Multipoint Optimal Minimum Entropy Deconvolution and Convolution Fix: Application to Vibration Fault Detection

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

[a]*Anti-Virus Researcher, Microsoft Malware Protection Center, Microsoft, Vancouver, British Columbia, Canada*
[b]*Advanced Control Systems Laboratory, Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Alberta, Canada*

Abstract

Minimum Entropy Deconvolution (MED) has been applied successfully to rotating machine fault detection from vibration data, however this method has limitations. A convolution adjustment to the MED definition and solution is proposed in this paper to address the discontinuity at the start of the signal - in some cases causing spurious impulses to be erroneously deconvolved. A problem with the MED solution is that it is an iterative selection process, and will not necessarily design an optimal filter for the posed problem. Additionally, the problem goal in MED prefers to deconvolve a single-impulse, while in rotating machine faults we expect one impulse-like vibration source per rotational period of the faulty element. Maximum Correlated Kurtosis Deconvolution was proposed to address some of these problems, and although it solves the target goal of multiple periodic impulses, it is still an iterative non-optimal solution to the posed problem and only solves for a limited set of impulses in a row. Ideally, the problem goal should target an impulse train as the output goal, and should directly solve for the optimal filter in a non-iterative manner. To meet these goals, we propose a non-iterative deconvolution approach called Multipoint Optimal Minimum Entropy Deconvolution Adjusted (MOMEDA). MOMEDA proposes a deconvolution problem with an infinite impulse train as the goal and the optimal filter solution can be solved for directly. From experimental data on a gearbox with and without a gear tooth chip, we show that MOMEDA and its deconvolution spectrums according to the period between the impulses can be used to detect faults and study the health of rotating machine elements effectively.

Keywords: Gearbox, Fault detection, Minimum entropy deconvolution, Vibration, Rotating machine, Deconvolution

1. Introduction

Rotating machines are a common piece of equipment with applications such as power generation turbines, centrifuges, helicopters, washing machines, and more. Fault detection is often focused on their components such as gears, bearings, or their shafts. Monitoring these machines and their components by collecting vibration or acoustic emissions can be used as part of condition-based maintenance planning [1], can detect faults as they start to develop to reduce further damage [2], or help diagnose already-developed faults [3].

---

*Corresponding author Tel. 1-604-679-4211
Email addresses: `glmcdona@gmail.com` (Geoff L. McDonald), `qingzhao@ece.ualberta.ca` (Qing Zhao)

Some of the common fault detection methods include Wavelet Transform-based methods [4, 5, 6], Spectral Kurtosis [7, 8, 9], modeling approaches [10, 11, 12, 13, 14, 15], Cyclostationary analysis [16, 17], and deconvolution-based methods [18, 19, 8, 20]. Often these methods are combined together [19, 8], and are often used in conjunction with machine learning as features to identify faults [3, 21, 22, 23].

In this paper we will be focusing on the deconvolution methods and their application to rotating machine fault detection. Deconvolution approaches are based on defining a measure of a signal, often referred to as a norm, then a FIR filter is designed such that the filtered output vibration reaches a maximum according to the norm. Minimum Entropy Deconvolution (MED) was proposed by R. A. Wiggens in 1978 for seismic recordings as an iterative selection procedure for a filter which aimed to maximize the Kurtosis norm of the filtered output signal [24]. Kurtosis is a feature that is larger for impulse-like signals, and the author successfully applied MED to reconstruct the impulse-like sources from measured seismic recordings. Realizing that many rotating machine fault types are expected to have impulse-like fault sources, in 2007 H. Endo et. al. [19] first demonstrated MED's effectiveness when applied to rotating machine fault detection [19], and the method has been applied to rotating machine faults successfully in multiple studies since then [8, 25, 5, 18, 20]. Despite the successful results with MED, there are several major drawbacks. Firstly, MED is optimizing the norm Kurtosis which prefers a solution of a single impulse. For rotating machines, we instead expect a series of periodic impulses-like features as fault vibration sources. Secondly, MED is an iterative approach that involves iteratively finding a good filter solution. Ideally, we would be able to solve for the solution directly. Lastly, MED selects a 'good' solution and not necessarily the optimal solution to the posed maximization problem.

To address some of the limitations of MED, in 2012 a new deconvolution problem called Maximum Correlated Kurtosis Deconvolution (MCKD) was proposed based on an introduced Correlated Kurtosis (CK) norm [20]. This problem was designed to deconvolve periodic impulses separated by a known period. Although this partially addressed the need for a periodic impulse deconvolution goal, it was still an iterative procedure, selected a 'good' filter solution (not optimal), required priori knowledge of the fault period, and for non-integer fault periods required an additional resampling preprocessing stage. The computational cost of designing the filter was expensive, meaning spectrums across possible fault periods were not practical.

In 1984, C. A. Cabrelli proposed a similar deconvolution problem to MED using a norm called the D-Norm - which we call Optimal Minimum Entropy Deconvolution (OMED) [26]. It was demonstrated that OMED is geometrically similar to MED. But unlike MED the deconvolution problem was shown to have an exact optimal solution for the filter solution without the need for an iterative process. The underlying solution maximally deconvolves a single point impulse in the signal with respect to the rest of the signal. In this paper, we show that although OMED is able to solve for the optimal solution to its problem, this solution performs worse in rotating machine fault detection than MED - likely as a result of its even greater preference to deconvolve only a single impulse solution.

We propose an adjustment to the convolution definition used by the MED and OMED problems to remove a discontinuity between the assumed-zero input signal and the start of the input signal. If not adjusted for, these algorithms tend to erroneously deconvolve a single impulse at this discontinuity. We recommend using these adjusted solutions going forward when processing rotating machine vibration data.

Addressing the limitations of MED, OMED, and MCKD in application to rotating machines, we propose in this paper a new deconvolution method called Multipoint Optimal Minimum Entropy Deconvolution Adjusted (MOMEDA). With MOMEDA, a target vector defines the location and weightings of the impulses to deconvolve - allowing for periodic impulse train deconvolution target goals that are well-suited to the nature of rotating machine faults of a single impulse-like vibration source per rotation. MOMEDA has a non-iterative optimal solution directly for the filter, so no iterating is required for filter selection. Sets of

target vectors can be solved simultaneously, allowing for spectrums of fault condition versus period analyzed to be plotted. Unlike MCKD, MOMEDA works with non-integer fault periods without a resampling stage. Finally, we show using simulated data and experimental data from a gearbox with and without a gear chip fault that MOMEDA can be used to effectively detect the presence of faults in rotating machines. Experimental results are compared for MED adjusted, OMED adjusted, MCKD, and MOMEDA; each with and without auto-regressive model prediction residual preprocessing.

In Section 2 we will provide background on existing deconvolution methods MED, MCKD, and OMED. For MED and OMED we propose the convolution adjustment to fix the discontinuity that can cause spurious impulses to be deconvolved. In Section 3 we present the MOMEDA problem formulation and solution for a single target vector goal. Next in Section 4 we demonstrate how periodic impulse train targets can be proposed and how MOMEDA can be used to solve for a set of targets simultaneously to generate a fault spectrum according to the period of the fault. Experimental results are presented in Section 5 from a gearbox setup with and without a gear tooth chip before presenting our conclusions and future work suggestions in Section 6.

2. Background and Convolution Adjustments

2.1. Minimum Entropy Deconvolution

At the core of MED in rotating machine fault detection is designing a filter that extracts the periodic impulse-like features associated with some faults. Given a sampled vibration signal composed of multiple components:

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

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

where $\vec{x}$ is the measured machine vibration, $\vec{u}$ is unknown input responsible for the system dynamics, $\vec{d}$ is an impulse train modeling a fault, and $\vec{e}$ is white noise. The characteristic responses, $\vec{h}_u, \vec{h}_d, \vec{h}_e$, generally represent the system dynamics, vibration transmission paths, and characteristics. Figure 1 illustrates an example of how these components may present in a rotating machine under fault.

Kurtosis is large for a single impulse and the fault component, $\vec{d}$, is a signal of high Kurtosis when compared to the other signal components. As a result, it is postulated that selecting a finite length filter $\vec{f}$ to maximize the Kurtosis may design a filter that approximately extracts the high Kurtosis source fault impulses with a time shift, while minimizing the low Kurtosis system dynamics and noise components. This Kurtosis maximization problem under assumed zero mean output, $\vec{y}$, is described as follows:

$$\max_{\vec{f}} \text{kurtosis} = \max_{\vec{f}} \frac{\sum_{n=1}^{N} y_n^4}{\left(\sum_{n=1}^{N} y_n^2 \right)^2},$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad \vec{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_L \end{bmatrix},$$
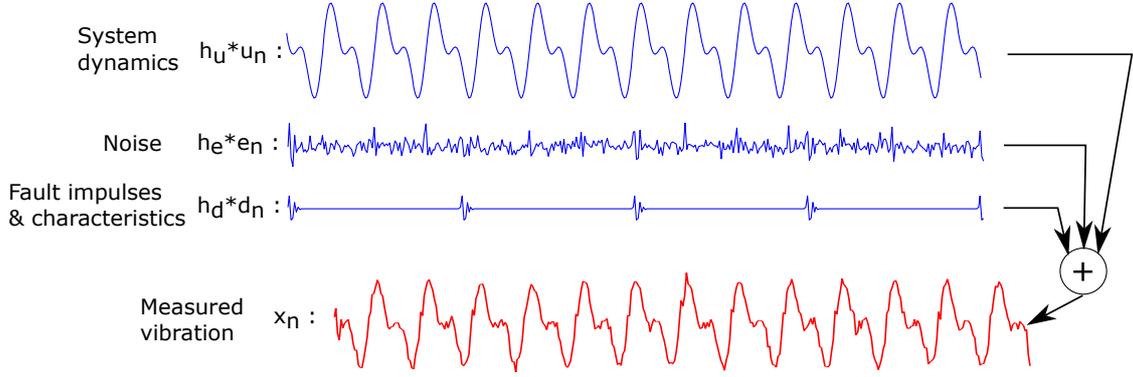
3

Figure 1: Overview of the signal components in a simple rotating machine vibration signal

With convolution definition of output length equal to the input length:

$$\vec{y} = \vec{f} * \vec{x}$$

$$\vec{y}_k = \sum_{l=1}^{L} f_l x_{k-l+1}, \quad k = 1, 2, \ldots, N, \quad x_n = 0 \text{ for } n \neq 1, 2, \ldots, N$$

Or in matrix form:

$$\vec{y} = \bar{X}_0^T \vec{f},$$

$$\bar{X}_0 = \begin{bmatrix} x_1 & x_2 & x_3 & \ldots & \ldots & x_N \\ 0 & x_1 & x_2 & \ldots & \ldots & x_{N-1} \\ 0 & 0 & x_1 & \ldots & \ldots & x_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & \ldots & x_{N-L+1} \end{bmatrix}_{L \text{ by } N}$$

The iterative MED filter selection picks a 'good' solution to this maximization problem, not necessarily the optimal solution. The iterative selection method is derived by taking the derivative, equating it to $\vec{0}$, and iteratively solving for $\vec{f}$. The iterative $\vec{f}$ selection as derived by Wiggens is described as:

$$\vec{f} = \frac{\sum_{n=1}^{N} y_n^2}{\sum_{n=1}^{N} y_n^4} \left(\bar{X}_0 \bar{X}_0^T \right)^{-1} \bar{X}_0 \left[y_1^3 \ y_2^3 \ \ldots \ y_N^3 \right]^T \tag{1}$$

Starting with a centered initial difference filter guess of $\vec{f} = [0, \ldots, 0, 1, -1, 0, \ldots, 0]$, Eq. 1 is repeatedly applied to calculate filter $\vec{f}$, and this new filter is used to calculate the updated output $\vec{y}$ before each iteration. Termination is typically defined as either a number of iterations [20], or a minimum change in filter coefficients between iterations [19]. Kurtosis of the resulting output is often used as a measure of fault. An implementation in MATLAB is available in the External Resources Section.

Several problems with MED exist in application to rotating machine fault detection. If a large filter length L is chosen, MED can design a filter to approximately extract a single impulse even from a white noise signal, often referred to as a spurious impulse. See Figure 2 illustrating this issue when deconvolving a single impulse from 1000 samples of Gaussian white noise and a 100 sample filter length L. Mitigations

a) Gaussian white noise

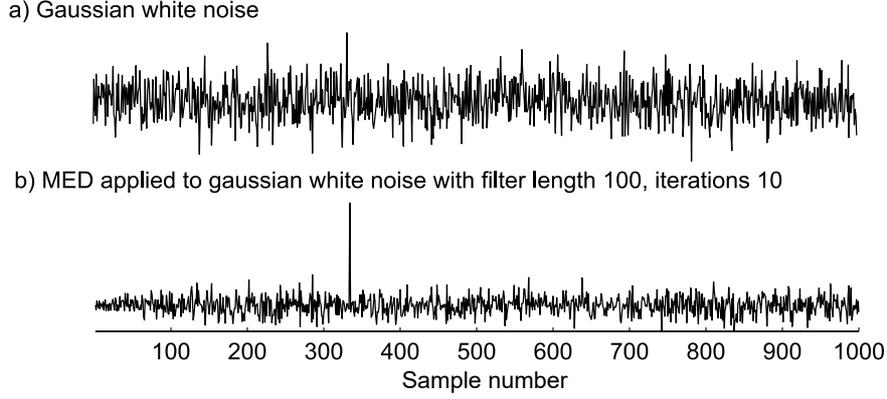b) MED applied to gaussian white noise with filter length 100, iterations 10

Sample number

Figure 2: Applying MED to Gaussian white noise yields a single deconvolved impulse. a) Zero-mean Gaussian white noise, and corresponding b) MED output y signal with filter size $L = 100$ and 10 iterations.

include selecting smaller filter lengths L or terminating the iterative selection early before this solution can be reached.

Another problem is that the solution to MED is iterative and may not correspond to an optimal solution. Fault indication performance may vary depending on the termination condition. In some cases, the resulting signal more closely extracts the periodic fault signal at an earlier termination condition [20].

MED is posing a deconvolution proposition that is not well-posed for rotating machine faults. While MED is prefers to deconvolve a single impulse (maximum Kurtosis), in rotating machine faults we are looking to deconvolve an impulses train with one impulse per rotation of the faulty element. In the next section we will briefly review our previous work to adjust the deconvolution problem to be better suited for rotating machine fault detection.

2.2. Maximum Correlated Kurtosis Deconvolution

In 2012, MCKD was proposed [20] to iteratively deconvolve a series of impulses using the proposed norm CK:

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

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

where M is the number of sequential impulses that are to be deconvolved and T is the period of separation for these impulses. The deconvolution is posed similarly to MED as:

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

It has an iterative selection procedure to select a good $\vec{f}$ aiming to maximize this problem. Although MCKD was found to improve deconvolution results in simulated and experimental data by deconvolving a series of periodic impulses [20], it has many limitations including still being an iterative algorithm, not

5

solving for the optimal solution to the posed problem, and only being able to deconvolve a small series of impulses in a row as opposed to an infinite train of impulses.

2.3. Minimum Entropy Deconvolution with Convolution Adjustment

Application of MED to rotating machine data should use a different convolution definition that reduces the tendency to deconvolve a single impulse at the start of the output signal $\vec{y}$. With the definition of MED, the convolution definition assumes zero data for $x_n = 0, n < 1$, which creates a discontinuity between assumed zero sample x_0 and the first sample x_1. In the original proposed application of MED to seismic recordings [24], the data at the start of $\vec{x}$ were generally close to 0 and this definition was not a significant issue. However, in application to rotating machine vibration this can cause a significant disturbance to be identified between samples observed at x_0 and x_1 - causing a spurious impulse to be deconvolved at this location or within L samples of it due to a delay.

H. Endo et. al. [19] proposed applying an AR model residual preprocessing before applying MED to the vibration signal (AR-MED), and this added step partially mitigates this discontinuity. Instead, for rotating machines the MED convolution definition should be adjusted to only consider the output range without use of any zero-assumed input data:

$$\vec{y} = \vec{f} * \vec{x}$$

$$y_k = \sum_{l=1}^{L} f_l x_{k+L-l}, \quad k = 1, 2, \ldots, N - L + 1$$

Or in matrix form:

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

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

Resulting in the MED adjusted (MEDA) iterative selection:

$$\vec{f} = \frac{\sum_{n=1}^{N-L} y_n^2}{\sum_{n=1}^{N-L} y_n^4} \left(X_0 X_0^T \right)^{-1} X_0 \begin{bmatrix} y_1^3 & y_2^3 & \ldots & y_{N-L}^3 \end{bmatrix}^T$$

This is solved iteratively similarly to MED. First select $\vec{f} = [0, \ldots, 0, 1, -1, 0, \ldots, 0]$, then iteratively solve the above equation recalculating $\vec{y}$ using the new filter for each iteration. An implementation in MATLAB is available in the External Resources Section. Figure 3 demonstrates how MED deconvolves a spurious impulse at the discontinuity, how AR preprocessing mitigates the spurious impulse, and how MEDA achieves a similar result as AR-MED.

2.4. Optimal Minimum Entropy Deconvolution

In 1984, Carlos A. Cabrelli [26] proposed a new norm towards deconvolving impulses called the D-Norm, and geometrically demonstrated the deconvolution problem's similarity to the MED problem. The proposed D-Norm deconvolution problem has an exact non-iterative solution to solve for the filter coefficients. We refer to this method as Optimal Minimum Entropy Deconvolution (OMED) for its similarity
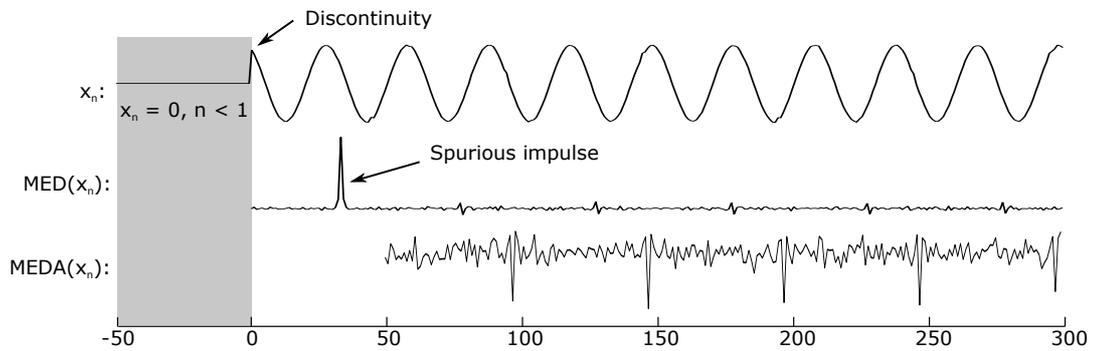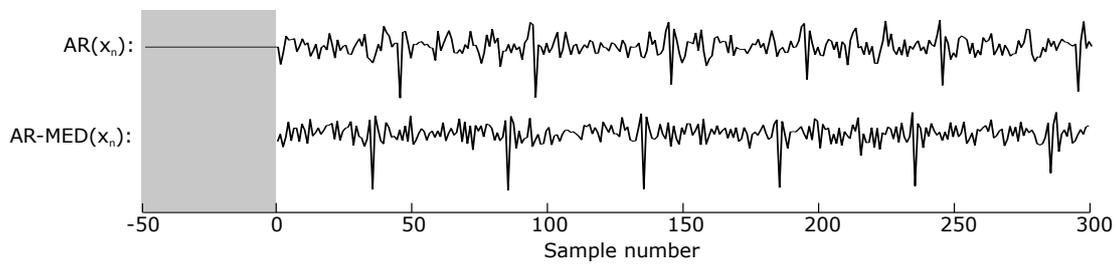
Figure 3: a) MED, MEDA, and b) 10th order AR-MED applied to convergence with filter length $L = 50$ and $N = 1000$ input samples of simulated data. The first 10 samples were discarded prior to MED and MEDA processing to align the AR model residual signal with x_n for better illustration.

7

to the Minimum Entropy Deconvolution problem, and optimal for the problem's ability to solve for the optimal solution.

The D-Norm maximization problem is formulated as the following:

$$\text{D-Norm} = D(\vec{y}) = \max_{k=1,2,\ldots,N} \frac{|y_k|}{||\vec{y}||}.$$

$$\text{OMED}: \quad \sup_{\vec{f}} D(\vec{y}) = \sup_{\vec{f}} \left(\max_{k=1,2,\ldots,N} \frac{|y_k|}{||\vec{y}||} \right),$$

It is sufficient to first solve for $\vec{f}$ over each k, then select the $\vec{f}$ corresponding to the maximum D-Norm:

$$\sup_{\vec{f}} D(y) = \max_{k=1,2,\ldots,N-L} \left(\sup_{\vec{f}} \frac{|y_k|}{||y||} \right).$$

To find the maxima and minimas, this is differentiated with respect to the filter $\vec{f}$ and solved to $\vec{0}$:

$$\frac{d}{d\vec{f}} \left(\frac{y_k}{||\vec{y}||} \right) = \vec{0} \tag{2}$$

Instead of following Cabrelli's OMED convolution definition with $x_n = 0, n < 1$ assumptions, we follow a similar procedure using the modified convolution definition forming Optimal Minimum Entropy Deconvolution Adjusted (OMEDA). This convolution adjustment is particularly important for OMED, otherwise it has a tendency to deconvolve the discontinuity - which is not a real vibration feature. From the convolution definition:

$$y_k = \sum_{l=1}^{L} f_l x_{k+L-l}, \quad k = 1, 2, \ldots, N - L,$$

it follows that the derivative is

$$\frac{dy_k}{df_l} = x_{k+L-l}.$$

Since the following is true,

$$\frac{d||\vec{y}||}{df_l} = ||\vec{y}||^{-1} \sum_{k=1}^{N-L} y_k x_{k+L-1},$$

it follows that Eq. (2) expands as

$$\frac{d}{df_l} \left(\frac{y_k}{||\vec{y}||} \right) = \frac{x_{k+L-l}||\vec{y}|| - y_k ||\vec{y}||^{-1} \sum_{k=1}^{N-L} y_k x_{k+L-1}}{||\vec{y}||^2}$$

$$= x_{k+L-l}||\vec{y}||^{-1} - y_k ||\vec{y}||^{-3} \begin{bmatrix} x_L \\ x_{L+1} \\ \vdots \\ x_N \end{bmatrix}^T \vec{y}, \tag{3}$$

and converting Eq. (3) to matrix form for $l = 1, 2, \ldots, L$ and solving to $\vec{0}$ we have:

8

$$\frac{d}{d\vec{f}}\left(\frac{y_k}{\|\vec{y}\|}\right) = \|\vec{y}\|^{-1}\vec{M}_k - \|\vec{y}\|^{-3}y_k X_0\vec{y} = \vec{0}, \tag{4}$$

$$\vec{M}_k = \begin{bmatrix} x_{k+L-1} \\ x_{k+L-2} \\ \vdots \\ x_k \end{bmatrix}$$

Since $\vec{y} = X_0^T\vec{f}$, and by rearranging we have:

$$\frac{y_k}{\|\vec{y}\|^2}X_0 X_0^T\vec{f} = \vec{M}_k$$

$(X_0 X_0^T)$ is the unnormalized Toeplitz autocorrelation matrix of $\vec{x}$ with no assumed-zero data. Assuming the inverse exists, we have:

$$\frac{y_k}{\|\vec{y}\|^2}\vec{f} = (X_0 X_0^T)^{-1}\vec{M}_k \tag{5}$$

We observe that if $\vec{f}$ is a solution to Eq. (5), then any multiple is also a solution, $\vec{f} = c\bar{f}$:

$$\vec{y} = cX_0\bar{f} = c\bar{y}$$

$$\frac{c\bar{y}_k}{c^2\|\bar{y}\|^2}c\bar{f} = (X_0 X_0^T)^{-1}\vec{M}_k$$

Therefore multiples of $(X_0 X_0^T)^{-1}\vec{M}_k$ are non-trivial solutions for $\vec{f}$. We pick the solution:

$$\vec{f} = (X_0 X_0^T)^{-1}\vec{M}_k \tag{6}$$

Finally, we need to solve for all the $\vec{f}$ solutions across $k = 1, 2, \ldots, N - L$. Expanding Eq. (6) for $k = 1, 2, \ldots, N - L$ we have an array of possible $\vec{f}$ solutions, $F = \left[\vec{F}_1, \vec{F}_2, \ldots, \vec{F}_{N-L}\right]$, which simplifies to:

$$F = (X_0 X_0^T)^{-1}X_0$$

Resulting in the array of possible outputs $Y = \left[\vec{Y}_1, \vec{Y}_2, \ldots, \vec{Y}_{N-L}\right]$:

$$Y = X_0^T(X_0 X_0^T)^{-1}X_0$$

Then the OMEDA solution for $\vec{f}$ is the column in F corresponding to the Y column with maximum D-Norm. An implementation of OMEDA in MATLAB is available in the External Resources Section.

Each potential output solution $\vec{Y}_k$ represents deconvolving only a single impulse at output sample y_k maximally. The end result of selecting the output maximizing the D-Norm over $\vec{Y}_k$ is selecting the filter that best extracts any single-point impulse while minimizing the amplitude of the rest of output. This solution, although it is the optimal solution to a similar problem to the MEDA maximization problem, tends to particularly prefer to deconvolve only single impulses over the expected periodic fault impulses. Consequently, OMEDA often is not as useful as MEDA. Comparing MEDA to OMEDA, Figure 4, we can see OMEDA is able to successfully deconvolve the goal of a single impulse using a smaller filter size than MEDA.

Although OMEDA generally under performs in comparison to MEDA in rotating machine fault extraction, it is an important building block towards the goal of solving for an optimal $\vec{f}$ that is solved for non-iteratively to extract periodic fault impulses.
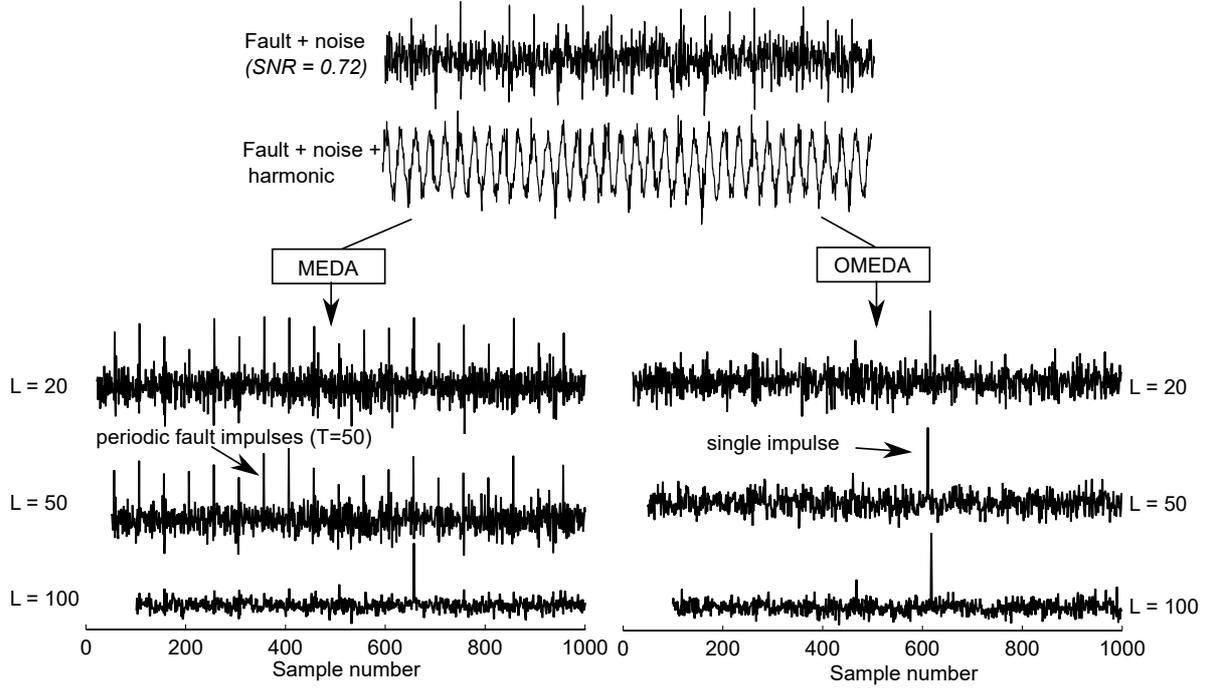
Figure 4: MEDA and OMEDA applied to a simulated fault signal. MEDA is iterated to convergence, while OMEDA solves for the optimal solution.

3. Multipoint Optimal Minimum Entropy Deconvolution

For rotating machine fault detection, we propose a deconvolution target of multiple impulses at known locations - as opposed to a target of a single impulse. We introduce this maximization problem as Multipoint Optimal Minimum Entropy Deconvolution Adjusted (MOMEDA):

$$\text{Multi D-Norm} = MDN(\vec{y}, \vec{t}) = \frac{1}{\|\vec{t}\| \, \|\vec{y}\|} \vec{t}^T \vec{y}$$

$$MOMEDA : \max_{\vec{f}} MDN(\vec{y}, \vec{t}) = \max_{\vec{f}} \frac{\vec{t}^T \vec{y}}{\|\vec{y}\|} \tag{7}$$

Where the target vector, $\vec{t}$, is a constant vector that defines the location and weightings of the goal impulses to be deconvolved. For example,

$$\vec{t} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T,$$

the above $\vec{t}$ will aim to deconvolve two impulses in the output signal: one impulse at $n = 4$ and the other at $n = 8$. With this level of control, the target vector $\vec{t}$ can be used to control the separation, location, and windows for the impulses to be deconvolved. The absolute value is discarded to require the deconvolved impulses to match in polarity - which is expected in rotating machine faults. This is a simpler problem than OMEDA, since no iteration through an array of solutions over $k = 1, 2, \ldots, N - L$ is required.

This Multi D-Norm is normalized to between 0 and 1, where a value of 1 indicates that the optimal target solution was reached. In Appendix A the derivation of this normalization factor is provided. This

10

normalization is important such that fault indicators from different fault periods or signals with different sampling rates can be comparable to one another in fault indication level.

Following a similar derivation as OMEDA, we solve the extremas of Eq. (7) by taking the derivative with respect the filter coefficients, $\vec{f} = f_1, f_2, \ldots, f_L$:

$$\frac{d}{d\vec{f}}\left(\frac{\vec{t}^T \vec{y}}{\|\vec{y}\|}\right) = \frac{d}{d\vec{f}}\frac{t_1 y_1}{\|\vec{y}\|} + \frac{d}{d\vec{f}}\frac{t_2 y_2}{\|\vec{y}\|} + \ldots + \frac{d}{d\vec{f}}\frac{t_{N-L} y_{N-L}}{\|\vec{y}\|} \tag{8}$$

We already solved the derivative for each of these terms as part of the OMEDA derivation, Eq. (4), so we know:

$$\frac{d}{d\vec{f}}\frac{t_k y_k}{\|\vec{y}\|} = \|\vec{y}\|^{-1} t_k \vec{M}_k - \|\vec{y}\|^{-3} t_k y_k X_0 \vec{y}$$

$$\vec{M}_k = \begin{bmatrix} x_{k+L-1} \\ x_{k+L-2} \\ \vdots \\ x_k \end{bmatrix}$$

Therefore Eq. (8) can be written:

$$\frac{d}{d\vec{f}}\left(\frac{\vec{t}^T \vec{y}}{\|\vec{y}\|}\right) = \|\vec{y}\|^{-1}\left(t_1 \vec{M}_1 + t_2 \vec{M}_2 + \ldots + t_{N-L} \vec{M}_{N-L}\right) - \|\vec{y}\|^{-3} \vec{t}^T \vec{y} X_0 \vec{y} \tag{9}$$

With the simplification,

$$t_1 \vec{M}_1 + t_2 \vec{M}_2 + \ldots + t_{N-L} \vec{M}_{N-L} = X_0 \vec{t},$$

and solving for extremas by equating to $\vec{0}$, Eq. (9) becomes:

$$\|\vec{y}\|^{-1} X_0 \vec{t} - \|\vec{y}\|^{-3} \vec{t}^T \vec{y} X_0 \vec{y} = \vec{0}$$

$$\frac{\vec{t}^T \vec{y}}{\|\vec{y}\|^2} X_0 \vec{y} = X_0 \vec{t}$$

Since $\vec{y} = X_0^T \vec{f}$ and assuming $\left(X_0 X_0^T\right)^{-1}$ exists:

$$\frac{\vec{t}^T \vec{y}}{\|\vec{y}\|^2} \vec{f} = \left(X_0 X_0^T\right)^{-1} X_0 \vec{t} \tag{10}$$

Since multiples of $\vec{f}$ are also solutions to Eq. (10), multiples of $\vec{f} = \left(X_0 X_0^T\right)^{-1} X_0 \vec{t}$ are solutions to the MOMEDA problem. In summary, the MOMEDA filter and output solutions are calculated simply as:

$$\vec{f} = \left(X_0 X_0^T\right)^{-1} X_0 \vec{t},$$

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

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

Where the target vector $\vec{t}$ is the same length as the output $(N - L + 1)$, and represents the location and weightings at which to deconvolve impulses in the output. This solution can be described as simply the inverse of the unnormalized Toeplitz autocorrelation matrix of $\vec{x}$ with no-assumed zero input, $\left(X_0 X_0^T\right)^{-1}$, multiplied by the filtered signal $\vec{x} * \vec{t}$ to yield the optimal filter $\vec{f}$ solution for MOMEDA. An implementation of MOMEDA in MATLAB is available in the External Resources Section.

Controlling the locations of the multiple impulses to be deconvolved along with MOMEDA's non-iterative optimal solution provide the basis for a machine diagnosis tool. In the next section we will present how the target vector selection can be used to analyze and generate spectrums to identify faults in rotating machines.

4. Target Propositioning and Fault Indication

For rotation machine fault detection with MOMEDA, solution targets of impulse trains separated by a proposed fault period should be considered:

$$t_n = P_n(T) = \delta_{\text{round}(T)} + \delta_{\text{round}(2T)} + \delta_{\text{round}(3T)} + \dots,$$

$$\vec{t} = \vec{P}(T),$$

where δ_n denotes an impulse at sample n. Non-integer T should be considered and for each impulse position rounded to nearest position, since in reality the fault period is unlikely to coincide exactly with a multiple of the sampling period. If $L \geq T$ the time-domain location of the fault impulse train $\vec{t}$ does not need to be swept across all position shifts, since the designed filter can adjust in time delay to the necessary position. Usage of filter lengths less than the fault period, $L < T$, is not discussed in this paper, but can be implemented by shifting the target vector at multiple positions and selecting the best match.

Figure 5 plots the result for MEDA versus MOMEDA for a simulated fault signal, where in this case we have increased the noise to raise the deconvolution difficulty. Due to the larger filters sizes and higher noise level, MEDA is only able to extract a single impulse in each case, while MOMEDA deconvolves the periodic fault impulses. Note that in the case of $L = 500 = N/2$ the filter length is equal to the length of the output signal $\vec{y}$ resulting in MEDA and MOMEDA being able to exactly deconvolve their optimal goals.

MOMEDA can be calculated for a spectrum of M target vector candidates $\vec{t_1}, \vec{t_2}, ..., \vec{t_M}$ by:

$$F = \begin{bmatrix} \vec{f_1} & \vec{f_2} & \dots & \vec{f_M} \end{bmatrix} = \left(X_0 X_0^T\right)^{-1} X_0 \begin{bmatrix} \vec{t_1} & \vec{t_2} & \dots & \vec{t_M} \end{bmatrix}$$

$$Y = \begin{bmatrix} \vec{y_1} & \vec{y_2} & \dots & \vec{y_M} \end{bmatrix} = X_0^T F$$

We introduce Multipoint Kurtosis (MKurt) to use as a measure of fault:

$$\text{Multipoint Kurtosis} = \frac{\left(\sum_{n=1}^{N-L} t_n^2\right)^2}{\sum_{n=1}^{N-L} t_n^8} \frac{\sum_{n=1}^{N-L} (t_n y_n)^4}{\left(\sum_{n=1}^{N-L} y_n^2\right)^2}$$

This definition is based on Kurtosis, however it has been expanded to multiple impulses at controlled locations according to the target vector and is normalized such that it reaches a value of 1 if a solution in the direction of $\vec{t}$ is reached. See Appendix B for the derivation of the normalization factor.
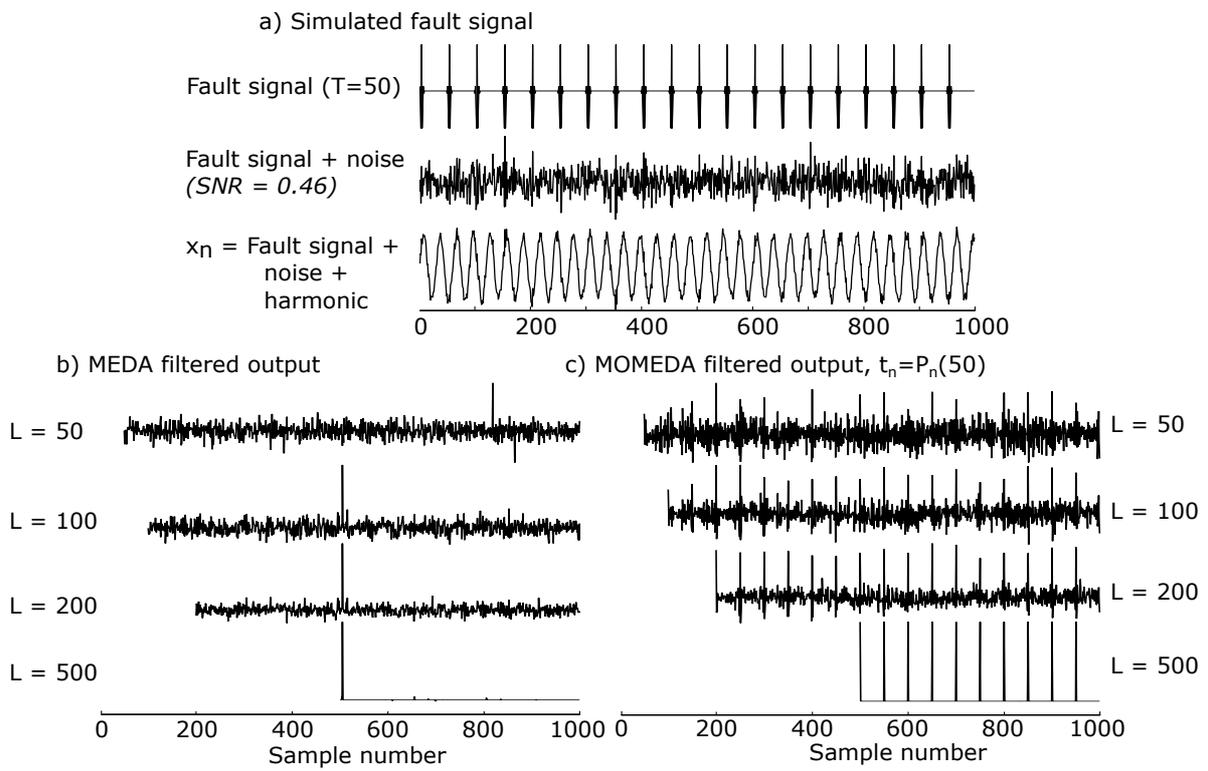
Figure 5: a) A simple simulated signal with a fault at $T_{fault} = 50$ under strong noise, b) MEDA-filtered output after iterating to convergence, and c) MOMEDA-filtered output.
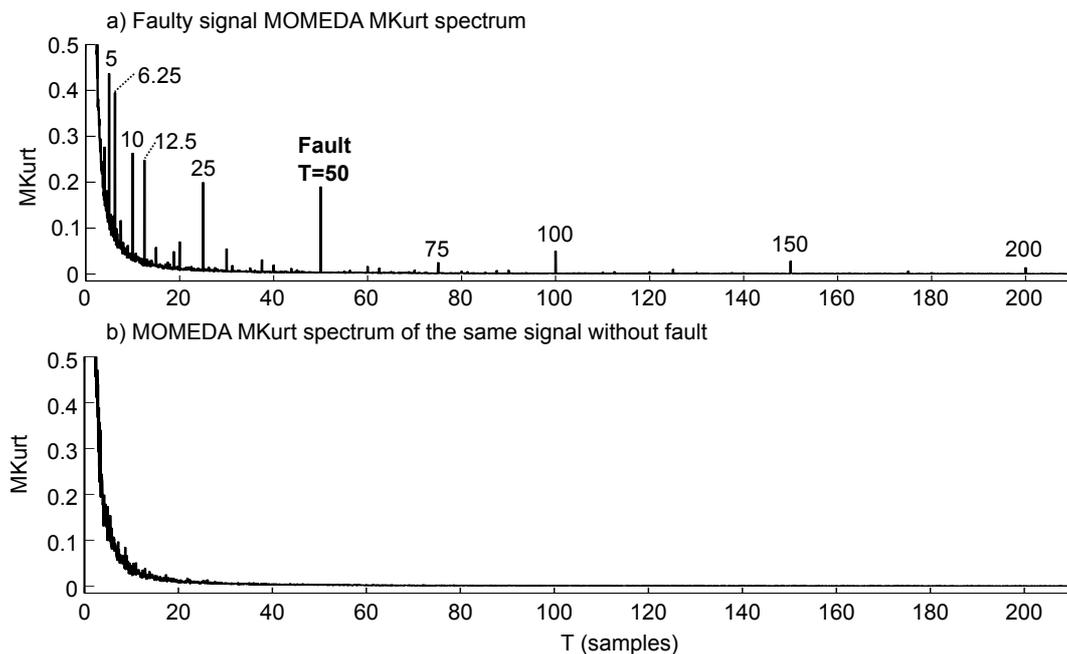
Figure 6: MOMEDA spectrum with a step of $\Delta T = 0.05$ and a filter length of $L = 200$, applied to a simulated fault signal seen in Figure 5 of period $T = 50$ and length $N = 5000$ for a) a faulty signal under strong noise, and b) the same signal without the fault.

Using a spectrum of periodic impulse train target vectors that step over a range of periods, and plotting the MKurt result produces a spectrum that can be used to identify faults. Figure 6 plots the spectrum from $T = 2$ to $T = 220$ with a step of $\Delta T = 0.05$ for the same simulated fault signal under heavy noise. Plotting the spectrum allows for clear identification of peaks corresponding to the fault period of 50, along with its multiples, factors and factors of the multiples (eg 75 being a factor of $3T = 150$). From this spectrum, it is clear that MOMEDA is able to differentiate between the faulty period and its related factors from the surrounding non-fault periods.

When working with experimental data, it was found that further extending the target vector, $\vec{t}$, by introducing a windowing may improve the spectrum clarity. By adding a windowing function, a tolerance for bearing slippage or slight machine speed variations is allowed for and larger period steps, ΔT, can be used. For example,

$$\vec{t} = [1, 1, 1, 1, 1] * \vec{P}(T)$$

The above $\vec{t}$ introduces a rectangular window function. This window function is not studied in this paper, and an empty window of [1] is used for all processing. The next section presents results from an experimental setup with a seeded fault.

5. Experimental Results

Experimental results are compared for a gearbox with and without a seeded gear tooth chip from an experiment performed by X. Tian et. al. [27]. This data has been used previously in other machine fault detection studies [4, 20].

The experimental setup, Figure 7, includes a motor, gearbox, and loading break. Gear 1 is switched between a healthy gear and a faulty gear with a seeded tooth chip for comparison. The rotational frequency
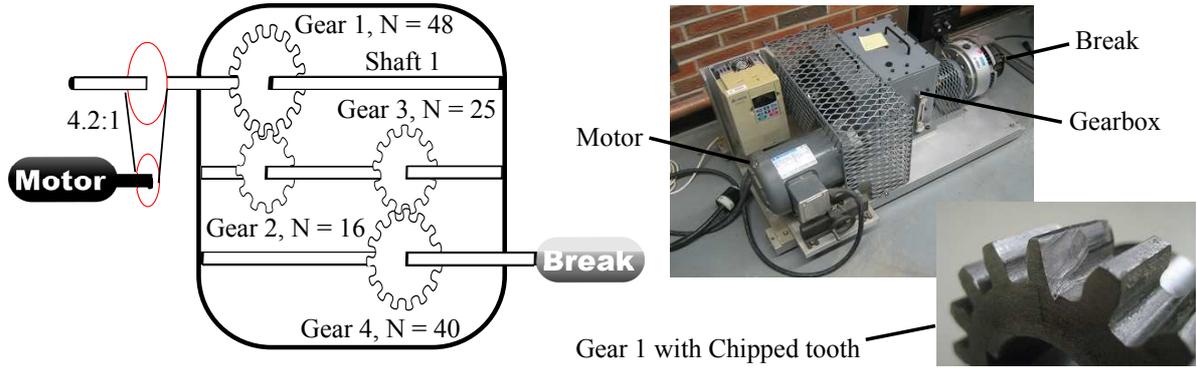
14

Figure 7: Machine setup for a seeded gear tooth chip experiment, and Gear 1 with the seeded tooth chip. [20]

of Gear 1 is varied from 10 to 40 Hz at 5 Hz intervals, and for each condition two recordings are performed. A PCB 352C67 accelerometer is attached to the machine casing, and sampled at rates of 2560 and 5120 Hz depending on the rotational frequency. 8192 samples are recorded per measurement. Since there was no direct measurement of the rotational speed of Shaft 1, the rotational period of Gear 1 is estimated for each vibration signal by peak finding on the MOMEDA MKurt spectrum within ±20 samples from the expected period at a step of 0.05 samples. The period corresponding to this peak is used as the rotational period. Table 1 illustrates the characteristics and calculated speed for each measurement under each condition.

Table 1: Measured vibration signals and estimated periods

Dataset	F_s (Hz)	T_{fault} (samples)	$T_{healthy}$ (samples)
10Hz #1	2560	273.35	260.95
10Hz #2	2560	273.35	260.95
15Hz #1	2560	177.52	173.32
15Hz #2	2560	177.47	173.32
20Hz #1	5120	263.55	259.25
20Hz #2	5120	263.55	259.20
25Hz #1	5120	209.65	207.05
25Hz #2	5120	209.60	207.05
30Hz #1	5120	174.22	172.37
30Hz #2	5120	174.17	172.37
35Hz #1	5120	148.99	147.64
35Hz #2	5120	149.09	147.64
40Hz #1	5120	130.40	129.10
40Hz #2	5120	130.45	116.20

Figure 8 illustrates the relationship between filter length L, the fault metric MKurt, and the MKurt difference between healthy and faulty states at each setup condition. Only filter lengths of L equal to or larger than the Gear 1 rotational period were considered for each vibration measurement. Under all filter lengths and machine conditions, the MKurt value under fault were larger than the values under healthy condition. A filter length of $L = 500$ is chosen for the rest of the results as a balance between the lower variance associated with smaller filter lengths, and higher MKurt difference associated with larger filter lengths.
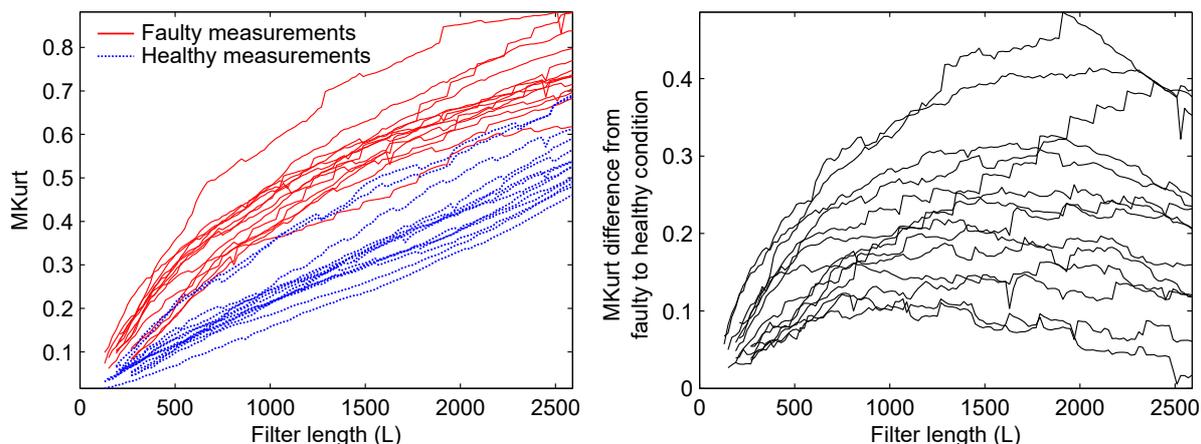
15

Figure 8: MOMEDA fault detection results versus filter length L.

For the first measurement of the machine vibration at 40 Hz, Figure 9 illustrates the faulty and healthy MOMEDA MKurt spectrums, inputs, outputs, and designed FIR filters by MOMEDA. In both the faulty and healthy vibration measurements there is a strong peak that is identifiable at the gear period, however under faulty conditions the peak is larger in magnitude. The presence of a peak even under healthy condition is not unexpected, since factors such as non-uniform force between the gears as the teeth mesh may cause features that can be deconvolved. Figure 10 illustrates the spectrums of all measurements under faulty and healthy conditions.

To compare MOMEDA to the other deconvolution algorithms and to study the impact of AR-model residual preprocessing, we compare the following methods:

MED Minimum Entropy Deconvolution

AR-MED Autoregressive Minimum Entropy Deconvolution

MEDA Minimum Entropy Deconvolution Adjusted

AR-MEDA Autoregressive Minimum Entropy Deconvolution Adjusted

MCKD3 Maximum Correlated Kurtosis Deconvolution 3rd Shift

AR-MCKD3 Autoregressive Maximum Correlated Kurtosis Deconvolution 3rd Shift

MCKD5 Maximum Correlated Kurtosis Deconvolution 5th Shift

AR-MCKD5 Autoregressive Maximum Correlated Kurtosis Deconvolution 5th Shift

OMEDA Optimal Minimum Entropy Deconvolution Adjusted

AR-OMEDA Autoregressive Optimal Minimum Entropy Deconvolution Adjusted

MOMEDA Multipoint Optimal Minimum Entropy Deconvolution Adjusted

AR-MOMEDA Autoregressive Multipoint Optimal Minimum Entropy Deconvolution Adjusted

For parameter selection of the MCKD and MED-based algorithms, we used the same deconvolution iteration count and auto-regressive model parameters as studied previously with this experimental data [20]:

- Termination condition: 100 iterations. This is chosen as larger number to approximate convergence.
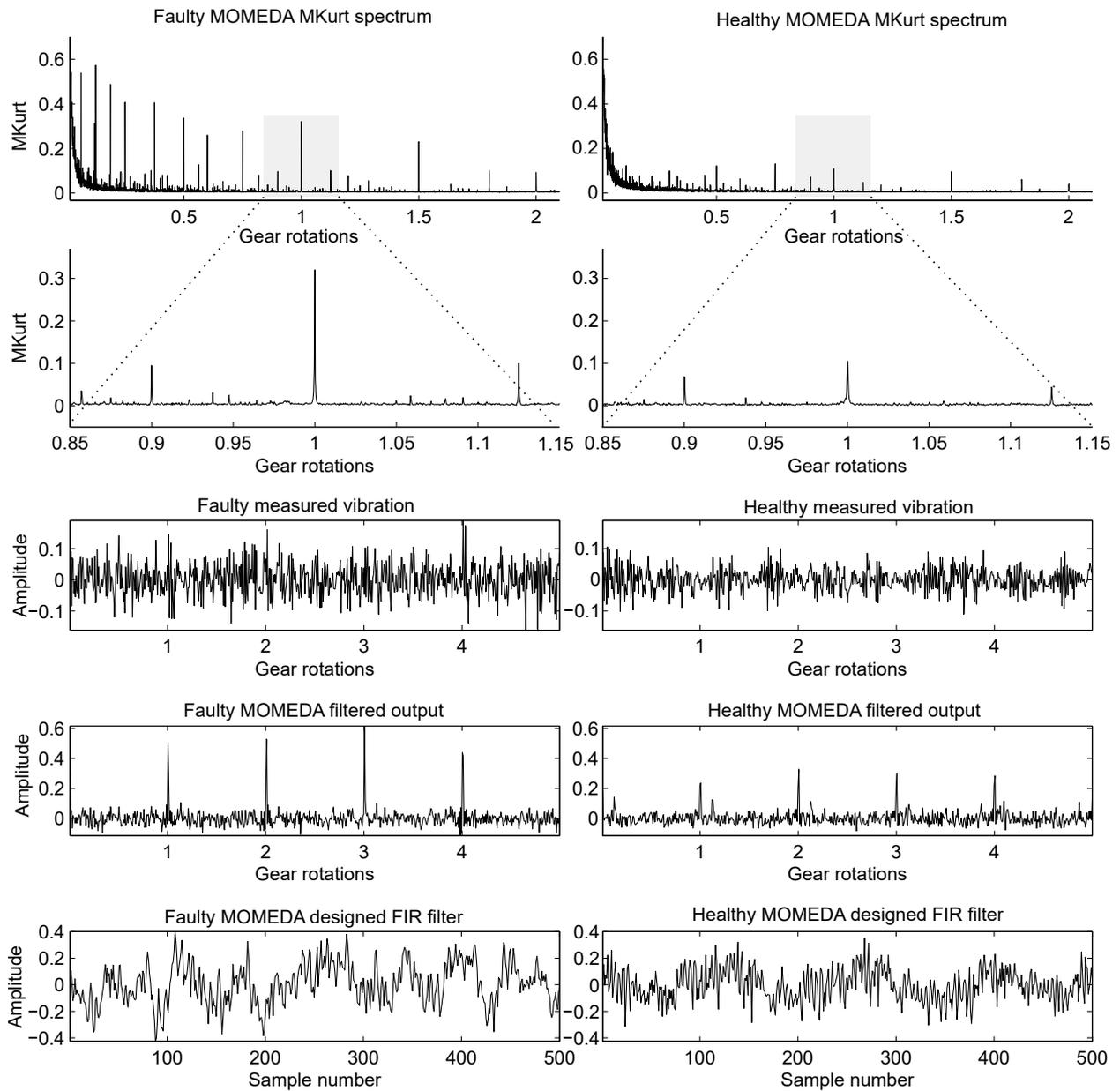
16

Figure 9: Faulty and healthy MOMEDA MKurt spectrums, measured vibrations, MOMEDA filtered outputs, and MOMEDA-designed FIR filters for the first measurement of the machine at 40 Hz using a filter length of 500, $L = 500$.
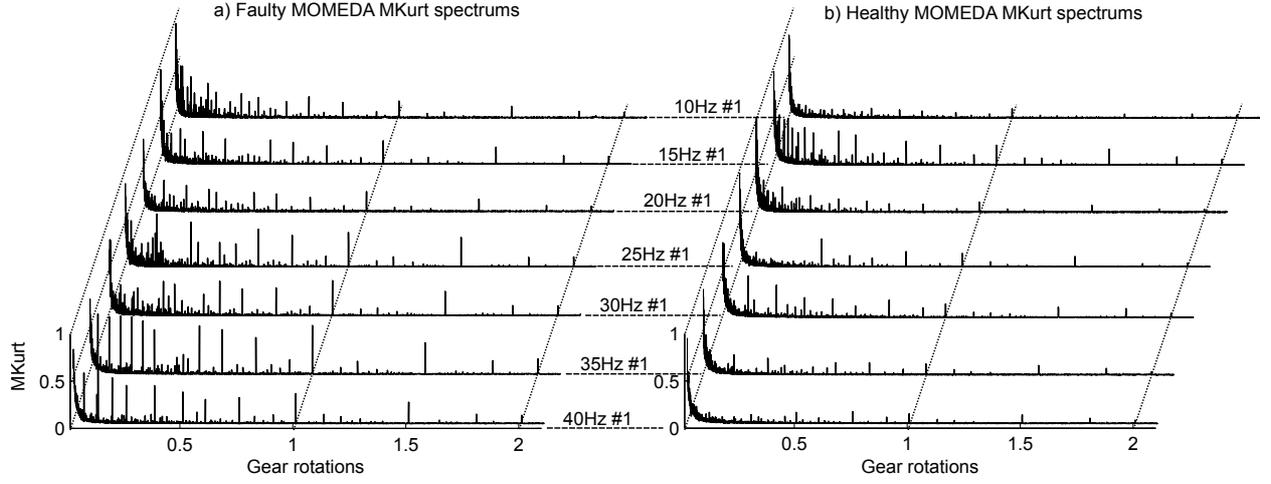
Figure 10: MOMEDA MKurt spectrums for vibrations measured a) with and b) without a gear tooth chip. A step of $\Delta T = 0.05$ and filter length of $L = 500$ is used.

- AR model order: 40% the number of samples in a single rotation.

We found that using the smaller filter lengths for the MED and MCKD-based methods (80% the number of samples in a single rotation) and comparing to the longer filter length results for MOMEDA ($L = 500$) did not fairly represent the non-MOMEDA based methods. Instead, we found the MED and MCKD-based methods had better fault indication performance using a filter length of $L = 500$ and therefore we instead used this filter length with all methods for fair comparison.

Fault indicators vary across methods. MED proposes using Kurtosis as the fault indicator, MCKD proposes using Correlated Kurtosis (CK), and MOMEDA uses MKurt. To compare fault metrics, we analyzed the fault indication performance of all deconvolution methods using each fault metric of Kurtosis, third-shift CK, Multi D-Norm, and MKurt. To compute Multi D-Norm and MKurt for output signals we calculate the maximum result for all shifts of the target impulse train at a step of $\Delta T = 0.05$ where the period of rotation of the faulty gear, T_{gear}, separates the impulses:

$$\text{Peak Multipoint Kurtosis} = \max_{s=1,1+\Delta T,1+2\Delta T,...,T_{gear}} MKurt(\vec{y}, \vec{t_s})$$

$$\text{Peak Multi D-Norm} = \max_{s=1,1+\Delta T,1+2\Delta T,...,T_{gear}} MDN(\vec{y}, \vec{t_s})$$

$$\vec{t_s} = \delta_{\text{round}(s+T_{gear})} + \delta_{\text{round}(s+2T_{gear})} + \delta_{\text{round}(s+3T_{gear})} + \dots$$

For better illustration, we adjusted the scale of the third-shift CK by taking the cubed root of this feature. We processed the faulty versus healthy condition difference between each fault indicator for each recorded signal. Figure 11 plots this fault indication difference along with the means, standard deviations, and 95% confidence intervals. Statistical significances between the mean fault indication differences are evaluated by a one-way repeated measures analysis of variance (ANOVA) and Tukey's honest significance test and considered significant if $p < 0.05$. From this, we draw several observations:

- MOMEDA is able to indicate the fault when using any of the fault indicators. We can not say that the fault detection performance is necessarily better or worse in fault detection compared to other
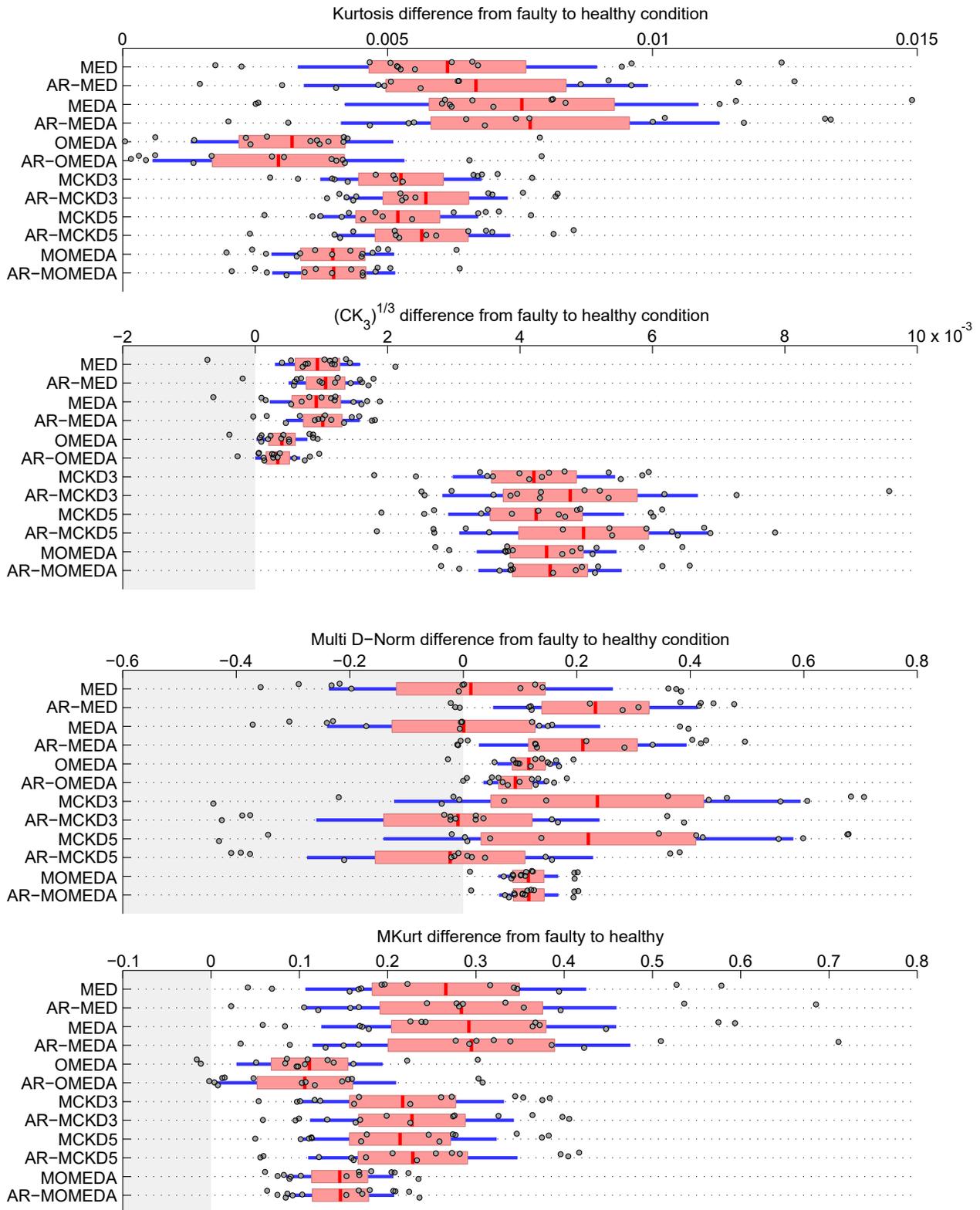
18

Figure 11: Experimental fault indication difference between faulty and healthy conditions for the marked methods with raw difference data markers, means, standard deviation boxes, and 95% confidence intervals.

19

methods. A larger fault indicator difference does not necessarily translate to a better algorithm performance. Ideally, a good fault indication algorithm would indicate a small fault metric increase by a similar amount regardless of machine rotational speed, and this metric would increase in value as fault level increases. Too large of a fault indicator difference in our results could indicate a lack in sensitivity to detect the difference between the level of the fault present in the machine.

- AR model preprocessing did not have a clear positive fault detection result when using Kurtosis, CK, or MKurt fault indicators on any method. When using the Multi D-Norm fault indicator, it appears to have improved the MED and MEDA algorithms but decreased the performance of the MCKD3 and MCKD5 methods. We could not identify a statistically significant effect on fault indication performance by preprocessing with AR models. There may be a significance that could be detected with a larger sample size, or it may have more of an effect on different fault types, machine configurations or equipment setups.

- Although MEDA had a higher mean difference than MED when using the Kurtosis fault indicator, this difference was not found to be statistically significant. A larger sample size may be able to detect if there is a significance. We suggest using MEDA over MED since there are some measurement conditions, such as measurements that have larger low-frequency components or DC offsets, that would cause MED to erroneously deconvolve the discontinuity.

- Care needs to be taken when selecting and comparing the fault indicators for each deconvolution methods. Performance of each method differs according the fault indicator studied.

6. Conclusion

In this paper, we introduced a new deconvolution-based method called MOMEDA for detecting faults in rotating machines. The introduced MOMEDA method has advantages over the previous applied deconvolution methods of MED and MCKD, since it provides a solution for the filter that is an optimal solution, can be solved for directly (non-iteratively), and the target goal in the problem more closely matches the expected fault dynamic sources of an infinite impulse train. By simultaneously solving for a range of impulse-train period targets, we proposed a spectrum analysis approach called MOMEDA MKurt spectrum that can be used to identify faults in rotating machines. We applied MOMEDA and the resulting MKurt spectrum to successfully differentiate between faulty and healthy condition in both simulated data and data from a gearbox with chipped tooth experiment.

We proposed a correction to MED called MEDA. The correction adjusts the convolution definition to remove the input discontinuity. As a result, this resolves some spurious impulse deconvolution issues faced when applying MED to rotation machine signals. We recommend using MEDA over MED going forward to improve reliability, especially when dealing with signals with lower frequency components with respect to the sampling frequency.

When comparing the difference in the fault indicators between healthy and faulty conditions, we found that using filter lengths larger than the period of the fault improved the fault detection results of MED, MEDA, OMEDA, and MCKD methods. Restricting to filter lengths shorter than the fault period should not be considered a requirement, and can limit the fault indication performance of the methods.

In our experimental results, we found no statistically significant difference for fault detection by preprocessing with AR models prior to applying any deconvolution method. There may still be an effect on different machine configurations, different fault types, or there may be a difference that could be detected if a larger number of measurements were studied. Until this difference is demonstrated as significant and

positive, the need to preprocess data with AR models before applying deconvolution remains unproven and perhaps not necessary.

As further work, application of MOMEDA to more experimental setups, fault types, and machine elements could be performed. Studying the effect of the fault indicators and methods according to the level of fault present in the machine is important. Modification of the target vector to instead perform an iterative path-finding approach may be researched as an improvement for machines under slightly varying rotational speeds or for bearings with slippage. For machines with significantly changing speeds, the filter is expected to change drastically as it does so - to better handle these cases, using target vectors designing filters to deconvolving only a few sequential impulses at controlled positions while stepping through the signal may be investigated to track the speed and fault level of specific elements. A comparison of fault identification between MOMEDA, MEDA, MCKD, and non-deconvolution methods is needed to help direct future work as to which direction is more effective. Fundamentally, the fault indicators of Kurtosis, CK, Multi D-Norm, and MKurt are all amplitude invariant and do not represent the magnitude of the impulses or their effect on the measured vibration. Further research should be invested into studying the resulting designed filters along with the magnitude of their result, and using this information to study the effect of the fault on the vibration of the machine to perhaps better indicate the severity or characteristics of the fault.

External Resources

Minimum Entropy Deconvolution MATLAB implementation (MED: overlapMode='full', MEDA: overlapMode='valid'):
http://www.mathworks.com/matlabcentral/fileexchange/29151-minimum-entropy-deconvolution-med-1d-and-2d

M-Shift Maximum Correlated Kurtosis Deconvolution (MCKD):
http://www.mathworks.com/matlabcentral/fileexchange/31326

Optimal Minimum Entropy Deconvolution Adjusted (OMEDA):
http://www.mathworks.com/matlabcentral/fileexchange/53482-optimal-minimum-entropy-deconvolution-with-convolution-fix–non-iterative-solution-

Multipoint Optimal Minimum Entropy Deconvolution Adjusted (MOMEDA):
http://www.mathworks.com/matlabcentral/fileexchange/53483-multipoint-optimal-minimum-entropy-deconvolution-with-convolution-adjustment–momeda-

Appendix A. Normalized Multi D-Norm Derivation

Starting from the Unnormalized Multi D-Norm,

$$\text{Unnormalized Multi D-Norm} = UMDN(\vec{y}, \vec{t}) = \frac{\vec{t}^T \vec{y}}{\|\vec{y}\|},$$

we introduce a normalization factor k:

$$k\,UMDN(\vec{y}, \vec{t}) = k\frac{\vec{t}^T \vec{y}}{\|\vec{y}\|},$$

We normalize Multi D-Norm to a value of 1 when the output $\vec{y}$ reaches a multiple of the goal vector $\vec{t}$, and solve for the normalization factor:

$$1 = k\frac{\vec{t}^T \vec{t}}{\|\vec{t}\|},$$

$$k = \frac{1}{\|\vec{t}\|}.$$

Therefore we use the normalized Multi D-Norm formula as follows:

$$\text{Multi D-Norm} = \frac{1}{\|\vec{t}\|}MDN(\vec{y}, \vec{t}) = \frac{1}{\|\vec{t}\|}\frac{\vec{t}^T \vec{y}}{\|\vec{y}\|}$$

Appendix B. Normalized Multipoint Kurtosis Derivation

Based on Kurtosis, we expand Kurtosis to instead consider multiple points controlled by target vector $\vec{t}$:

$$\text{Unnormalized Multipoint Kurtosis} = UMKurt(\vec{y}, \vec{t}) = \frac{\sum_{n=1}^{N-L} (t_n y_n)^4}{\left(\sum_{n=1}^{N-L} y_n^2\right)^2}$$

We introduce a normalization factor k:

$$\text{Multipoint Kurtosis} = MKurt(\vec{y}, \vec{t}) = k\frac{\sum_{n=1}^{N-L} (t_n y_n)^4}{\left(\sum_{n=1}^{N-L} y_n^2\right)^2},$$

We normalize MKurt to a value of 1 when the output $\vec{y}$ reaches a multiple of the goal vector $\vec{t}$, and solve for the normalization factor:

$$1 = k\frac{\sum_{n=1}^{N-L} \left(t_n^2\right)^4}{\left(\sum_{n=1}^{N-L} t_n^2\right)^2},$$

$$k = \frac{\left(\sum_{n=1}^{N-L} t_n^2\right)^2}{\sum_{n=1}^{N-L} t_n^8}.$$

Therefore we use the normalized MKurt formula as follows:

$$\text{Multipoint Kurtosis} = MKurt(\vec{y}, \vec{t}) = \frac{\left(\sum_{n=1}^{N-L} t_n^2\right)^2}{\sum_{n=1}^{N-L} t_n^8}\frac{\sum_{n=1}^{N-L} (t_n y_n)^4}{\left(\sum_{n=1}^{N-L} y_n^2\right)^2}$$

References

[1] A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical systems and signal processing*, vol. 20, no. 7, pp. 1483–1510, 2006.

[2] R. Rubini and U. Meneghetti, "Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings," *Mechanical systems and signal processing*, vol. 15, no. 2, pp. 287–302, 2001.

[3] S. K. Goumas, M. E. Zervakis, and G. S. Stavrakakis, "Classification of washing machines vibration signals using discrete wavelet analysis for feature extraction," *Instrumentation and Measurement, IEEE Transactions on*, vol. 51, no. 3, pp. 497–508, 2002.

[4] X. Fan and M. J. Zuo, "Gearbox fault detection using hilbert and wavelet packet transform," *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 966 – 982, 2006.

[5] R. Jiang, J. Chen, G. Dong, T. Liu, and W. Xiao, "The weak fault diagnosis and condition monitoring of rolling element bearing using minimum entropy deconvolution and envelope spectrum," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, p. 0954406212457892, 2012.

[6] N. G. Nikolaou and I. A. Antoniadis, "Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted morlet wavelets," *Mechanical Systems and Signal Processing*, vol. 16, no. 4, pp. 677 – 694, 2002.

[7] T. Barszcz and R. B. Randall, "Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1352 – 1365, 2009.

[8] N. Sawalhi, R. Randall, and H. Endo, "The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2616 – 2633, 2007.

[9] J. Antoni and R. Randall, "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines," *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308 – 331, 2004.

[10] R. Isermann, "Model-based fault-detection and diagnosis - status and applications," *Annual Reviews in Control*, vol. 29, no. 1, pp. 71 – 85, 2005.

[11] W. Wang and A. K. Wong, "Autoregressive model-based gear fault diagnosis," *Journal of Vibration and Acoustics*, vol. 124, no. 2, pp. 172–179, 2002.

[12] G. McDonald and Q. Zhao, "Model-based adaptive frequency estimator for gear crack fault detection," in *American Control Conference (ACC), 2011*, pp. 792–797, IEEE, 2011.

[13] M. Yang and V. Makis, "Arx model-based gearbox fault detection and localization under varying load conditions," *Journal of Sound and Vibration*, vol. 329, no. 24, pp. 5209 – 5221, 2010.

[14] X. Wang and V. Makis, "Autoregressive model-based gear shaft fault diagnosis using the kolmogorov-smirnov test," *Journal of Sound and Vibration*, vol. 327, no. 3-5, pp. 413 – 423, 2009.

[15] J. Liu, D. Xu, and X. Yang, "Sensor fault detection in variable speed wind turbine system using $h - \infty$ method," in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, pp. 4265 –4269, june 2008.

[16] A. C. McCormick and A. K. Nandi, "Cyclostationarity in rotating machine vibrations," *Mechanical Systems and Signal Processing*, vol. 12, no. 2, pp. 225 – 242, 1998.

[17] C. Capdessus, M. Sidahmed, and J. L. Lacoume, "Cyclostationary processes: Application in gear faults early diagnosis," *Mechanical Systems and Signal Processing*, vol. 14, no. 3, pp. 371 – 385, 2000.

[18] T. Barszcz and N. Sawalhi, "Fault detection enhancement in rolling element bearings using the minimum entropy deconvolution," *Archives of acoustics*, vol. 37, no. 2, pp. 131–141, 2012.

[19] H. Endo and R. Randall, "Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 906 – 919, 2007.

[20] G. L. McDonald, Q. Zhao, and M. J. Zuo, "Maximum correlated kurtosis deconvolution and application on gear tooth chip fault detection," *Mechanical Systems and Signal Processing*, vol. 33, pp. 237–255, 2012.

[21] X. Lou and K. A. Loparo, "Bearing fault diagnosis based on wavelet transform and fuzzy inference," *Mechanical Systems and Signal Processing*, vol. 18, no. 5, pp. 1077 – 1095, 2004.

[22] R. Huang, L. Xi, X. Li, C. R. Liu, H. Qiu, and J. Lee, "Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 193 – 207, 2007.

[23] B. A. Paya, I. I. Esat, and M. N. M. Badi, "Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor," *Mechanical Systems and Signal Processing*, vol. 11, no. 5, pp. 751 – 765, 1997.

[24] R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 16, no. 1-2, pp. 21 – 35, 1978.

[25] H. Endo, R. Randall, and C. Gosselin, "Differential diagnosis of spall vs. cracks in the gear tooth fillet region: Experimental validation," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 636 – 651, 2009.

[26] C. A. Cabrelli, "Minimum entropy deconvolution and simplicity: A noniterative algorithm," *Geophysics*, vol. 50, no. 3, pp. 394–413, 1985.

[27] X. Tian, J. Lin, M. Agelinchaab, J. Zhang, M. Akhtar, M. Zuo, and K. Fyfe, "Technical report for experiment 02-05 on gearbox," tech. rep., Reliability Research Laboratory, Dept. Mech. Eng., University of Alberta, 2002.