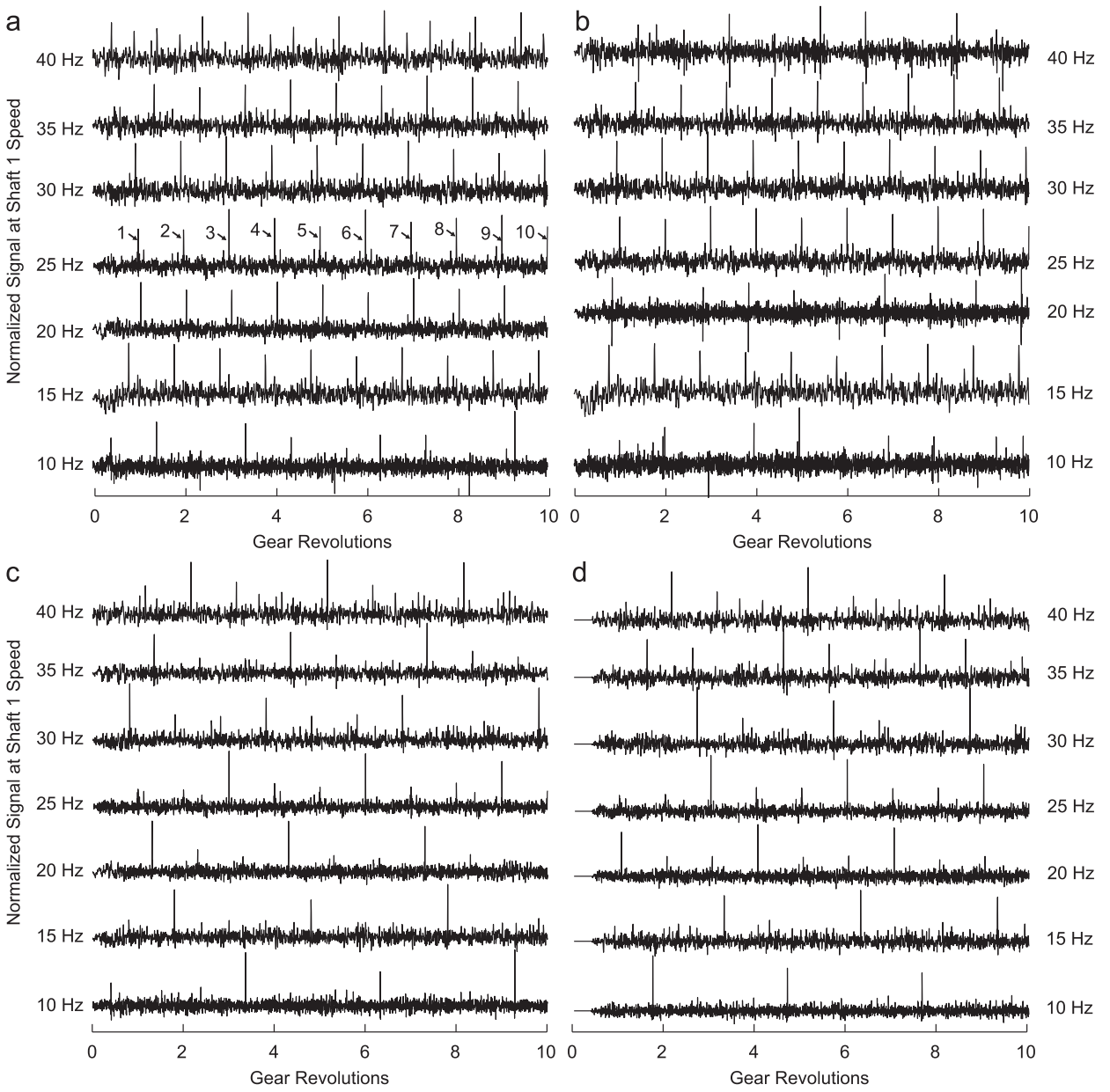

Fig. 12. Experimental acceleration data processed and normalized with a faulty gear chip using (a) the $MCKD_1$ method, (b) the $MCKD_3$ method, (c) the MED method, and (d) the ARMED method. All signals are normalized by peak values and some signals are polarity flipped for better illustration.

7. Online condition monitoring implementation

In industry application, fault detection during machine operation is critical in preventing equipment damage or failure. Presented here is a modification of the MCKD technique to monitor the health of a machine online and generate fault alarms.

Two operations for buffering and downsampling are defined as

$$\text{Data buffering by factor R and no overlap, } B_R : [x_1 \;\; x_2 \;\; x_3 \;\; \cdots] \overset{B_R(\ldots)}{\rightarrow} \begin{bmatrix} x_1 & x_{R+1} & x_{2R+1} & \cdots \\ x_2 & x_{R+2} & x_{2R+2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ x_R & x_{2R} & x_{3R} & \cdots \end{bmatrix}$$

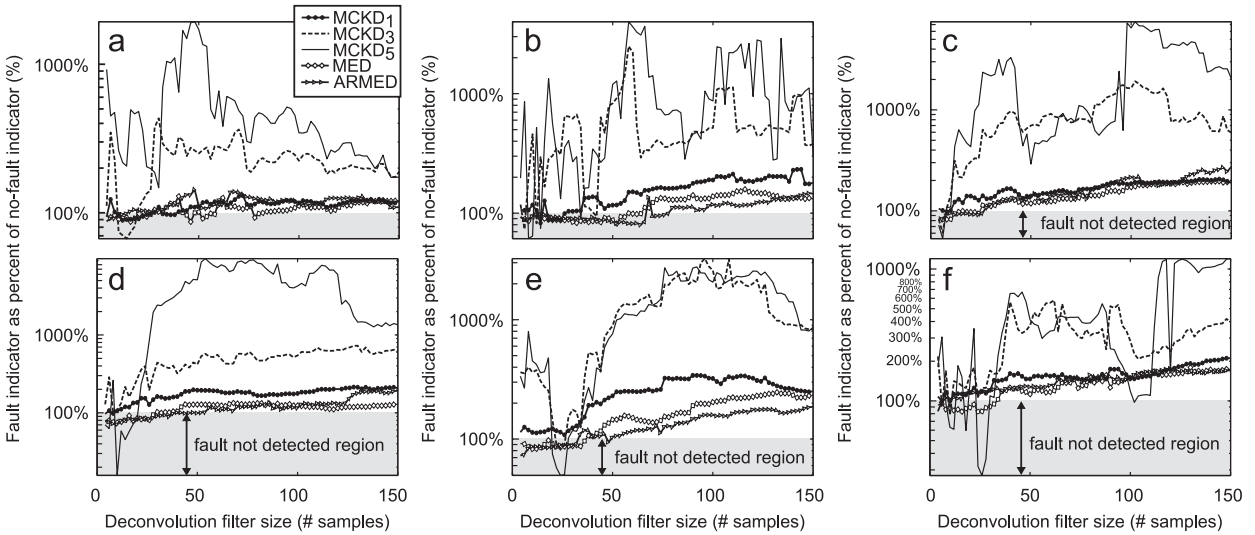

Fig. 13. Experimental fault detection by a method with varied deconvolution filter size for the machine operating at (a) 15 Hz, (b) 20 Hz, (c) 25 Hz, (d) 30 Hz, (e) 35 Hz, and (f) 40 Hz. Note the log-scale on the *y*-axis, meaning the $MCKD_3$ and $MCKD_5$ typically perform significantly better at fault detection.

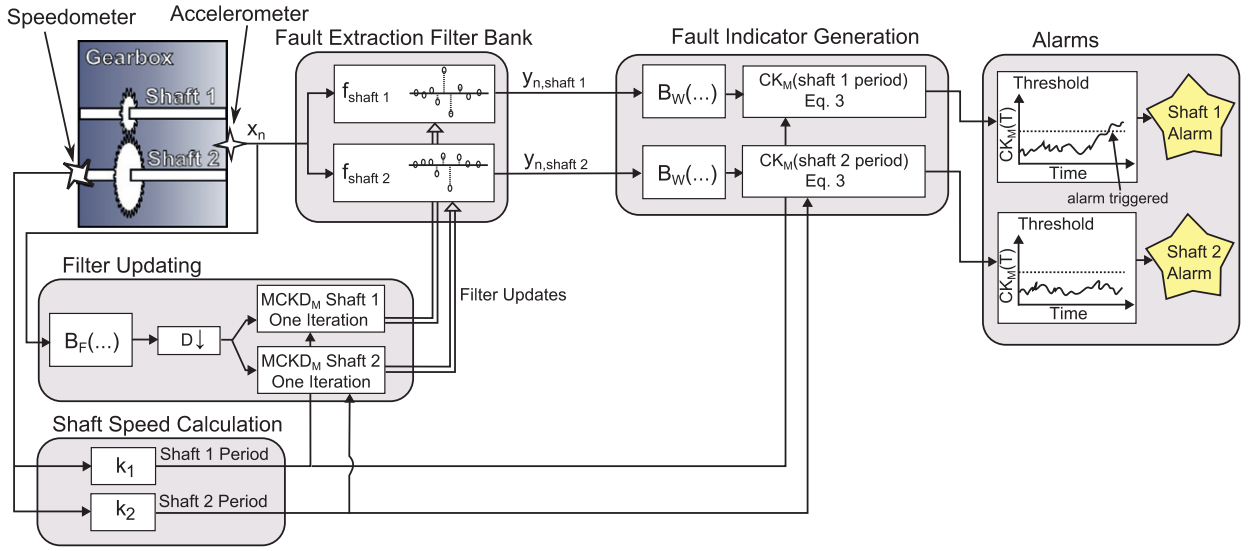

Fig. 14. Online implementation of MCKD-based fault detection for a two-shaft gearbox with a single accelerometer. Threshold and FIR filter plots are creative illustrations, not data.

$$\text{Downsample by factor D, } D\downarrow: \begin{bmatrix} x_{0,0} & x_{0,1} & x_{0,2} & \cdots \\ x_{1,0} & x_{1,1} & x_{1,2} & \cdots \\ x_{2,0} & x_{2,1} & x_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ x_{I,0} & x_{I,1} & x_{I,2} & \cdots \end{bmatrix} \overset{D\downarrow}{\rightarrow} \begin{bmatrix} x_{0,0} & x_{0,D} & x_{0,2D} & \cdots \\ x_{1,0} & x_{1,D} & x_{1,2D} & \cdots \\ x_{2,0} & x_{2,D} & x_{2,2D} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ x_{I,0} & x_{I,D} & x_{I,2D} & \cdots \end{bmatrix}$$

Fig. 14 illustrates the online implementation schematic. The presented implementation is for a single accelerometer near both Shaft 1 and Shaft 2, but performance can be improved by placing two different accelerometers close to each shaft. A speed measurement of one of the shafts is recommended, but if the system is expected to be operating at only a single operating speed then discrete Fourier transform peak detection on the acceleration signal can estimate the exact machine speed. To reduce complexity, the *MCKD* block performs only a single deconvolution filter update iteration for each column of input data and the filter update downsampling factor D controls how often a filter update iteration is performed. The buffer factor F controls the window size of MCKD iteration and a value resulting in around 10 times the period of the

shaft is recommended. The buffer factor W controls the window size for CK fault indicator generation, a larger factor will result in a more slowly changing fault indicator and slower time response to a fault while a smaller factor results in a more noisy fault indicator but quicker time response to a fault. Constant gain values of k_1 and k_2 convert the measured speed to the periods of Shafts 1 and 2 respectively. The alarm thresholds should be selected based on indicator values under no-fault. f_{shaft1} and f_{shaft2} refer to the current FIR deconvolution filters for the respective shafts.

For testing the online implementation, basic input sequences are formed for each operating frequency by looping the experimental no-fault data 10 times followed by looping fault data 10 times. The data is looped by direct concatenation of the datasets reduced to the number of samples corresponding to the largest number of complete shaft revolutions in the dataset. A deconvolution filter size of $L=200$, no decimation factor $D=1$, and filter update and fault indicator buffer sizes of 5000 is used $W=F=5000$.

Table 3 illustrates the number of samples per dataset after looping, the time taken to measure the data, and the processing time for the proposed online fault detection implementation. The processing times are measured on Intel Core 2 Duo CPUs at 2.00 GHz with implementation in MATLAB without parallel processing. From the processing times, it is clear that the proposed method is easily achievable in an online application. Fig. 15 illustrates the online fault detection results for first-shift MCKD. It is clear that the fault detection performance is good in online application, and the proposed threshold alarm on this signal is feasible. The trough in the fault indicator as the fault is introduced is as a result of both the dataset looping not being phase aligned during dataset concatenation between the no-fault and fault data, and the slight machine speed difference between the two datasets. Fig. 16 presents the fifth-shift MCKD results. Clearly the MCKD method is very strong at detecting the faults, with the fifth-shift MCKD method providing clearer results than the first-shift implementation but at a higher computational cost.

Table 3
Computation times for online *MCKD* gear chip fault detection on looped experimental data.

Shaft 1 frequency (Hz)	Number of samples (kSamples)	Duration of data acquisition (s)	$MCKD_1$ data processing time (s)	$MCKD_5$ data processing time (s)
15	326	128	26	44
20	324	63	27	45
25	325	64	24	45
30	325	64	34	53
35	324	63	25	45
40	326	64	35	54

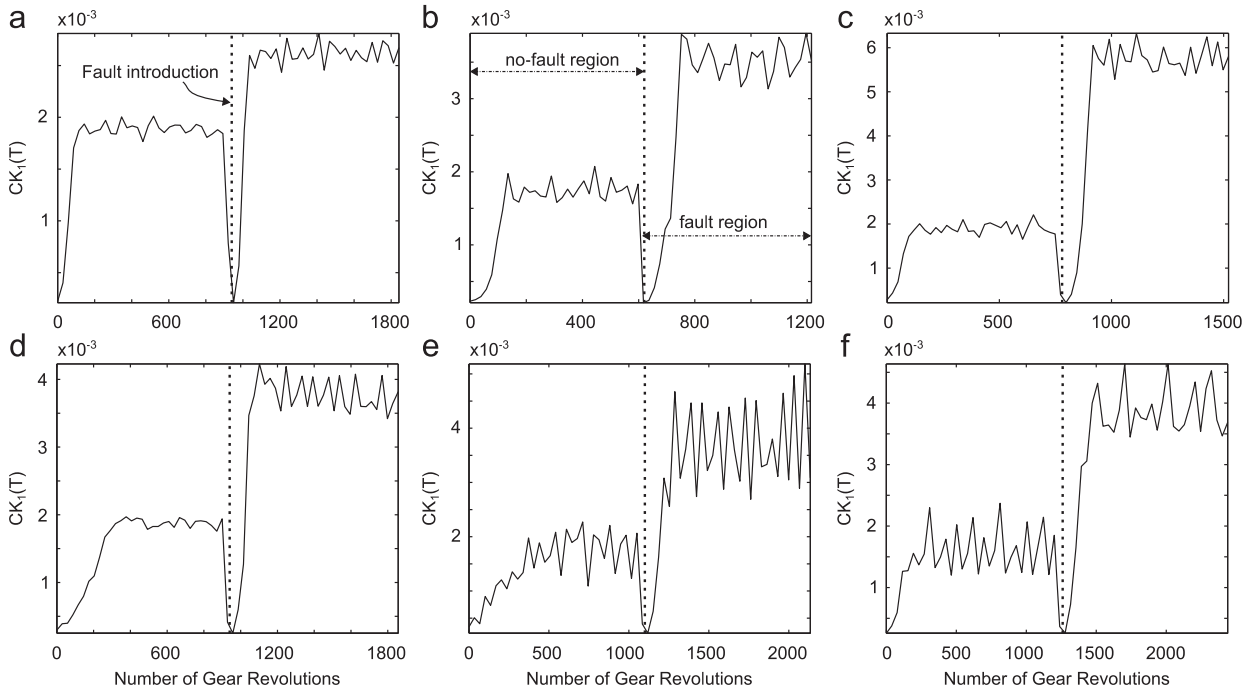

Fig. 15. Online results for first-shift MCKD fault detection for machine under 10% load at (a) 15 Hz, (b) 20 Hz, (c) 25 Hz, (d) 30 Hz, (e) 35 Hz, and (f) 40 Hz. The data is from looped experimental data.

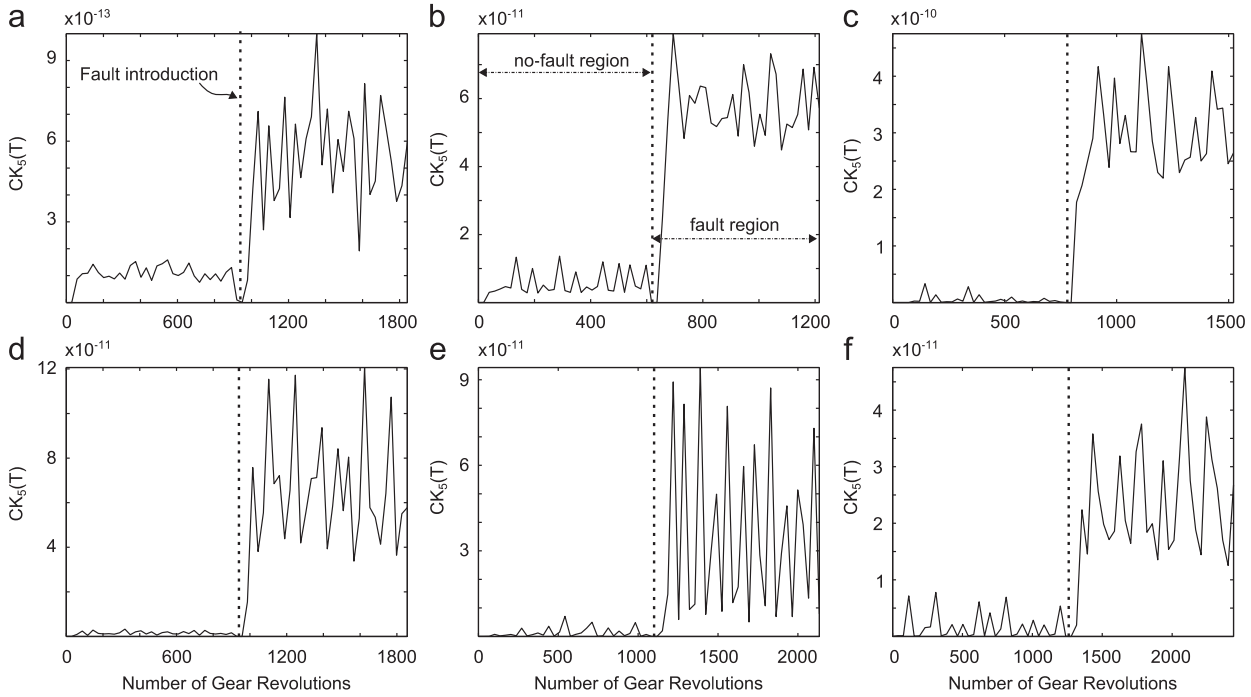

Fig. 16. Online results for fifth-shift MCKD fault detection for machine under 10% load at (a) 15 Hz, (b) 20 Hz, (c) 25 Hz, (d) 30 Hz, (e) 35 Hz, and (f) 40 Hz.

8. Conclusion

This paper introduces a new deconvolution process, MCKD, which aims to deconvolve periodic impulse faults from a machine vibration signal. Simulation data and experimental gear chip vibration data were compared between MCKD and the established methods of AR residual, MED, and ARMED. The experimental results indicate that the AR residual method is inadequate, which is as a result of the fault impulses being of comparable amplitude to the system noise and non-LTI characteristics of the signal. Additionally, the experimental results indicate that the ARMED method performs similarly to direct MED application; indicating that the AR stage of ARMED is unnecessary on some datasets. Finally, the simulation and experimental results indicate that the MCKD method is the most successful in deconvolving the periodic faults and has significantly better gear chip fault detection results, often on the order of 10 or more times better at fault detection on the experimental data. There is a significant performance advantage gain from increasing the shift-order from first to third, but fifth-shift MCKD did not significantly outperform third-shift. As a result the third-shift MCKD method is recommended in application. Clearly the MCKD method is a notable improvement upon the existing state-of-the-art methods and may include applications beyond rotating machine fault detection.

Online application of the MCKD method is shown to be computationally implementable on a personal computer, and results from fault detection on looped experimental data indicate that it is a strong method in online application and a simple threshold alarm is feasible.

Further work should investigate the application of MCKD towards detecting other types of gear faults, bearing faults, and rotor-to-stator rubbing. Concurrent fault detection should be investigated on an experimental setup. The choice of polyphase filter resampling should be investigated further, since it may have a significant effect on the computational cost and algorithm fault detection performance.

Acknowledgement

The authors would like to thank X. Tian, J. Lin, M. Agelinchaab, J. Zhang, M. Akhtar, and K.R. Fyfe for their work on the gearbox with tooth chip experiment. Many thanks to the anonymous reviewers, who provided valuable feedback and recommendations. Thanks to Jian Qu and Tejas Patel from the Reliability Research Lab, Mechanical Engineering University of Alberta, for their help with access to the equipment and navigating the data archives. We also wish to graciously thank NSERC and Syncrude Canada Ltd. for making this possible through funding support.

External resources: Minimum Entropy Deconvolution MATLAB implementation: http://www.mathworks.com/matlabcentral/fileexchange/29151-minimum-entropy-deconvolution-med-1d-and-2d.

M-Shift Maximum Correlated Kurtosis Deconvolution MATLAB implementation: http://www.mathworks.com/matlabcentral/fileexchange/31326.

References

[1] T. Barszcz, R.B. Randall, Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine, Mech. Syst. Signal Process. 23 (2009) 1352–1365.
[2] P.D. Samuel, D.J. Pines, A review of vibration-based techniques for helicopter transmission diagnostics, J. Sound Vib. 282 (2005) 475–508.
[3] F. Harris, E. Kasper, L. Isler, US Civil Rotorcraft Accidents 1963 Through 1997, Technical Report NASA/TM2000-209597, USAAMCOM-TR-00-A-006, NASA, Ames Research Center, 2000.
[4] W. Wang, A.K. Wong, Autoregressive model-based gear fault diagnosis, J. Vib. Acoust. 124 (2002) 172–179.
[5] X. Wang, V. Makis, Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov–Smirnov test, J. Sound Vib. 327 (2009) 413–423.
[6] M. Yang, V. Makis, Arx model-based gearbox fault detection and localization under varying load conditions, J. Sound Vib. 329 (2010) 5209–5221.
[7] G.L. McDonald, Q. Zhao, Model-based adaptive frequency estimator for gear crack fault detection, in: American Control Conference, 2011.
[8] R. Isermann, Model-based fault-detection and diagnosis—status and applications, Annu. Rev. Control 29 (2005) 71–85.
[9] H. Endo, R. Randall, Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter, Mech. Syst. Signal Process. 21 (2007) 906–919.
[10] N. Sawalhi, R. Randall, H. Endo, The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis, Mech. Syst. Signal Process. 21 (2007) 2616–2633.
[11] J. Antoni, R. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, Mech. Syst. Signal Process. 20 (2004) 308–331.
[12] F. Combet, L. Gelman, Optimal filtering of gear signals for early damage detection based on the spectral kurtosis, Mech. Syst. Signal Process. 23 (2009) 652–668.
[13] R. Randall, Gearbox fault diagnosis using cepstrum analysis, in: Proceedings of the 4th World Congress on Theory of Machines and Mechanisms, vol. 1, pp. 169–171.
[14] A.C. McCormick, A.K. Nandi, Cyclostationarity in rotating machine vibrations, Mech. Syst. Signal Process. 12 (1998) 225–242.
[15] M.E. Badaoui, F. Guillet, J. Danière, New applications of the real cepstrum to gear signals, including definition of a robust fault indicator, Mech. Syst. Signal Process. 18 (2004) 1031–1046.
[16] W.J. Staszewski, G.R. Tomlinson, Application of the wavelet transform to fault detection in a spur gear, Mech. Syst. Signal Process. 8 (1994) 289–307.
[17] N.G. Nikolaou, I.A. Antoniadis, Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets, Mech. Syst. Signal Process. 16 (2002) 677–694.
[18] X. Fan, M.J. Zuo, Gearbox fault detection using Hilbert and wavelet packet transform, Mech. Syst. Signal Process. 20 (2006) 966–982.
[19] S. Abbasion, A. Rafsanjani, A. Farshidianfar, N. Irani, Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine, Mech. Syst. Signal Process. 21 (2007) 2933–2945.
[20] Y. Pan, J. Chen, L. Guo, Robust bearing performance degradation assessment method based on improved wavelet packet-support vector data description, Mech. Syst. Signal Process. 23 (2009) 669–681.
[21] C. Capdessus, M. Sidahmed, J.L. Lacoume, Cyclostationary processes: application in gear faults early diagnosis, Mech. Syst. Signal Process. 14 (2000) 371–385.
[22] I. Antoniadis, G. Glossiotis, Cyclostationary analysis of rolling-element bearing vibration signals, J. Sound Vib. 248 (2001) 829–845.
[23] H. Endo, R. Randall, C. Gosselin, Differential diagnosis of spall vs. cracks in the gear tooth fillet region: experimental validation, Mech. Syst. Signal Process. 23 (2009) 636–651.
[24] W. Wang, An evaluation of some emerging techniques for gear fault detection, Struct. Health Monit. 2 (2003) 225–242.
[25] R.A. Wiggins, Minimum entropy deconvolution, Geoexploration 16 (1978) 21–35.
[26] S.L. Marple Jr., Digital Spectral Analysis with Applications, , 1987.
[27] X. Tian, J. Lin, M. Agelinchaab, J. Zhang, M. Akhtar, M. Zuo, K. Fyfe, Technical Report for Experiment 02-05 on Gearbox, Technical Report, Reliability Research Laboratory, Department of Mechanical Engineering, University of Alberta, 2002.
[28] X. Fan, M.J. Zuo, Gearbox fault detection using Hilbert and wavelet packet transform, Mech. Syst. Signal Process. 20 (2006) 966–982.