The design of optimal indicators for early fault detection using a generalized likelihood ratio test

Kayacan Kestel[1 2], Jérôme Antoni[2], Cédric Peeters[1], Quentin Leclère[2],
François Girardin[2], Ted Ooijevaar[3], Jan Helsen[1]

[1]Vrije Universiteit Brussel - VUB, Department of Mechanical Engineering,
Pleinlaan 2, Brussels, 1050, Belgium
[2]Univ Lyon, INSA-Lyon, LVA, EA677, F-69621 Villeurbanne, France
[3]Corelab MotionS, Gaston Geenslaan 8, 3001 Leuven, Belgium

Abstract

This study elaborates on a methodology to design health indicators for vibration-based condition monitoring of rotating machines for early fault detection. These indicators are optimal in maximizing the probability of detection given a constant rate of false alarm. Two probability density functions (PDF) modeling the vibration signals' healthy and faulty states are exploited to generate health indicators using a generalized likelihood ratio test. The key point is formulating a framework to express the health indicators as the function of the partial derivatives of the PDF of the healthy state and a modulation function. The modulation function allows the detection of slight deviations from the healthy state of the vibration signals. Furthermore, it is shown that specific choices of the modulation functions recover conventional health indicators such as kurtosis, skewness, ℓ_p/ℓ_q-norms, the negentropy of the squared envelope, etc. Since the indicators are asymptotically distributed as a chi-squared distribution, a statistical threshold can be estimated to assess the machine's state with respect to its healthy state. The performance of the thresholds is demonstrated on simulated and experimental vibration signals. It is shown that the transition from the healthy to the faulty state of the machine can be detected via the threshold. An important conclusion of this study is that many conventional health indicators are optimal if and only if the healthy state of the machines is Gaussian. The proposed methodology shows how to design indicators for non-Gaussian descriptions of the healthy state and can pave the way for the development of many other health indicators by carefully selecting the PDF of the healthy state and the modulation function.

1 Introduction

Vibration-based condition monitoring has extensively utilized various health indicators (HI) such as kurtosis, skewness, Jarque-Bera statistic, and peak factor [1]. Dyer and Stewart first proposed kurtosis for fault detection, and it has been used as a measure of a signal's distribution's heavy tails or as a sparsity measure in numerous studies [2, 3, 4, 5, 6]. Skewness, introduced by Martin and Honarvar, evaluates the distribution's asymmetry and is useful for detecting non-Gaussian features in fault signals. It has also been used as an objective function for an optimization algorithm for bearing fault detection [7]. The Jarque-Bera statistic, which combines kurtosis and skewness, was initially introduced in economics and derived from the Lagrange multiplier test [8]. The peak factor, also known as crest factor, is another indicator employed in vibration-based condition monitoring, and its usage alongside kurtosis is discussed in [9, 10]. Sparsity indicators include the ℓ_p/ℓ_q-norm families and negentropy of the squared envelope [1], with the former being the most commonly used [11, 12, 3] and additional statistical indicators assessing impulsiveness using ℓ_p/ℓ_q being derived [13]. Bozchalooi and Liang introduced the geometric-to-arithmetic ratio of signals, or smoothness index [1], to the field of vibration-based condition monitoring to address the issue of extremely low signal-to-noise ratios [14]. Negentropy has also been utilized as an indicator of deviation from Gaussianity [12, 15]. Many of these indicators lack a comprehensive mathematical foundation to justify their application and they have been introduced to the field through heuristic means [1].

The development of HIs in vibration-based condition monitoring has generally been based on existing indicators [2]. However, literature has also introduced statistical tests to interpret common conventional indicators.

Kurtosis and skewness were demonstrated to be powerful tools for normality tests [16], and recovered using Lagrange multiplier tests [17, 8]. Statistical tests such as the Rao, Wald, and generalized likelihood ratio (GLRT) have also been used for vibration-based fault detection [18]. Log-envelope-based indicators have been proposed based on the Priestly and Subba Rao test to detect a fault in multiple second-order cyclostationary noises [19]. Recently, Antoni and Borghesani proposed new condition monitoring indicators developed using the GLRT [2]. These were generated using probability density functions (PDF) of a Gaussian, a generalized Gaussian (GG), and a Bernoulli-Gaussian mixture distribution. These indicators were found to be similar to conventional indicators like Kurtosis. Antoni et al. proposed an extension of these indicators [2] to construct condition indicators using GLRT for the frequency domain [20]. However, the proposed indicators have limitations, as the forms of the rooting PDFs are limited because the GLRT requires the tested PDFs to be nested.

This study employs the approach proposed in the previous work of Antoni et al. [2] to develop HIs in an optimal manner, which maximizes the probability of detecting faults while maintaining a constant false alarm probability. This study focuses on early fault detection when the healthy state's PDF undergoes slight changes. The HIs are rooted in two PDFs describing the healthy and faulty conditions. The main contribution of this study is the derivation of a theorem that provides a closed-form expression for the optimal HIs in terms of the partial derivatives of the PDF under the healthy state and the modulation function that characterizes its distortion in the faulty state. Additionally, this study provides the asymptotic PDF of the HIs, which facilitates a comparison to a threshold. Furthermore, the theorem shows that many classical HIs can be produced as specific cases, including kurtosis, ℓ_1/ℓ_2 indicator, geometric to arithmetic mean ratio, and negentropy of the squared envelope. To the authors' knowledge, the thresholds associated with many of these HIs have not been previously reported in the field of condition monitoring.

2 Theoretical framework

This section presents a general framework for designing optimal HIs to detect early faults. The methodology is based on Neyman-Pearson theory, which maximizes the probability of detecting faults while maintaining a constant false alarm rate.

Consider a sequence of data samples $\mathbf{x} = (x_0, \ldots, x_{L-1})^T \in \mathbb{R}^L$ of length L. For simplicity, the samples are assumed to be identically and independently distributed (iid) and centered (i.e., $\mathbb{E}x = 0$) without loss of generality. The healthy state of the system is denoted as H_0, and the faulty states as H_1, each characterized by a distinct PDF. A health indicator is defined as a scalar quantity $I(\mathbf{x})$, computed from the data $\mathbf{x}$, which takes different values depending on whether the data originates from the healthy or faulty state. Ideally, the HI is zero in the healthy state and strictly positive in the faulty state. Formally,

Definition 2.1 *An HI is a functional $I(\mathbf{x}) : \mathbb{R}^L \to \mathbb{R}$ such that*

$$\begin{cases} \lim_{L\to\infty} I(\mathbf{x}|H_0) & = & 0 \quad \text{(healthy state)} \\ \lim_{L\to\infty} I(\mathbf{x}|H_1) & > & 0 \quad \text{(faulty state).} \end{cases} \tag{1}$$

The optimal HI is designed in the Neyman-Pearson sense. It maximizes the probability of detection $\mathbb{P}(I(\mathbf{x}) > \gamma|H_1)$ given a constant probability of false alarm $\mathbb{P}(I(\mathbf{x}) > \gamma|H_0) = \alpha$, where $\gamma > 0$ is a threshold.

The optimal detector is returned by the likelihood ratio $p(\mathbf{x}|H_1)/p(\mathbf{x}|H_0)$, where $p(\mathbf{x}|H_1)$ and $p(\mathbf{x}|H_0)$ stand for the PDFs of the data under H_1 and H_0, respectively. In general, for iid data, one will have $p(\mathbf{x}|H_i) = \Pi_{n=0}^{L-1} p_i(x_n; \boldsymbol{\theta}_i)$, with p_i the PDF of a single observation x_n specified by the vector of parameters $\boldsymbol{\theta}_i$ under H_i, $i = 0, 1$. In most instances, the latter will be partially or even fully unknown, and therefore must be estimated from the data $\mathbf{x}$ themselves. Based on these ideas, reference [2] proposed to define an HI as the average of the logarithm of the generalized likelihood ratio (GLR),

$$I(\mathbf{x}) = \left\langle \ln p_1(x_n; \hat{\boldsymbol{\theta}}_1) - \ln p_0(x_n; \hat{\boldsymbol{\theta}}_0) \right\rangle, \tag{2}$$

where $\hat{\boldsymbol{\theta}}_i$ stands for the maximum likelihood estimate (MLE) of $\boldsymbol{\theta}_i$ under H_i, $i = 0, 1$ and $\langle \bullet \rangle = \frac{1}{L}\sum_{n=0}^{L-1}(\bullet)$ for the time averaging operator. The so-defined HI has many nice properties, which were thoroughly studied in [2].

This is now specialized to the case where one wants to detect a small – i.e. early – transition from H_0 to H_1. For simplicity, let $\boldsymbol{\theta}_0 = \theta_0$ be a scalar parameter, and $\boldsymbol{\theta}_1 = (\theta_1, \nu)$ a two-element vector. In order to model a slight change of p_0 towards p_1, let us consider the particular but rather flexible form

$$p_1(x;\theta_1,\nu) = p_0(x;\theta_1)\phi(x;\theta_1,\nu) \tag{3}$$

where $\phi(x_n;\theta_1,\nu)$ is a positive continuous function of ν such that

$$(a) \quad \phi(x_n;\theta_1,0) = 1, \quad \text{and} \quad (b) \quad \int_{\mathbb{R}} p_0(x;\theta)\phi(x_n;\theta,\nu)dx = 1, \tag{4}$$

thus ensuring that p_1 is a PDF such that $p_1(x;\theta,0) = p_0(x;\theta)$. The function $\phi(x;\theta_1,\nu)$ modulates the PDF p_0 to transform it into p_1 as the system transits from state H_0 to H_1, when ν departs from zero ($\nu = 0$ under H_0). The tuning parameter ν determines the magnitude of this deviation, and the maximum likelihood estimate (MLE) $\hat{\nu}$ is expected to be closely linked to the design of a health indicator (HI), such that $\hat{\nu} \approx 0$ only under H_0.

Suppose $p_0(x;\theta)$ and $\phi(x;\theta,\nu)$ are twice differentiable with respect to θ and ν. Then, if (i) the parameterization (3) is locally identifiable and if (ii) $\left\langle \frac{\partial \ln \phi}{\partial \nu}(x_n;\hat{\theta}_0,0) \right\rangle \neq 0$, the MLE of ν given a small deviation from state H_0 is

$$\hat{\nu} = \frac{\mathbb{I}_0}{|\mathbb{I}_1|} \left\langle \frac{\partial \ln \phi}{\partial \nu}(x_n;\hat{\theta}_0,0) \right\rangle + \mathscr{O}(\hat{\nu}^2) \tag{5}$$

where $\mathscr{O}(\hat{\nu}^2)$ represents the high-order terms, $\hat{\theta}_0$ is the MLE of θ under the H_0 assumption, i.e. such that $\frac{\partial}{\partial \theta}\langle \ln p_0(x_n;\hat{\theta}_0)\rangle = 0$, $\mathbb{I}_0 = -\mathbb{E}_{H_0}\left\{ \frac{\partial^2}{\partial \theta^2} \ln p_0(x;\theta_0) \right\}$ is the Fisher information of the statistical model described by H_0 and

$$\mathbb{I}_1 = -\mathbb{E}_{H_0} \begin{pmatrix} \frac{\partial^2 \ln p_1}{\partial \theta^2}(x;\theta_0,0) & \frac{\partial^2 \ln p_1}{\partial \nu \partial \theta}(x;\theta_0,0) \\ \frac{\partial^2 \ln p_1}{\partial \nu \partial \theta}(x;\theta_0,0) & \frac{\partial^2 \ln p_1}{\partial \nu^2}(x;\theta_0,0) \end{pmatrix} = -\mathbb{E}_{H_0} \begin{pmatrix} \frac{\partial^2 \ln p_0}{\partial \theta^2}(x;\theta_0,0) & \frac{\partial^2 \ln \phi}{\partial \nu \partial \theta}(x;\theta_0,0) \\ \frac{\partial^2 \ln \phi}{\partial \nu \partial \theta}(x;\theta_0,0) & \frac{\partial^2 \ln \phi}{\partial \nu^2}(x;\theta_0,0) \end{pmatrix} \tag{6}$$

is the Fisher information matrix of the statistical model described by H_0, all evaluated at $\theta_0 = \mathbb{E}_{H_0}\{\hat{\theta}_0|H_0\}$, with $\mathbb{E}_{H_0}\{\cdots\} = \mathbb{E}\{\cdots|H_0\}$ the expected value with respect to H_0. Accordingly, under the abovementioned condition,

$$\begin{cases} \lim_{L\to\infty} \hat{\nu}|H_0 & = & \mathbb{E}\{\hat{\nu}|H_0\} & = 0 \\ \lim_{L\to\infty} \hat{\nu}|H_1 & = & \mathbb{E}\{\hat{\nu}|H_1\} & \neq 0. \end{cases} \tag{7}$$

Thus, the HI defined by Eq. (2) has the asymptotic expression

$$I(\mathbf{x}) = \frac{\left\langle \frac{\partial \ln \phi}{\partial \nu}(x_n;\hat{\theta}_0,0) \right\rangle^2}{2\left(\mathbb{E}\left\{ \frac{\partial^2 \ln \phi}{\partial \theta \partial \nu}(x;\theta_0,0)|H_0 \right\}^2 / \mathbb{E}\left\{ \frac{\partial^2 \ln p_0}{\partial \theta^2}(x;\theta_0)|H_0 \right\} - \mathbb{E}\left\{ \frac{\partial^2 \ln \phi}{\partial \nu^2}(x;\theta_0,0)|H_0 \right\} \right)} + \mathscr{O}(\hat{\nu}^3), \tag{8}$$

which follows the Chi-squared distribution $\chi_1^2/(2L)$.

In accordance with Eq. (8), $I(\mathbf{x})$ clearly satisfies definition (2.1). The so-defined HI can be used as such, yet any positive and monotonic function of it will also be a valid HI. This remark will be useful to recover some classical indicators used in condition monitoring.

2.1 Interpretation of the optimal HIs

The proposed HI can be interpreted either in light of the perturbation parameter ν or, in some instances, in light of another known indicator derived from it.

According to condition given in Eq. (7), $\hat{\nu}$ reflects the deviation from H_0 to H_1, and squaring it will make it a valid metric to measure the distance between p_0 and p_1. In addition, from Eq. (8), one has $\mathbb{E}\{I(\mathbf{x})\} = \mathbb{I}_0/(2|\mathbb{I}_1|)\mathbb{V}\{\hat{\nu}|H_0\}$ where $\mathbb{E}\{I(\mathbf{x})\} = 1/(2L)$ since $I(\mathbf{x}) \sim \chi_1^2/(2L)$; therefore $\mathbb{V}\{\hat{\nu}|H_0\}\{\hat{\nu}^2\} = |\mathbb{I}_1|/(L\mathbb{I}_0)$ and

$$I(\mathbf{x}) = \frac{1}{2L}\left(\frac{\hat{\nu}}{\sqrt{\mathbb{V}\{\hat{\nu}|H_0\}}} \right)^2. \tag{9}$$

In other words, the HI $I(\mathbf{x})$ turns out to be the square of the standardised perturbation parameter $\hat{\nu}$.

It happens that in many cases investigated hereafter in the paper, the perturbation parameter can be related to a known indicator $\hat{s}(\mathbf{x})$ through the relationship

$$\hat{\nu} = a(\hat{s}(\mathbf{x}) - s_0), \tag{10}$$

for some constant a and where $s_0 = \mathbb{E}\{\hat{s}(\mathbf{x})|H_0\}$. Therefore,

$$I(\mathbf{x}) = \frac{1}{2L} \frac{(\hat{s}(\mathbf{x}) - s_0)^2}{\mathbb{V}\{\hat{s}(\mathbf{x})\}\{\hat{\nu}|H_0\}}. \tag{11}$$

This expression confers on the optimal HI $I(\mathbf{x})$ an interpretation in terms of the square of the centred and standardized indicator $\hat{s}(\mathbf{x})$. As will be seen in the following examples, it also offers a systematic way to calculate the mean and variance of many classical HIs.

2.2 The linear modulation

The design of the modulation function $\phi(x;\theta,\nu)$ is crucial in the proposed methodology. It should ideally be informed by the theoretical knowledge of the PDF p_1 representing the faulty state, in comparison to the reference PDF p_0 representing the healthy state. However, considering the interest in small deviations of p_0 characterized by small values of ν, it is expected that ϕ can be approximated by a polynomial form (e.g., a Taylor expansion) in ν. Since Eq. (8) only involves derivatives up to the second order, the polynomial can be truncated to the second order:

$$\phi(x;\theta,\nu) = 1 + \nu g_1(x;\theta) + \nu^2 g_2(x;\theta) + R(x;\theta,\nu), \tag{12}$$

with $R(x;\theta,\nu)$ a residual of higher order. In order for $\phi(x;\theta,\nu)$ to be a valid modulation function, condition (b) of Eq. (4) requires that $\mathbb{E}\{g_1(x;\theta)|H_0\} = \mathbb{E}\{g_2(x;\theta)|H_0\} = 0$ and $\mathbb{E}\{R(x;\theta,\nu)|H_0\} = 0$, $\forall\nu$. Therefore, one has

$$\begin{cases} \left\langle \frac{\partial}{\partial\nu} \ln\phi(x_n;\theta,0) \right\rangle = \langle g_1(x_n;\theta)\rangle & \text{(13a)} \\ \mathbb{E}\left\{ \frac{\partial^2}{\partial\nu^2} \ln\phi(x_n;\theta,0)|H_0 \right\} = -\mathbb{E}\{g_1(x;\theta)|H_0\}^2 = -\mathbb{E}\left\{ \frac{\partial}{\partial\nu} \ln\phi(x_n;\theta,0)|H_0 \right\}^2 & \text{(13b)} \\ \mathbb{E}\left\{ \frac{\partial^2}{\partial\theta\nu} \ln\phi(x_n;\theta,0)|H_0 \right\} = \mathbb{E}\left\{ \frac{\partial}{\partial\theta} g_1(x_n;\theta,0)|H_0 \right\}. & \text{(13c)} \end{cases}$$

The definitions above are plugged into the expression defined in Eq. (8) to obtain the health indicators.

2.3 Recovery of the indicators

This section focuses on constructing optimal HIs based on the Gaussian PDF p_0, which holds significant practical importance. In this context, the widely accepted assumption that the normal distribution accurately describes the healthy state of a system is adopted. This assumption is often supported by the central limit theorem, which states that the sum of many random variables with finite and somewhat comparable variances tends to follow a Gaussian distribution[1]. Relying on the Gaussian assumption enables the derivations of well-known indicators commonly used for condition monitoring.

The Gaussian has the standard deviation θ for its free parameter. Its MLE is the solution of the equation $\langle \partial_\theta \ln p_0(x_n;\theta)\rangle = 0$, which reads

$$\hat{\theta}_0 = \sqrt{\langle |x_n|^2 \rangle}. \tag{14}$$

The expected curvature at $\theta_0 = \mathbb{E}\{\hat{\theta}_0|H_0\}$ is then returned by

$$\mathbb{E}\left\{ \frac{\partial^2 \ln p_0}{\partial\theta^2}(x_n;\theta_0) \right\} = -\frac{2}{\theta_0^2}. \tag{15}$$

[1] A precise statement is given by Lindeberg's condition [21]

2.3.1 Recovery of the ℓ_q/ℓ_2 indicators

The distortion of the Gaussian PDF by a simple polynomial (with respect to x) modulation:

$$\begin{cases} H_0: & p_0(x;\theta) = \frac{e^{-\frac{x^2}{2\theta^2}}}{\sqrt{2\pi}\theta}, \quad 0<\theta \\ H_1: & p_1(x;\theta,\nu) = p_0(x;\theta)\underbrace{\left(1+\nu\left(\frac{1}{m_\beta}\left|\frac{x}{\theta}\right|^q-1\right)\right)}_{\phi(x;\theta,\nu)}, \quad 0<q, \quad 0\leq\nu\leq 1. \end{cases} \tag{16}$$

The modulation function models an inflation of the tails of a PDF, for instance because of a series of transients in the signals due to a fault. Condition (4)(b) then implies that

$$m_\beta = \mathbb{E}\{|x/\theta|^\beta|H_0\} = \int_{\mathbb{R}} p_0(x;1)\,|x|^\beta\,dx = 2^{\frac{\beta}{2}}\frac{\Gamma((\beta+1)/2)}{\sqrt{\pi}}, \tag{17}$$

and the condition of positivity of ϕ that $0\leq\nu\leq 1$, which places a limit on the distortion imposed to p_0 in the faulty state. Alternatively, the β-th moment under H_1 is given by

$$\begin{aligned} \mathbb{E}\{|x/\theta|^q|H_1\} &= \int_{\mathbb{R}} p_1(x;1,\nu)|x|^q dx = \int_{\mathbb{R}} p_0(x;1)(1+\nu(|x|^q/m_q-1))|x|^q dx \\ &= m_q+\nu\left(\frac{m_{2q}}{m_q}-m_q\right). \end{aligned} \tag{18}$$

This proves that $\mathbb{E}\{|x/\theta|^q|H_1\}-\mathbb{E}\{|x/\theta|^q|H_0\}$ linearly increases with ν.

The implementation of Eq. (8) starts by calculating the partial derivatives

$$\begin{cases} \frac{\partial\ln\phi}{\partial\nu}(x;\theta,0) = g_1(x;\theta) = \frac{1}{m_q}\left|\frac{x}{\theta}\right|^q-1 \\ \frac{\partial^2\ln\phi}{\partial\nu^2}(x;\theta,0) = -g_1(x;\theta) = -\frac{1}{m_q^2}\left(\left|\frac{x}{\theta}\right|^q-1\right)^2 \\ \frac{\partial^2\ln\phi}{\partial\nu\theta}(x;\theta,0) = \frac{\partial g_1}{\partial\theta}(x;\theta) = -\frac{1}{m_q}\frac{q}{\theta}\left|\frac{x}{\theta}\right|^q \end{cases} \tag{19}$$

Substituting the MLE (14) for θ into the first partial derivative and taking the average then returns

$$\left\langle\frac{\partial\ln\phi}{\partial\nu}(x_n;\hat{\theta}_0,0)\right\rangle = \frac{1}{m_q}\left(\frac{\langle|x_n|^q\rangle}{\langle x_n^2\rangle^{q/2}}-1\right). \tag{20}$$

Then, taking the expected value under H_0 of the last two partial derivatives evaluated at $\theta_0=\mathbb{E}_{H_0}\{\hat{\theta}_0\}=m_2^{1/2}$,

$$\begin{cases} \mathbb{E}_{H_0}\frac{\partial^2\ln\phi}{\partial\nu^2}(x;\theta_0,0) = -m_{2q}+m_q^2 \\ \mathbb{E}_{H_0}\frac{\partial^2\ln\phi}{\partial\nu\theta}(x;\theta_0,0) = -\beta m_q/m_2^{1/2}. \end{cases} \tag{21}$$

Then, the definition in Eq. (5) yields

$$\hat{\nu} = \frac{(\hat{s}_{q,2}(\mathbf{x})/m_q-1)}{m_{2q}-m_q^2(1+q^2/2)}, \quad \text{where} \quad \hat{s}_{q,2}(\mathbf{x}) = \frac{\langle|x_n|^q\rangle}{\langle x_n^2\rangle^{q/2}}. \tag{22}$$

Similarly, application of Eq. (8) yields

$$I(\mathbf{x}) = \frac{(\hat{s}_{q,2}(\mathbf{x})-m_q)^2}{2\left(m_{2q}-m_q^2(1+q^2/2)\right)}, \tag{23}$$

which is asymptotically distributed like $2/L\cdot\chi^2_{1,1-\alpha}$ under H_0. The ℓ_q/ℓ_2 indicator particularizes to two popular cases commonly used in condition monitoring, when $q=4$ and $q=1$, which are investigated hereafter.

2.3.2 Recovery of the kurtosis

With $q=4$,

$$\hat{s}_{4,2}(\mathbf{x}) = \frac{\langle|x_n|^4\rangle}{\langle x_n^2\rangle^2} \tag{24}$$

is the kurtosis. In this case, $m_4=3$ and $m_8=105$, so that

$$I(\mathbf{x}) = \frac{(\hat{s}_{4,2}(\mathbf{x})-3)^2}{48}, \quad 0\leq\nu\leq\frac{1}{3}. \tag{25}$$

2.3.3 Recovery of the ℓ_1/ℓ_2 indicator

Similarly, with $q = 1$, the ℓ_1/ℓ_2 indicator is recovered:

$$\hat{s}_{1,2}(\mathbf{x}) = \frac{\langle |x_n| \rangle}{\sqrt{\langle x_n^2 \rangle}}. \tag{26}$$

Then, $m_1 = \sqrt{2/\pi}$ and $m_2 = 1$, yielding

$$I(\mathbf{x}) = \frac{(\hat{s}_{1,2}(\mathbf{x}) - \sqrt{2/\pi})^2}{2(1-3/\pi)}, \quad 0 \leq \nu \leq \sqrt{\frac{\pi}{2}}. \tag{27}$$

2.3.4 Recovery of the negentropy of the squared envelope

A different modulation function is now considered, which continuously transforms the Gaussian PDF into a generalized Gaussian. The statistical model reads

$$\begin{cases} H_0: & p_0(x;\theta) & = & \frac{e^{-\frac{x^2}{2\theta^2}}}{\sqrt{2\pi}\theta}, \quad 0 < \theta \\ H_1: & p_1(x;\theta,\nu) & = & \frac{1+\nu/2}{\sqrt{2}\theta\Gamma\left(\frac{1}{2+\nu}\right)} e^{-\left|\frac{x}{\sqrt{2}\theta}\right|^{2+\nu}} = p_0(x;\theta)\phi(x;\theta,\nu), \quad \nu > -2 \end{cases} \tag{28}$$

with $\phi(x;\theta,\nu) = \left(1+\frac{\nu}{2}\right)\frac{\sqrt{\pi}}{\Gamma\left(\frac{1}{2+\nu}\right)} e^{\frac{x^2}{2\theta^2} - \left|\frac{x}{\sqrt{2}\theta}\right|^{2+\nu}}$.

Following the same calculation steps as in the previous example, Eq. (8) then yields

$$I(\mathbf{x}) = \frac{(\hat{H}_2(\mathbf{x}) + \ln 2 + \gamma - 2)^2}{3\pi^2 - 7} \simeq \frac{(\hat{H}_2 - 0.7296)^2}{1.6091}, \tag{29}$$

where $\gamma \simeq 0.5772$ is Euler's constant, and

$$\hat{H}_2(\mathbf{x}) = \left\langle \frac{x_n^2}{\langle x_n^2 \rangle} \ln \frac{x_n^2}{\langle x_n^2 \rangle} \right\rangle \tag{30}$$

is the negentropy of the squared envelope x_n^2.

To the best of the authors' knowledge, the use of the negentropy of the squared envelope as an HI for condition monitoring has never been justified on a formal basis before, although it has been suggested on intuitive grounds [15]. Even less obvious is the threshold that comes with it.

2.3.5 Recovery of the Geometric-to-Arithmetic mean of the squared envelope

Let $\mathbf{y} = (x_0^2, \ldots, x_{L-1}^2)^T$ collects the squares of the data $\mathbf{x}$, also known as the squared envelope. When $\mathbf{x}$ is sampled from the Gaussian distribution under H_0, $\mathbf{y}$ will follow a Chi-squared. The Chi-squared PDF $p_0(y;\theta)$ is a special case of the Gamma distribution, which can represent the PDF $p_1(y;\theta,\nu)$ under assumption H_1. This is expressed as

$$\begin{cases} H_0: & p_0(y;\theta) & = & \frac{\sqrt{\theta}}{\Gamma(1/2)} y^{-\frac{1}{2}} e^{-\theta y}, \quad \theta = \frac{1}{2\sigma^2}, \\ H_1: & p_1(y;\theta,\nu) & = & \frac{\theta^{\nu+\frac{1}{2}}}{\Gamma(\nu+1/2)} y^{\nu-\frac{1}{2}} e^{-\theta y} = p_0(x;\theta)\phi(x;\theta,\nu), \quad \nu > -1/2, \end{cases} \tag{31}$$

with the modulation function $\phi(x;\theta,\nu) = y^\nu \frac{\theta^\nu \sqrt{\pi}}{\Gamma(\nu+1/2)}$.

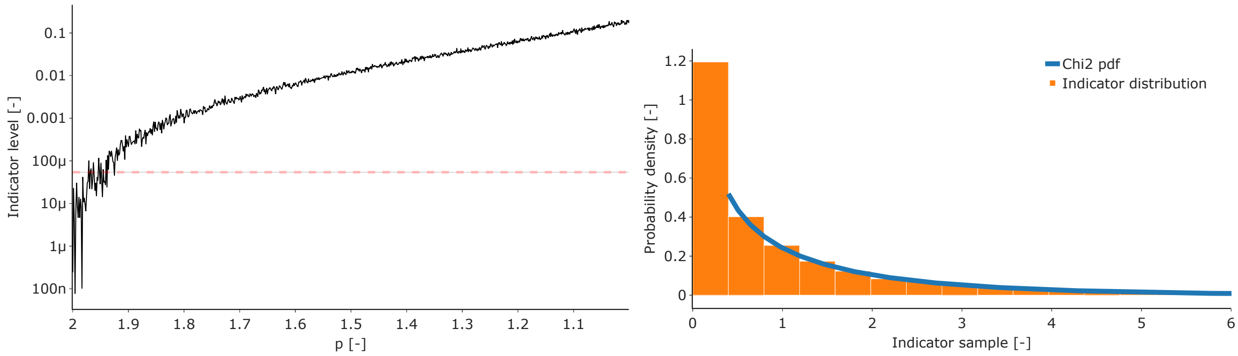
Following theorem (8), the GLRT yields

$$I(\mathbf{x}) = \frac{(\hat{s}_{\mathrm{GA}}(\mathbf{x}) + \ln 2 + \gamma)^2}{2(\pi^2 - 4)}, \tag{32}$$

where γ is the Euler's constant as mentioned above and

$$\hat{s}_{\mathrm{GA}}(\mathbf{x}) = \ln\left(\frac{\left(\prod_{n=0}^{L-1} |x_n|^2\right)^{\frac{1}{L}}}{\langle |x_n|^2 \rangle}\right) = \langle \ln |x_n|^2 \rangle - \ln(\langle |x_n|^2 \rangle), \tag{33}$$

is the log-ratio of the geometric mean to the arithmetic mean of the square envelope x_n^2. This is also known as the smoothness index, applied on the squared envelope.



(a) Kurtosis indicator, threshold with $\alpha = 0.999$ (b) Kurtosis indicator distribution via Monte-Carlo

Figure 1: Statistical features of Kurtosis indicator on simulated signals.

2.4 Proof of concept

The proof of the concept is performed on simulated signals sampled from the generalized Gaussian (GG) distribution. Altering the shape factor p of the GG between 2 and 1, signals are simulated. Decrease in the p from 2 results in more leptokurtic signals. The performances of the recovered HIs are tested on signals with 100000 points and scale $\theta = 1$ of the GG PDF. As mentioned above, the distribution of an HI follows the Chi-squared distribution scaled by twice the signal's length with a given confidence level. Furthermore, the degree of freedom of the Chi-squared is inherently one because the parameter space of H_0 differs from H_1 merely with the tuning parameter. Therefore, as signal length and the confidence level of 0.1% is kept constant, the threshold is the same for all indicators.

Health indicators derived from a Gaussian kernel are expected to exceed the threshold as the shape factor p of the signals slightly deviates from $p = 2$, for which the signal's distribution is Gaussian. The trend of the Kurtosis index is demonstrated in Fig. 1, along with its tuning parameter and distribution. The kurtosis index exceeds the threshold just after the shape factor $p = 1.93$, demonstrating an ever-increasing trend as p decreases. Furthermore, the distribution of the Kurtosis index obtained via Monte-Carlo simulation and the Chi-squared with a single degree of freedom is demonstrated in Fig. 1b. As can be seen, the distribution of the Kurtosis indicator well respects the Chi-squared probability distribution function. The performance of three other indicators is not tested on simulated signals since they demonstrate a similar trend to Kurtosis; nonetheless, they will be tested on experimental signals in the next section.

3 Experimental results

The experimental signals are obtained by Flanders Make using a lab test rig, which performs accelerated run-to-failure tests. The test rig comprises a FAG 6205-C-TVH ball bearing with a Rockwell-C indentation of 300 micrometer in diameter in the inner race of the bearing. The test rigs were operating at a constant speed of 2000 rpm and a radial load of 9 kN. Acceleration signals are sampled at a rate of 50 kHz for one second. These signals are classified as quasi-Gaussian because the maximum likelihood estimate of the shape factor for the *healthy* state of the bearing is approximately 2. The evolution of the shape factor, denoted as p, for the entire measurement set, is depicted in Fig. 2. The value of p remains just below two until the fault becomes pronounced enough to alter the distribution significantly.

The trends of the HIs on experimental signals are illustrated in Fig. 3. It is important to highlight that the thresholds depicted in Fig. 3 are estimated with a 99.9% confidence level, ensuring that the number of HIs falsely rejecting the healthy hypothesis remains below 1 for every 1000 measurements. In Fig. 3, multiple HI values surpass the threshold for the Kurtosis index, ℓ_1/ℓ_2-index, and negentropy of the squared envelope. The percentage of HIs exceeding the threshold for healthy measurements does not exceed 3% for these indicators. On the contrary, the trend of the geometric-to-arithmetic mean aligns with the threshold for all the "healthy" measurements. The reason behind this behavior is uncertain to the authors, but this particular indicator may be less sensitive to signal distribution variations.

Authors also tested the skewness, Jarque-Bera statistics, and derivates of ℓ_β/ℓ_2-index. However, due to the

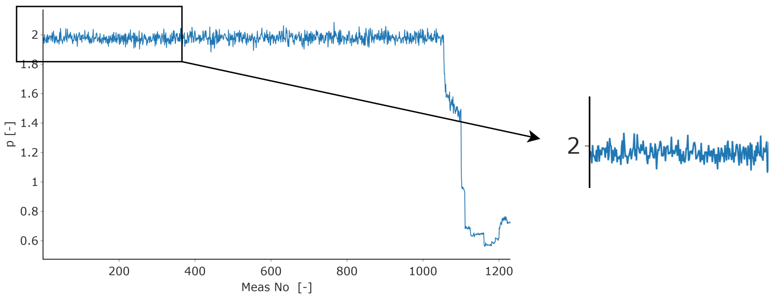
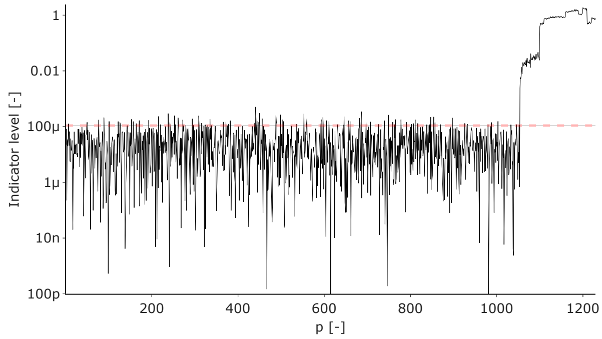
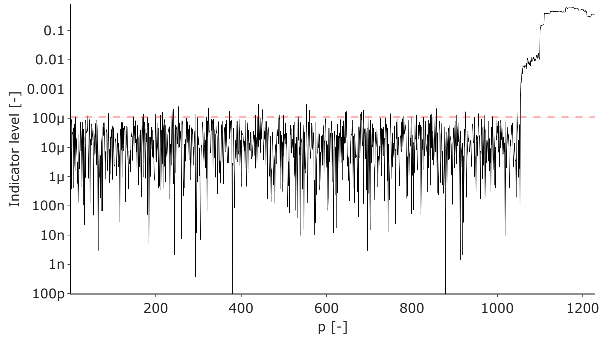


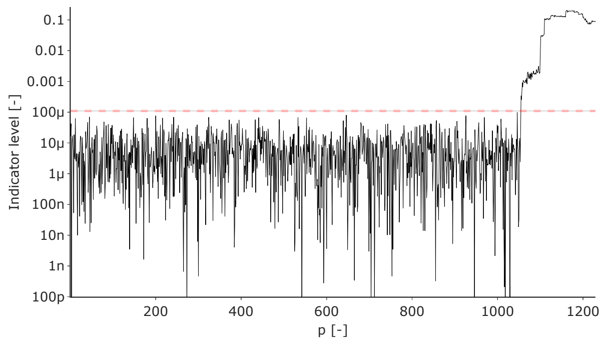
Figure 2: Evolution of shape factors of the quasi-Gaussian experimental signals.



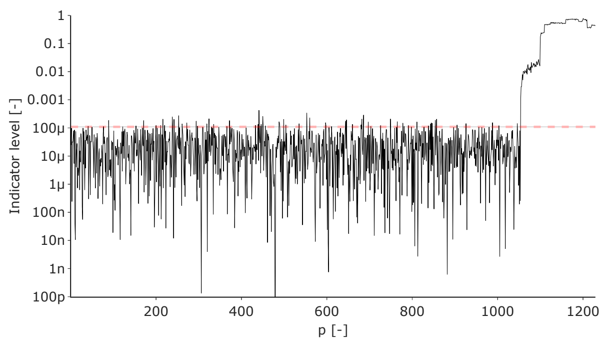
(a) Kurtosis indicator



(b) ℓ_1/ℓ_2 index



(c) Geometric-to-Arithmetic mean



(d) Negentropy of squared envelope

Figure 3: Trends of the HIs along with their statistical threshold on experimental signals. The statistical threshold is estimated at significance level of 0.1%.

page limitations, results are not demonstrated in this text.

4 Conclusion

This study presents a framework for designing health indicators optimized for early detection of faults, along with a statistical threshold for a given risk level. The methodology uses the generalized likelihood ratio test to compare two probability distribution functions representing the healthy and faulty states of the signals. The faulty PDF is derived from the healthy one through Taylor expansions, ensuring they are nested. The approach is designed to detect slight deviations from the distribution of healthy signals. Furthermore, the healthy PDFs are customized to recover classical health indicators, such as kurtosis, negentropy of the squared envelope, and ℓ_q/ℓ_2-norms. The optimality of several classical health indicators is validated concerning a test comparing two PDFs using the proposed approach. The authors claim that most of the decision rules accompanying the recovered indicators are established for the first time in the context of condition monitoring, despite their widespread use.

Both simulated and experimental vibration signals are used to evaluate the proposed approach's detection performance. HIs rooted in the Gaussian null hypothesis are tested on quasi-Gaussian signals. The results show that the trend of health indicators follows the threshold as long as the signals' distribution is similar to

Gaussian. The proposed approach can be used with different probability distribution functions or mixtures.

The formalism laid out in this study can extend the utilization of such indicators in the condition monitoring field by using PDF mixtures that better represent the healthy distribution of signals while being mathematically complex. An extension of the proposed method using GG PDF to construct new HIs may also have the potential to employ on the signals whose healthy state is not purely Gaussian.

Acknowledgements

This research received funding from the Flemish Government under the ”Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” program. The authors would like to acknowledge FWO (Fonds Wetenschappelijk Onderzoek) for their support through the post-doctoral grant of Cedric Peeters (1282221N) and SBO project Robustify (S006119N). The authors also would like to acknowledge the support of SIM SBO SeaFD and the ETF POSEIDON project. They would also like to thank FWO for providing funding (V440622N) for a long stay abroad of Kayacan Kestel at INSA Lyon. The experimental data was provided by Flanders Make, the strategic research centre for the manufacturing industry.

References

[1] Haoxuan Zhou, Xin Huang, Guangrui Wen, Zihao Lei, Shuzhi Dong, Ping Zhang, and Xuefeng Chen. Construction of health indicators for condition monitoring of rotating machinery: A review of the research. *Expert Systems with Applications*, 203:117297, October 2022.

[2] Jérôme Antoni and Pietro Borghesani. A statistical methodology for the design of condition indicators. *Mechanical Systems and Signal Processing*, 114:290–327, January 2019.

[3] Niall Hurley and Scott Rickard. Comparing Measures of Sparsity. *IEEE Trans. Inform. Theory*, 55(10):4723–4741, October 2009.

[4] Jérôme Antoni and R.B. Randall. The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. *Mechanical Systems and Signal Processing*, 20(2):308–331, February 2006.

[5] P. Borghesani, P. Pennacchi, and S. Chatterton. The relationship between kurtosis- and envelope-based indexes for the diagnostic of rolling element bearings. *Mechanical Systems and Signal Processing*, 43(1-2):25–43, February 2014.

[6] R.B.W. Heng and M.J.M. Nor. Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Applied Acoustics*, 53(1-3):211–226, January 1998.

[7] Aziz Kubilay Ovacikli, Patrik Paajarvi, and James P. LeBlanc. Skewness as an objective function for vibration analysis of rolling element bearings. In *2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pages 462–466, Trieste, Italy, September 2013. IEEE.

[8] Carlos M Jarque and Anil K Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259, 1980.

[9] C. Pachaud, R. Salvetat, and C. Fray. CREST FACTOR AND KURTOSIS CONTRIBUTIONS TO IDENTIFY DEFECTS INDUCING PERIODICAL IMPULSIVE FORCES. *Mechanical Systems and Signal Processing*, 11(6):903–916, November 1997.

[10] Joel Igba, Kazem Alemzadeh, Christopher Durugbo, and Egill Thor Eiriksson. Analysing RMS and peak values of vibration signals for condition monitoring of wind turbine gearboxes. *Renewable Energy*, 91:90–106, June 2016.

[11] Xiaodong Jia, Ming Zhao, Yuan Di, Pin Li, and Jay Lee. Sparse filtering with the generalized lp / lq norm and its applications to the condition monitoring of rotating machinery. *Mechanical Systems and Signal Processing*, 102:198–213, March 2018.

[12] Dong Wang. Some further thoughts about spectral kurtosis, spectral L2/L1 norm, spectral smoothness index and spectral Gini index for characterizing repetitive transients. *Mechanical Systems and Signal Processing*, 108:360–368, August 2018.

[13] Bingchang Hou, Dong Wang, Tongtong Yan, and Zhike Peng. A Comparison of Machine Health Indicators Based on the Impulsiveness of Vibration Signals. *Acoust Aust*, 49(2):199–206, June 2021.

[14] I. Soltani Bozchalooi and Ming Liang. A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection. *Journal of Sound and Vibration*, 308(1-2):246–267, November 2007.

[15] Jerome Antoni. The infogram: Entropic evidence of the signature of repetitive transients. *Mechanical Systems and Signal Processing*, 74:73–94, June 2016.

[16] Ralph B D'agostino, Albert Belanger, and Ralph B D'Agostino Jr. A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4):316–321, 1990.

[17] Carlos M. Jarque and Anil K. Bera. A Test for Normality of Observations and Regression Residuals. *International Statistical Review / Revue Internationale de Statistique*, 55(2):163, August 1987.

[18] S.M. Kay. *Fundamentals of Statistical Signal Processing: Detection theory*. Fundamentals of Statistical Si. Prentice-Hall PTR, 1998.

[19] Pietro Borghesani and Md Rifat Shahriar. Cyclostationary analysis with logarithmic variance stabilisation. *Mechanical Systems and Signal Processing*, 70-71:51–72, March 2016.

[20] Jérôme Antoni, Pietro Borghesani, Souyab Kass, Amani Raad, Konstantinos Gryllias, and Flanders Make. Methodologies for designing new condition indicators. In *Proceedings of ISMA 2018-International Conference on Noise and Vibration Engineering and USD 2018-International Conference on Uncertainty in Structural Dynamics*, pages 883–890, 2013.

[21] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley, 1986.