# Subspace-based damage identification of roller bearing

G. Gautier[a], R. Serra[b] and J.-M. Mencik[c]

INSA Centre Val de Loire, Université Francois Rabelais de Tours, LMR EA 2640, Campus de Blois,
3 rue de la chocolaterie, CS 23410, 41034 Blois Cedex, France

**Abstract.** This study presents a frequency-band subspace-based damage identification method for faults diagnosis in roller bearing. The main purpose is to obtain damage sensitive indicators in the frequency range where a damage is expected to appear. The proposed method is validated using the data collected from bearing run-to-failure tests. It is shown that the proposed method can detect bearing defects at an early stage of development and appears therefore as an efficient tool for diagnosis purposes.

## 1. Introduction

Traditional preventative maintenance not only leads to wasteful machine downtime but also premature replacement of mechanical parts. Successfully implementing a condition monitoring programme allows the machine to operate to its full capacity without to need to stop the machine at fixed periods for inspection. In this context, the vibration based structural health monitoring (SHM) process [10] basically involves observation of a structure over time using dynamic response measurements from an array of sensors, extraction of damage sensitive indicators and statistical analysis of these features to determine the current state of structural health [13].

The background of the proposed method is concerned with the concept of damage diagnosis from subspace identification (SubId) methods [4,5]. The vibration response data are periodically collected from the monitored structure and are used to obtain system matrices that are well suited for structural damage identification [12]. Subid methods require the knowledge of the system order, that can increase when damage occurs in a rotating machinery. For industrial applications, where the amount of data to be processed can be huge, the system order selection for each data set can be unsuited.

In order to circumvent these difficulty, a band-pass filtering approach is employed [16]. The original signal is filtered around some expected damage frequencies obtained from given knowledge about rotating machinery components. A frequency can usually be associated with a particular component [2] and a specific subspace based damage indicator can be obtained, that is well suited for fault diagnosis. In this case the order of each subspace damage indicator is known a priori, producing an automatic damage identification procedure. Also, since the identification is focused on a particular frequency band, the global dynamic of the machinery is removed of the procedure, leading to a more sensitive fault identification.

[a] e-mail: guillaume.gautier@insa-cvl.fr
[b] e-mail: roger.serra@insa-cvl.fr
[c] e-mail: jean-mathieu.mencik@insa-cvl.fr

The proposed method is validated using the data collected from bearing run-to-failure tests [1]. A bearing is characterized by four fundamental frequencies [2]. Therefore four dedicated indicators are designed and tracked over time in order to identify faults.

## 2. Subspace based damage identification

Consider a linear state space systems of the form [7,11]:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{v}_k,$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{w}_k, \tag{1}$$

where $\mathbf{x}_k$ is a $2n \times 1$ state vector ($2n$ being the order of the system) defined at the discrete time $t_k$ ($k = 1, 2, \ldots$), with the convention that $t_{k+1} = t_k + \Delta t \; \forall k$ ($\Delta t$ being the time step); $\mathbf{u}_k$ and $\mathbf{y}_k$ are $m \times 1$ and $l \times 1$ vectors of input and output data, respectively, while $\mathbf{v}_k$ and $\mathbf{w}_k$ are $2n \times 1$ and $l \times 1$ vectors of process and measurement noises (respectively), which are supposed to be white Gaussian with zero-mean distributions. Also, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are state-space matrices of respective sizes $2n \times 2n$, $2n \times m$, $l \times 2n$ and $l \times m$.

Within the framework of SubId methods [4,5,7,8], Eq. (1) is rewritten into a linear system with a larger size that involves matrices having interesting features. These are particularly used for removing the influence of inputs and noises. Then, a so-called extended observability matrix $\boldsymbol{\Gamma}$ can be derived which is well suited for the identification of damages since it is directly linked to the dynamic properties of the system [9].

The subspace identification approach makes use of the following matrix equation:

$$\mathbf{Y} = \boldsymbol{\Gamma}\mathbf{X} + \boldsymbol{\Theta}\mathbf{U} + \boldsymbol{\Xi}\mathbf{V} + \mathbf{W}, \tag{2}$$

where **Y** is the output Hankel matrix defined as:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{N-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_N \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}_{\alpha-1} & \mathbf{y}_\alpha & \cdots & \mathbf{y}_{N+\alpha-2} \\ \hline \mathbf{y}_\alpha & \mathbf{y}_{\alpha+1} & \cdots & \mathbf{y}_{N+\alpha-1} \\ \mathbf{y}_{\alpha+1} & \mathbf{y}_{\alpha+2} & \cdots & \mathbf{y}_{N+\alpha} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}_{2\alpha-1} & \mathbf{y}_{2\alpha} & \cdots & \mathbf{y}_{N+2\alpha-2} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_p \\ \hline \mathbf{Y}_f. \end{bmatrix} \quad (3)$$

where the subscripts "p" and "f" stand for past and futur, respectively.

The matrices **U**, **V** and **W** are defined in a similar way. By making use of a projection matrix $\Pi_{\mathbf{U}^T}^\perp = \mathbf{I} - \mathbf{U}^T(\mathbf{U}\mathbf{U}^T)^{-1}\mathbf{U}$, so that $\Pi_{\mathbf{U}^T}^\perp \mathbf{U}^T = \mathbf{0}$, i.e., $(\Pi_{\mathbf{U}^T}^\perp \mathbf{U}^T)^T = \mathbf{U}\Pi_{\mathbf{U}^T}^\perp = \mathbf{0}$ it becomes possible to remove the influence of the inputs.

The strategy for removing the matrices of noises **V** and **W** in equation (10) is to make use of the so-called instrumental variables (IV) approach [7]. The key idea here is to assume that the matrices of noises **V** and **W** are asymptotically uncorrelated with the matrices $\mathbf{U}_p$ and $\mathbf{Y}_p$ in the following sense:

$$\frac{1}{N}(\Xi\mathbf{V}\Pi_{\mathbf{U}^T}^\perp + \mathbf{W}\Pi_{\mathbf{U}^T}^\perp)\mathbf{P}^T \approx \mathbf{0} \quad \text{for } N \text{ large;} \quad (4)$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{U}_p \\ \mathbf{Y}_p \end{bmatrix}. \quad (5)$$

As a result, Eq. (10) leads to

$$\frac{1}{N}\mathbf{\Gamma}\mathbf{X}\Pi_{\mathbf{U}^T}^\perp \mathbf{P}^T \approx \frac{1}{N}\mathbf{Y}\Pi_{\mathbf{U}^T}^\perp \mathbf{P}^T. \quad (6)$$

Equation (6) means that the column space of $\mathbf{\Gamma}$ is almost equal to the space spanned by the left singular vectors of the matrix $\mathbf{Y}\Pi_{\mathbf{U}^T}^\perp \mathbf{P}^T$. This yields

$$\text{ran}(\mathbf{\Gamma}) \approx \text{ran}(\widetilde{\mathbf{Q}}), \quad (7)$$

where $\widetilde{\mathbf{Q}}$ is a $l\alpha \times 2n$ matrix whose columns are the left singular vectors of $\mathbf{Y}\Pi_{\mathbf{U}^T}^\perp \mathbf{P}^T$ that are associated with its $2n$ largest singular values (namely, $\{\tilde{\sigma}_j\}_{j=1,\ldots,2n}$). The order $2n$ is determined by inspecting the amplitudes of normalized singular values as shown in Fig. 1. From Eq. (7), one has

$$\mathbf{\Gamma} \approx \mathbf{\Gamma}^e\mathbf{T} \text{ where } \mathbf{\Gamma}^e = \widetilde{\mathbf{Q}}, \quad (8)$$

where **T** is a $2n \times 2n$ full rank matrix, while $\mathbf{\Gamma}^e$ is the so-called experimental extended observability matrix [14] that is built from output signals.

Let us assume the observability matrix denoted as $\mathbf{\Gamma}_0^e$ which has been determined from one reference undamaged data set (indicated by subscript 0). Consider the observability matrix $\mathbf{\Gamma}_k^e$ corresponding to the $k$th data set. Until a damage occurs, the following property remains valid [6]:

$$\mathbf{\Gamma}_0^e = \mathbf{\Gamma}_k^e\widetilde{\mathbf{T}} + \mathbf{E} \quad (9)$$
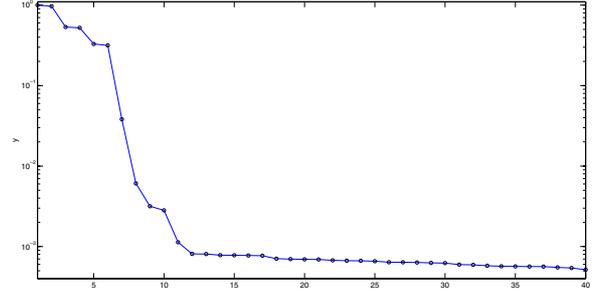


**Figure 1.** Example of the normalized singular values plot for a system of order $2n = 6$.

where **E** is a residue term due to noise effects. The equation (9) can be rewritten as:

$$\mathbf{E} = \left(\mathbf{I} - \mathbf{\Gamma}_k^e\mathbf{\Gamma}_k^{e+}\right)\mathbf{\Gamma}_0^e \quad (10)$$

where $\mathbf{\Gamma}_k^{e+}$ is the pseudo-inverse of $\mathbf{\Gamma}_k^e$. Here, $\left(\mathbf{I} - \mathbf{\Gamma}_k^e\mathbf{\Gamma}_k^{e+}\right)$ can be computed in an efficient numerical manner through a QR decomposition [15].

The relation (10) can be taken as damage sensitive indicator. It is expected that structural damages are detected by an increase in **E**.

The important point is to identify significant changes, i.e. those which can not be attributed to fluctuations in the damage indicator due to noise [3]. In this regard, a damage indicator is defined as:

$$\varepsilon = \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{\Sigma}} \quad (11)$$

where $\mathbf{e} = \text{vec}\{\mathbf{E}\}$ and $\mathbf{\Sigma}$ is a covariance matrix defined as:

$$\mathbf{\Sigma} = \frac{1}{n_{ref}}\sum_{k=1}^{n_{ref}}\mathbf{e}_k^T\mathbf{e}_k \quad (12)$$

where $n_{ref}$ is the number of reference data set (assumed undamaged).

In rotating machinery, vibration signals carry dynamic information from mechanical elements. The energy of vibration is increased when a mechanical element is damaged or worn. For example, the frequencies induced by rolling element bearings, when rollers pass over a surface anomaly, are called fundamental fault frequencies. These frequencies are functions of the bearing geometry (that is number of balls, the contact angle, the pitch diameter and roller diameters) and the relative speed between the two raceways. When bearing geometry (Fig. 2) is known, the fundamental fault frequencies can be calculated using the following equations [2]:

$$\text{BPFI} = \frac{N_b}{2}f_r\left(1 + \frac{B}{P}\cos(\theta)\right) \quad (13)$$

$$\text{BPFO} = \frac{N_b}{2}f_r\left(1 - \frac{B}{P}\cos(\theta)\right) \quad (14)$$

$$\text{FTF} = \frac{f_r}{2}\left(1 - \frac{B}{P}\cos(\theta)\right) \quad (15)$$

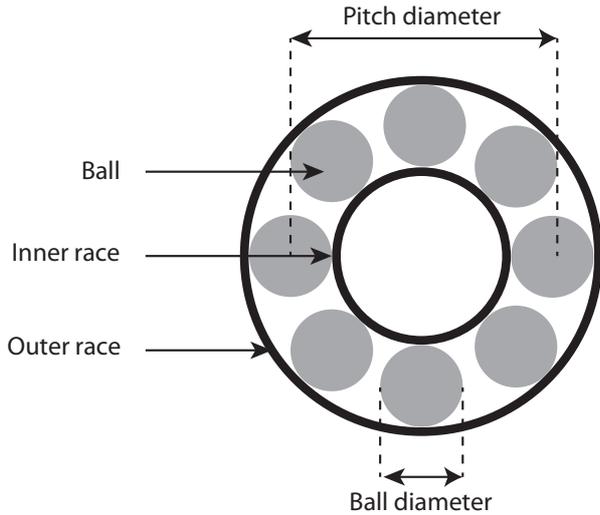**Figure 2.** Schematic representation of a roller bearing.

$$\mathrm{BSF} = \frac{P}{2B} f_r \left[ 1 - \left( \frac{B}{P} \cos(\theta) \right)^2 \right] \qquad (16)$$

where:

- BPFI: ball pass frequency inner race,
- BPFO: ball pass frequency outer race,
- FTF: fundamental train frequency,
- BPF: ball pass frequency,
- $N_b$: number of balls,
- $f_r$: shaft frequency,
- $B$: ball diameter,
- $P$: pitch diameter,
- $\theta$: contact angle.

In order to diagnosis fault in rotating machinery a subspace based fault indicator is assessed for each fundamental fault frequency. This operation is done by band-pass filtering the vibration data [16]. The definition of the band-pass center frequency is set as the value of a fundamental frequency. The filter bandwidth is defined so that the filter does not overlap another filter. In roller bearing case, four filtered data are obtained, associated to the four fundamental frequencies (Fig. 3). Therefore, four indicators are obtained as depicted in Sect. 2, where it is expected that the increase of an indicator reflects the health condition of any associated component.

## 3. Experimental application

The aim of this case study is to detect the occurence of bearing faults. The bearings test rig hosts four test bearing on one shaft [1] ans depicted in Fig. 4. The shaft is driven by an AC motor and coupled by rubber belts. The rotation speed of the shaft is constant at 2000 rpm. A magnetic plug installed in the oil feedback pipe collects debris from the oil as an evidence of bearing degradation.

Four Rexnord ZA-2115 double row bearings are installed on the shaft. The bearings have 16 rollers on each row, a pitch diameter of 71.5 mm, roller diameter of 8.4 mm and a tapered contact angle of 15.17 deg. From
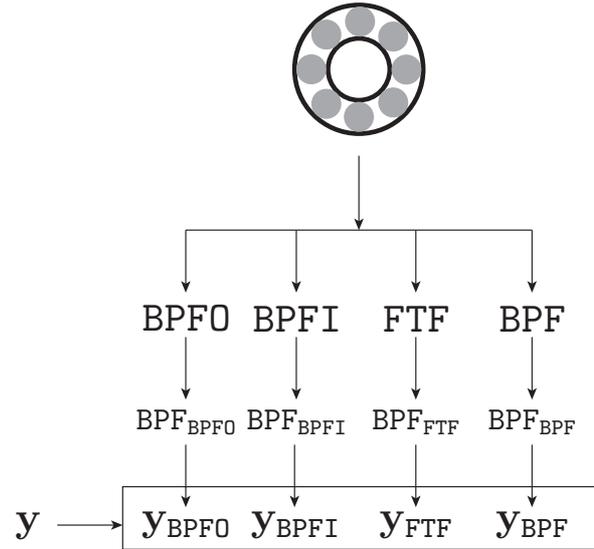


**Figure 3.** Data filtering from the knowledge of bearing characteristics.
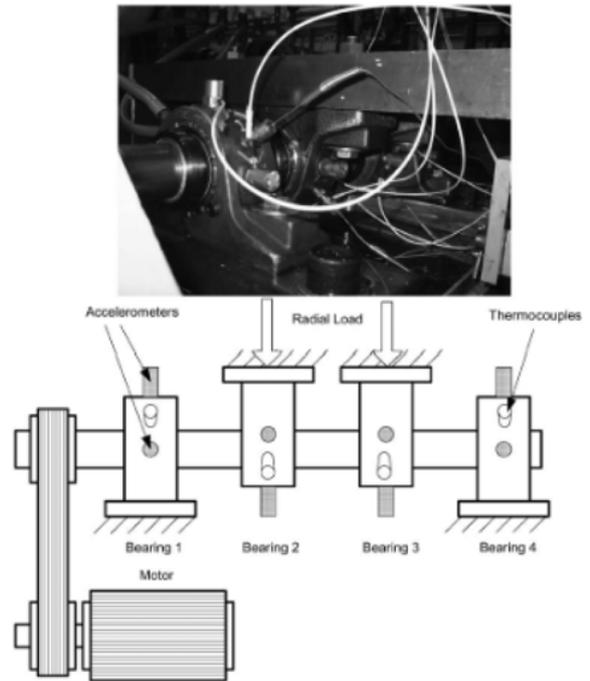


**Figure 4.** Bearing test rig.

Eqs. (13)–(16) the expected damaged frequencies are: BPFI = 296.90 Hz, BPFO = 236.43 Hz, FTF = 14.78 Hz and BPF = 280.08 Hz.

Accelerometers are placed on each bearing housing. Vibration data are collected every 10 minutes with a sampling rate of 20 kHz and a data length of 20480 points.

The test data are recorded during 7 days, until a significant amount of metal debris is found on the magnetic plug of the test bearing. From visual inspection, an outer race defect is discovered in the test bearing 1.

The damage identification is carried out for the four bearing fundamental frequencies with damage indicators
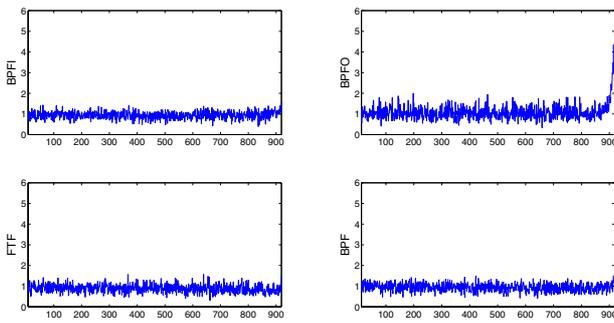
**Figure 5.** Damage indicators for the four bearing fundamental frequencies (`BPFI`, `BPFO`, `FTF` and `BPF`).
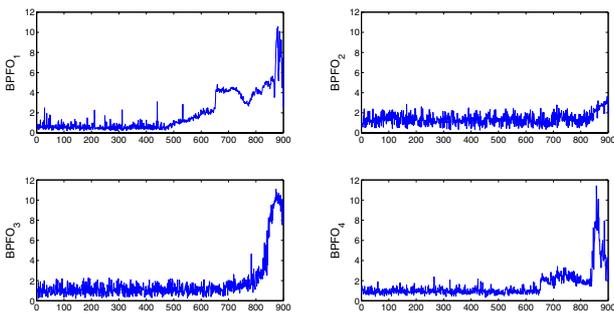


**Figure 6.** `BPFO` indicators for the four accelerometer signals.

obtained from the whole set of accelerometers data. The number of reference data set is $n_{ref} = 50$ and the filter bandwidth is 10 Hz. The observability matrices are determined for $\alpha = 10$ and $2n = 2$. Results of the damage identification are reported in Fig. 5. It can be shown that a damage is identified for the ball pass frequency of the bearing outer race, close to the end of the run-to-failure test (approximatively 3 hours before the experiment interruption). Since the four bearings have the same characteristics, the identification of the damaged bearing is not possible.

This drawback is avoided by assuming that a fault is more visible on an accelerometer close to the damaged bearing. For each accelerometer data (i.e. 1 to 4), a pass-band filter is designed. The damage diagnosis results are reported in Fig. 6. The first bearing reveals damage approximatively 2 days before the experiment interruption. One expects the first bearing to be the most damaged. This result is validated by visual inspection. It can be shown that the `BPFO` indicators for bearings 3 and 4 also increases. There are at least two reasons for this. First, the dynamic of the damaged bearing is recorded by the whole set of accelerometers. In this case, it should be expected that the $BPFO_2$ indicator also increases. The other reason be that bearings 3 and 4 also deteriorate at the end of

the run-to-failure test. It is the most probable, but is not validated by visual inspection.

## 4. Conclusion

A damage identification procedure has been proposed, which make use of subspace identification and band-pass data filtering. A bank of indicators has been proposed, which can serve for the diagnosis of specific rotating machinery faults. The present methodology has been applied for roller bearing damage identification. It is shown that a bearing outer race defect can be identified at an early stage.

## References

[1] H. Qiu and J. Lee and J. Lin and G. Yu, Journal of Sound and Vibration **289**, 1066–1090 (2006)
[2] B.-P. Graney and K. Starry, Materials Evaluation **70(1)**, 78–85 (2011)
[3] C.P. Fritzen, G. Mengelkamp, A. Guemes, Structural Health Monitoring, 1530–1538 (2003)
[4] M. Verhaegen, Automatica, **30(1)**, 61–74 (1994)
[5] P. Van Overschee and B. De Moor, Automatica, **30(1)**, 75–93 (1994)
[6] A. Swindlehurst and R. Roy and B. Ottersten and T. Kailath, IEEE Transactions on automatic control, **40(2)**, 311–316
[7] M. Viberg and B. Wahlberg and B. Ottersten, Automatica, **33(9)**, 1603–1616 (1997)
[8] S J. Qin, Computers and Chemical Engineering, **30**, 1502–1513 (2006)
[9] K F. Alvin and A N. Robertson and G W. Reich and K C. Park, Computers and Structures, **81**, 1149–1176 (2003)
[10] S W. Doebling and C R. Farrar and M B. Prime and D W. Shevitz (Los Alamos National Lab, 1996)
[11] M. Raffy and C. Gontier, Mechanical Systems and Signal Processing, **19(4)**, 714–735 (2005)
[12] N. Viet-Haand J.-C. Golinval, Mechanical Systems and Signal Processing, **24(6)**, 1831–1843 (2010)
[13] A.-M. Yan and J.-C. Golinval, Mechanical Systems and Signal Processing, **20**, 611–626 (2006)
[14] G. Gautier and R. Serra and J.-M. Mencik, Mechanics & Industry, **14(2)**, 145–149 (2013)
[15] D.-M. Sima and S. Van Huffel, Journal of Computational and Applied Mathematics, **203(1)**, 264–278 (2007)
[16] S.Y Nalakath and K. Saeed and K.-A. Alnefaie, International Journal of Emerging Technology and Advanced Engineering, **4(8)** (2014)