Simultaneous signal separation and prognostics of multi-component systems: the case of identical components

Li Hao, Nagi Gebraeel & Jianjun Shi

To cite this article: Li Hao, Nagi Gebraeel & Jianjun Shi (2015) Simultaneous signal separation and prognostics of multi-component systems: the case of identical components, IIE Transactions, 47:5, 487-504, DOI: 10.1080/0740817X.2014.955357

To link to this article: https://doi.org/10.1080/0740817X.2014.955357

Accepted author version posted online: 28 Oct 2014.
Published online: 23 Feb 2015.

Submit your article to this journal

Article views: 293

View Crossmark data

Citing articles: 7 View citing articles
Simultaneous signal separation and prognostics of multi-component systems: the case of identical components

LI HAO, NAGI GEBRAEEL* and JIANJUN SHI

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
E-mail: nagi@isye.gatech.edu

Received August 2012 and accepted May 2014

When monitoring complex engineering systems, sensors often measure mixtures of signals that are unique to individual components (component signals). However, isolating component signals directly from sensor signals can be a challenge. As an example, in vibration monitoring of a rotating machine, if different components generate vibration signals at similar frequencies, they cannot be distinguished using traditional spectrum analysis (non-inseparable). However, developing degradation signals from component signals is important to monitor the deterioration of crucial components and to predict their residual lifetimes. This article proposes a simultaneous signal separation and prognostics framework for multi-component systems with non-inseparable component signals. In the signal separation stage, an Independent Component Analysis (ICA) algorithm is used to isolate component signals from mixed sensor signals, and an online amplitude recovery procedure is used to recover amplitude information that is lost after applying the ICA. In the prognostics stage, an adaptive prognostics method to model component degradation signals as continuous stochastic processes is used to predict the residual lifetimes of individual components. A case study is presented that investigates the performance of the signal separation stage and that of the final residual-life prediction under different conditions. The simulation results show a reasonable robustness of the methodology.

Keywords: Multi-component systems, vibration monitoring, independent component analysis, degradation modeling, residual-life prediction, prognostics

1. Introduction

The condition monitoring of engineering systems involves collecting sensor signals, such as temperature, vibration, crack propagation, etc., to enable fault detection (diagnostics) and facilitate remaining lifetime predictions (prognostics). One of the most popular condition monitoring techniques is vibration monitoring, which is well suited for many applications, such as machine tools (Jantunen, 2002), power transformers (Booth and McDonald, 1998), engines (Basir and Yuan, 2007), electronic motors (Nandi et al., 2005), wind turbines (Caselitz and Giebhardt, 2005), and even structural systems (Fan and Qiao, 2011). Depending on the frequency range and other system characteristics, vibration can be measured in displacement, velocity, or acceleration. Typically, a sensor measures a mixture of vibration signals generated by the components of the system.

This mixture can be transformed using signal processing techniques, such as the Discrete Fourier Transformation (DFT), into a spectrum of individual frequencies that are related to individual components. Some of these frequencies are generated only when specific faults or component defects occur. Examples of defects include imbalance, misalignment, bearing defects, gear defects, etc. The frequency associated with a specific defect is referred to as the “defective frequency” in this article.

It is not uncommon to observe a correlation between the amplitude of the defective frequency and the severity of the defect that is generating that frequency. In fact, there are several examples in the literature where amplitudes of defective frequencies have been used to develop degradation signals for predicting failure times (Baruah and Chinnam, 2005; Gebraeel et al., 2005; Huang et al., 2007). In this article, we are interested in modeling the degradation of systems that consist of two or more identical components operating under similar conditions; for example, two identical bearings mounted on the same shaft. The key challenge here is that a defect in any of the components will

*Corresponding author
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uiie.
excite the same defective frequency. Without having dedicated sensors uniquely located near each component, it is not possible to distinguish between the defective and non-defective components simply from the spectrum analysis of sensor vibration signals. In this case, sensor vibration signals are considered as “inseparable” signals under traditional signal processing techniques. In reality, physical or structural constraints often create strong limitations for installing sensors close to each component of interest, thus presenting significant challenges in modeling degradation processes.

This article proposes a two-stage vibration-based prognostic methodology for modeling the degradation processes of components with identical defective frequencies. The first stage uses Independent Component Analysis (ICA) to separate component vibration signals from sensor vibration signals. ICA is a Blind Signal Separation (BSS) technique that is used to separate mixtures of signals without necessarily relying on information about component signals themselves or the process by which the signals are mixed (Cardoso and Souloumiac, 1993; Cardoso, 1998; Hyvärinen et al., 2001). In our model, we assume that the signal measured by each sensor is a mixture of the vibration signals generated by individual components. We develop a novel approach that relies on signal transformation and ICA to separate the “inseparable mixtures” of component signals; i.e., vibration signals with the same frequency. Once the signals are separated, they are used to construct the degradation signal for each component.

The second stage models the degradation signals for the purpose of predicting the remaining lifetime of each component. The degradation signals are characterized by incorporating an adaptive degradation-based prognostic model similar to the one developed in Gebraeel et al. (2005). Parameters of the degradation model are assumed to be random to account for the variability between the degradation processes of “identical” components. The random parameters are updated using real-time signals observed from the field in a Bayesian manner. Then, the updated models are used to compute the revised Remaining Life Distributions (RLDs), which is the distribution of the time that a component’s degradation signal is expected to cross a pre-determined failure threshold.

The remainder of the article is organized as follows: Section 2 provides a literature review on the state-of-the-art pertaining to the topic of this research. Section 3 discusses the model and defines the mixing relationship between components and sensor signals. Section 4 describes the signal separation procedure and the challenges of applying ICA. We also discuss an amplitude recovery procedure to approximately recover signal amplitude. Section 5 reviews the adaptive prognostics model and discusses how it fits within our framework. Section 6 focuses on a numerical study that investigates the effectiveness of signal separation and the accuracy of residual-life prediction. Finally, conclusions are drawn.

2. Literature review

The literature on the condition monitoring of multi-component systems has addressed several application domains, such as power transformers (Booth and McDonald, 1998; Huang and Huang, 2002), engines (Basir and Yuan, 2007), motors (Nandi et al., 2005), steam turbines (Beebe, 2003), wind turbines (Caselitz and Giebhardt, 2005), gearboxes (He et al., 2007), manufacturing processes (Lee et al., 2006), and even structural systems (Fan and Qiao, 2011). Traditionally, the condition monitoring literature has relied on two key assumptions: (i) sensors can be located on, or very close to, the component that is being monitored (Gebraeel et al., 2005; Huang et al., 2007) and (ii) components of a system are different and generate different defective frequencies that can easily be identified through spectral analysis (Rao, 1996; Gelle and Colas, 2001). These assumptions do not necessarily hold for many multi-component engineering systems. In reality, a system may consist of multiple identical components, and sensor placement can be constrained by the physical and/or functional limitations. In such systems, a sensor signal will consist of a mixture of multiple vibration signals that have same defective frequency. As a result, these signals cannot be separated using conventional signal processing techniques.

In this article, we focus on developing a prognostic degradation modeling approach that utilizes vibration data to develop degradation signals that can be modeled to predict the remaining lifetime. The literature on degradation modeling is relatively rich and has relied on numerous statistical as well as stochastic models for characterizing degradation. Examples of such work include nonlinear random-coefficient models developed by Lu and Meeker (1993) to model degradation and estimate time-to-failure distributions. Other papers have modeled degradation using the Wiener process (Doksum and Hoyland, 1992; Gebraeel et al., 2005; Wang, 2010). For example, Doksum and Hoyland (1992) proposed a time scale transformation of the Wiener process to model the accumulated decay of a component under variable stress levels. Wang (2010) incorporated random effects into the Wiener process to represent unit-to-unit variability among the degradation processes of different individuals. Most of these efforts focused primarily on estimating the lifetime distribution for a population of components. Gebraeel et al. (2005) developed two stochastic degradation models where real-time data from components operating in the field were used to update the model and estimate RLDs for partially degraded components.

It is important to note that most of the degradation modeling literature has focused uniquely on modeling the degradation of single components. There is very limited literature that addresses the degradation modeling in the context of multi-component systems and almost none that tackles the problem of inseparable mixtures of vibration signals. As mentioned earlier, our modeling approach involves
a signal separation stage that is based on ICA. ICA assumes that the signals measured by a sensor are a linear combination of statistically independent signals unique to individual signal sources, which in our context are assumed to be the components of the system. ICA separates the signals by maximizing their independence. Different ICA algorithms include FastICA (Hyvärinen and Oja, 1997), AMUSE (Tong et al., 1990), SOBI (Belouchrani et al., 1997), and JADE (Cardoso and Souloumiac, 1993). Other BSS algorithms include the Blind Deconvolution (BD), which models the relationship between sensor signals and component signals as a convolution (Thi and Jutten, 1995; Smaragdis, 1998). ICA and other BSS techniques have also been widely used for diagnostics and fault identification. For example, ICA has been applied on the detection of sensor drifts (Kermit and Tomic, 2003), the identification of variation patterns in a multi-stage manufacturing process (Apley and Lee, 2003), the investigation of the sources of noise in a diesel engine (Li et al., 2001), and the detection of gear faults (Roan et al., 2002). In addition, BD has been applied in the diagnosis of vibration systems (Gelle et al., 2000; Zhang et al., 2009).

One of the major limitations of BSS techniques, including ICA and BD, is that these methods do not retain any information pertaining to the absolute amplitude of the signal (Hyvärinen, 2001). That is, the relative relationship between the amplitudes of any two successive signals is preserved but the actual value is not, as will be shown later. As a result, ICA has not been widely used in prognostics or predictive degradation modeling since almost all prognostic models require the exact signal amplitude in order to predict lifetime and remaining lifetime. In Section 4, we will discuss how to overcome this limitation by proposing an amplitude recovery procedure.

3. System description and modeling approach

This article is based on the vibration monitoring of multicomponent systems. Typically in vibration monitoring, defective frequencies are only generated as a result of the initiation or existence of mechanical faults and defects. Similar to the assumptions made in Gebrael et al. (2005), we assume that the amplitude of a defective frequency is correlated with the severity of the underlying degradation process. The term “degradation signal” is used to define the behavior of the amplitude of the defective frequency over a component’s lifetime.

In this framework, we focus on the degradation phase that begins at the very first onset of a defect and ends at the time of failure. We consider a system that, among other constituent components, contains at least \( n \geq 2 \) identical components operating under the same conditions. An example is a wind turbine that has two or more identical bearings rotating at the same speed or two or more pairs of gears with the same mesh frequency. Consequently, a defect in any of these identical components will generate the same defective frequency, which we denote as \( f^* \). This implies that the signal measured by each sensor will consist of an in-separable mixture of component-specific vibration signals. Thus, conventional methods used for spectral analysis cannot be used to distinguish between the degradation levels of the components in question. Furthermore, we assume that vibration is measured periodically using \( m \geq n \) sensors, typically accelerometers. We consider applications where sensors cannot be placed in close proximity to the degrading components due to physical and/or functional constraints. Instead, sensors can only be placed on an external structure; for example, the outside of a gearbox. As a result, each sensor captures only a proportion of the vibration signals generated by each component. This proportion will depend on sensor location, material properties, damping, etc.

Our goal is to develop a prognostic methodology for modeling degradation and predicting the remaining lifetimes of partially degraded components with an identical defective frequency. To achieve this, we develop a two-stage methodology that involves a signal separation stage followed by degradation modeling and remaining life prediction. In the first stage, ICA is used to separate sensor signals into individual component-specific signals. The resulting component signals are then used to predict the remaining lifetimes of their corresponding components using the degradation-based prognostic modeling approach developed by Gebrael et al. (2005).

3.1. Problem formulation

We start by defining \( x_i(t_k) \) as the time domain vibration signal measured by sensor \( i \) at some observation time \( t_k \) (hereafter referred to as the “sensor vibration signal”), for \( i = 1, 2, \ldots, m \) and \( k = 1, 2, \ldots, M \), where \( t_M \) represents the last observation time. We assume that \( x_i(t_k) \) is a linear combination of the vibration signals generated by each component \( j \) (hereafter referred to as the “component vibration signal”), which we denote as \( s_j(t_k) \). Note that in this context, the linearity assumption is an approximation of the real mechanism by which sensors capture the vibration signals generated by the components. In reality, this mechanism is relatively complex and indeed more complicated than a simple linear model. However, there is a rich literature that has utilized the linearity assumption for similar purposes, especially for systems that are small in size and have a high rigidity. Examples include Li et al. (2001), Roan et al. (2002), Ypma et al. (2002), He et al. (2007), and Zhou and Chelidze (2007). An example presented in Roan et al. (2002) considered vibration monitoring of gearboxes and argued that the linear assumption can hold if the accelerometers are mounted on rigid gearboxes. In this article, we consider a similar setting, and we focus on applications that involve vibration monitoring of mechanical systems with similar properties; i.e., applications where the linearity assumption can be valid. Thus, a sensor vibration
signal $x_i(t_k)$ can be expressed as follows:

$$x_i(t_k) = a_{i,1}s_1(t_k) + a_{i,2}s_2(t_k) + \ldots + a_{i,n}s_n(t_k) + e_i(t_k),$$

for $k = 1, 2, \ldots, M,$

where $0 \leq a_{i,j} \leq 1 (j = 1, \ldots, n)$ is a real-valued element that represents the proportion of the vibration generated by component $j$ that is captured by sensor $i$. The constraint on $a_{i,j}$ implies that each sensor can capture no more than 100% of the vibration generated by an individual component. The term $e_i(t_k)$ represents the sensor noise, which is assumed to be Gaussian.

For a system with $m$ sensors and $n$ identical components, the relationship between sensor vibration signals and component vibration signals can be expressed in matrix form as follows:

$$x(t_k) = As(t_k) + e(t_k) \quad \text{for} \quad k = 1, 2, \ldots, M,$$ \hspace{1cm} (2)

where $x(t_k) = [x_1(t_k); x_2(t_k); \ldots; x_m(t_k)], \quad s(t_k) = [s_1(t_k); s_2(t_k); \ldots; s_n(t_k)], \quad$ and $e(t_k) = [e_1(t_k); e_2(t_k); \ldots; e_m(t_k)]$ represent the vectors of sensor vibration signals, component vibration signals, and sensor noise, respectively. The matrix $A = \{a_{i,j}\}, (i = 1, \ldots, m, j = 1, \ldots, n)$ defines the mixing process. The assumption that $m \geq n$ guarantees that $A$ is full column rank. This is necessary and sufficient for capturing the vibration signals generated by all $n$ identical components. Note that the special case, $m = n$, can be used to describe the scenario where a dedicated sensor is used to monitor each component.

### 3.2. Pre-processing using DFT

Equation (2) shows the relationship of the time domain vibration signals. In order to model the degradation of the $n$ identical components considered in this setting, we need to find their amplitudes of the defective frequency $f^*$. The amplitude of the defective frequency can be used as a proxy for the underlying physical degradation and can therefore be used to estimate the residual lifetime of the corresponding component (Gebraeel et al., 2005). For this purpose, the DFT is used to extract the frequency content of the time domain signals.

The DFT converts a signal into a set of complex sinusoids that are ordered by their frequencies. Thus, the DFT of component vibration signal $s_j(t_k)$ at defective frequency $f^*$ can be expressed in the complex form $s_j^R(t_k) + js_j^I(t_k)$, where $s_j^R(t_k)$ and $s_j^I(t_k)$ are the real and imaginary parts of the DFT of $s_j(t_k)$ at frequency $f^*$, respectively. Similarly, we denote the DFT of the vectors $s(t_k), x(t_k),$ and $e(t_k)$ in Equation (2) at $f^*$ by complex vectors $s^R_R(t_k) + js^I_R(t_k)$, $x^R(t_k) + jx^I(t_k)$, and $e^R(t_k) + je^I(t_k)$, respectively. Since DFT is a linear transformation, the linear relationship in Equation (2) remains valid. The model for the transformed frequency domain signals is therefore given as

$$\begin{align*}
\begin{bmatrix}
x^R_R(t_k) + jx^I_R(t_k)
\end{bmatrix} &= A \begin{bmatrix}
s^R_R(t_k) + js^I_R(t_k)
\end{bmatrix} \\
+ \begin{bmatrix}
e^R(t_k) + je^I(t_k)
\end{bmatrix} & \quad \text{for} \quad k = 1, 2, \ldots, M.
\end{align*} \hspace{1cm} (3)
$$

The amplitude of component vibration signal $s_j(t_k)$ at the defective frequency $f^*$ is given by $\sqrt{|s_j^R(t_k)|^2 + |s_j^I(t_k)|^2}$. We denote this amplitude as $s_j^R(t_k)$. As mentioned earlier, this amplitude represents the severity of degradation at time $t_k$. Therefore, the sequence $(s_j^R(t_1), \ldots, s_j^R(t_M))$ can be used to construct the degradation signal unique to the component (“component degradation signal” for short).

Since $A$ is real-valued matrix, Equation (3) can be separated into two partitions:

$$\begin{align*}
x^R_R(t_k) &= As^R_R(t_k) + e^R(t_k) \quad \text{for} \quad k = 1, 2, \ldots, M \\
x^I_R(t_k) &= As^I_R(t_k) + e^I(t_k) \quad \text{for} \quad k = 1, 2, \ldots, M.
\end{align*} \hspace{1cm} (4)
$$

Both the real part $x^R_R(t_k)$ and the imaginary part $x^I_R(t_k)$ can be derived by applying the DFT on the observed sensor vibration signal $x(t_k)$. Thus, we can estimate the mixing matrix $A$ using either the observations of real parts $x^R_R(t_k), \ldots, x^R_R(t_M)$ or the observations of imaginary parts $x^I_R(t_k), \ldots, x^I_R(t_M)$. After the matrix $A$ is estimated, component vibration signals can be calculated using linear regression. In this article, we focus only on using the real parts. We refer to $x^R_R(t_k), s^R_R(t_k)$, and $e^R(t_k)$ as the vectors of “real sensor signal,” “real component signal,” and “real sensor noise,” respectively.

Particularly, we assume that the phase angle of any component vibration signal, given by $\arctan \frac{s^I(t_k)}{s^R(t_k)}$, remains relatively constant for any point in time $t_k$. The reason is that the phase angle of a defective vibration signal is determined by the location of the defect on the component as well as the steadiness of its rotating speed. We assume that (i) once the defect occurs, the location of the defect on one component cannot be altered by external environment such as the degradation of other components and (ii) the rotation speed of the system is controlled to be steady. Thus, our assumption of constant phase angle holds. Consequently, $\frac{s^I_R(t_k)}{s^R_R(t_k)} \propto \frac{s^I_R(t_k)}{s^R_R(t_k)}$, and, as a result, $s^I_R(t_k) \propto s^I_R(t_k)$ for any time $t_k$. This represents that the real component signal preserves the shape of its corresponding component degradation signal and can be considered as an indicator of the severity of component degradation.

### 4. Signal separation using ICA

As we discussed earlier, the first step of our methodology involves signal separation. Specifically, we apply ICA on the real sensor signals in Equation (4) to estimate the matrix $A$. ICA focuses on separating statistically independent
signals from these linear mixtures without requiring specific knowledge of the mixing process. To guarantee the unique decomposition of the signal mixtures into statistically independent signals, ICA assumes that at most one of the statistically independent signals follows a Gaussian distribution and that the noise variables all follow Gaussian distributions. For detailed explanation, the reader may refer to Hyvärinen et al. (2001).

However, there are some key challenges associated with applying ICA to our problem setting. The first challenge is that ICA assumes that the linear mixture consists of statistically independent signals unique to individual components. However, the vector of real component signals \( s^R_j(t_k) \) in Equation (4) may not necessarily satisfy this requirement. In fact, its elements will always exhibit some level of correlation. To see this, consider the fact that the amplitude of a degradation signal tends to increase as the severity of physical degradation increases. In other words, the sequence of real component signals \( [s^R_j(t_1), \ldots, s^R_j(t_M)] \) of component \( j \) will exhibit an increasing trend. If multiple (identical) components begin to degrade, it is clear that their corresponding degradation signals will possess a significant level of correlation due to a common increasing trend.

To eliminate this correlation, we remove the common increasing trend by focusing on modeling the increments of the signals instead of the actual signals themselves. We let \( \Delta^f(t_k) = s^R_j(t_{k+1}) - s^R_j(t_k) \) represent the increments of the real component signals at time \( t_k \) (“component increments” for short) and define the sequence \( [\Delta^f(t_1), \ldots, \Delta^f(t_M)] \) as independent and identically distributed samples of a random vector \( \Delta S^f \). Then, we assume that all elements of \( \Delta S^f \) are mutually independent. This assumption means that the variations in the degradation signal among the individual components are independent of one another. Similarly, we define \( \Delta X^f(t_k) \) and \( \Delta E^f(t_k) \) as the sensor increments and error increments, respectively, with corresponding random vectors \( \Delta X^f \) and \( \Delta E^f \). The model in Equation (4) can now be re-expressed in terms of random vectors as follows:

\[
\Delta X^f = A \Delta S^f + \Delta E^f.
\]  

(5)

The second challenge with applying ICA to our problem setting is that it does not preserve amplitude information. To explain this, consider the following expansion of Equation (5):

\[
\Delta X^f = \sum_{j=1}^n a_j b_j \Delta S^f_j + \Delta E^f,
\]

where \( a_j \) is the \( j \)th column of \( A \), \( \Delta S^f_j \) is the \( j \)th component increment, and \( b_j \) is an arbitrary scalar. The linear combination as well as the assumption of independent component increments hold for any value of \( b_j \). Therefore, ICA cannot be used to determine the true amplitude of \( \Delta S^f_j \) or the true norm of the column \( a_j \) \( (j = 1, 2, \ldots, n) \). To address this challenge we develop an online amplitude recovery procedure for approximately recovering the amplitude content of the estimated degradation signals (as will be shown later in Section 4.2).

4.1. Applying ICA

ICA is applied to a historical training data set consisting of sensor increments with the goal of estimating the mixing matrix \( A \). To guarantee a unique solution, ICA typically assumes unit-variance of the source signals, which in our modeling framework are represented by the component increments \( \Delta S^f \). Consequently, we let \( \Sigma_S \) denote the covariance matrix of \( \Delta S^f \) and define \( \Delta \tilde{S}^f = \Sigma_S^{-\frac{1}{2}} \Delta S^f \), where \( \Delta \tilde{S}^f \) represents a random vector of signals increments with unit-variance and independent elements. Using this representation, we can now express Equation (5) as follows:

\[
\Delta X^f = \tilde{A} \Delta \tilde{S}^f + \Delta E^f
\]

\[
= \tilde{A} \Delta \tilde{S}^f + \Delta E^f,
\]  

(6)

where \( \tilde{A} = \Sigma_S^{-\frac{1}{2}} A \).

Our objective now is to use ICA to estimate the matrix \( \tilde{A} \). We begin by performing an eigen-decomposition of the covariance matrix of \( \Delta X^f \), which we denote by \( \Sigma_X \). Since \( \Delta \tilde{S}^f \) has unit-variance, \( \Sigma_X \) can be expressed as \( \Sigma_X = \tilde{A} \tilde{A}^T + \Sigma_E \), where \( \Sigma_E \) represents the covariance matrix of \( \Delta E^f \). If we assume that all elements in \( \Delta E^f \) are mutually independent and have equal variance \( \sigma_E^2 \), then \( \Sigma_E \) can be expressed as \( \Sigma_E = \sigma_E^2 I \). Consequently, the eigenvector matrix of \( \Sigma_X \), denoted as \( P \), is equal to that of \( \tilde{A} \tilde{A}^T \), and the eigenvalues of \( \Sigma_X \), denoted as \( \lambda_i \) \( (i = 1, 2, \ldots, m) \), are the eigenvalues of \( \tilde{A} \tilde{A}^T + \sigma_E^2 I \). For notation convenience, we denote the diagonal matrix that has \( \lambda_i, i = 1, 2, \ldots, m \) on the diagonal as \( \Lambda_X \). For the eigenvalues of \( \Lambda_X \) on the diagonal as \( \Lambda \), thus, \( \Lambda = \Lambda_X - \sigma_E^2 I \). Consequently, according to the definition of eigen-decomposition, \( \Lambda \tilde{A}^T = P \Lambda_X P^T \), which implies that \( \tilde{A} = P \Lambda \Lambda^{-\frac{1}{2}} Q \), where \( Q \) is an unknown orthogonal matrix.

The value of \( \sigma_E^2 \) is usually unknown and can be estimated from data. To do this, we rely on the fact that \( \Lambda \) is an \( m \)-by-\( n \) matrix with full column rank, which means that the first \((m-n)\) smallest eigenvalues of \( \Lambda \tilde{A}^T \) are equal to zero. Thus, by ranking the eigenvalues of \( \Sigma_X \), we have \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > \sigma_E^2 = \lambda_{n+1} = \lambda_{n+2} = \cdots = \lambda_m \). Therefore, \( \sigma_E^2 \) can be estimated by taking the average of \( \lambda_{n+1}, \lambda_{n+2}, \ldots, \lambda_m \).

The next step involves optimization where the objective is to estimate the orthogonal matrix \( Q \). For every feasible value of \( Q \), there will be a corresponding estimated component increments calculated based on Equation (6). Since the true component increments are statistically independent,
the optimal $Q$ will correspond to the estimated component increments that exhibit maximum independence. Usually independence is measured using fourth-order statistics, such as the fourth-order cumulants (Cardoso and Souloumiac, 1993) and kurtosis (Hyvärinen and Oja, 1997). In this article, we utilize an optimization algorithm based on the theory of the fourth-order cumulants known as “JADE” (proposed by Cardoso and Souloumiac (1993)).

For any random vector $w$, its fourth-order cumulants of the $p$th, $q$th, $u$th, and $v$th elements is defined as

$$C_{p,q,u,v}(w) = E[w_p w_q w_u w_v] - E[w_p w_q]E[w_u w_v] - E[w_p w_u]E[w_q w_v] - E[w_p w_v]E[w_q w_u].$$

According to the properties of fourth-order cumulants stated in Nikias and Petropulu (1993), if $w$ is a random vector with statistically independent elements, all of its cross-cumulants ($p, q, u, v \neq p, p, p, p$) are equal to zero. Thus, maximizing independence between estimated component increments is equivalent to minimizing all cross-cumulants between the estimated component increments. Using the JADE algorithm, $Q$ is estimated by minimizing the sum of squares of a subset of all cross-cumulants of estimated component increments. We denote the estimated value of $Q$ as $\hat{Q}$. Subsequently, the estimator of $A$, denoted as $\hat{A}$, can be estimated using $P \Lambda_h^{-1} \hat{Q}$.

Due to the loss of amplitude after ICA stated earlier, the estimated matrix $\hat{A}$ does not contain any amplitude information. Thus, if we want to estimate $\hat{s}_R^f(t_k)$ using $\hat{A}$ based on Equation (4), the least-squares estimation result, denoted by $\tilde{s}_R^f(t_k) = [\hat{A}^T \hat{A}]^{-1} \hat{A}^T \hat{s}_R^f(t_k)$ for $k = 1, 2, \ldots, M$, will not preserve the amplitude information either. Instead, it will only preserve the shape of $s_R^f(t_k)$. As mentioned earlier, $s_R^f(t_k)$ is an indicator of the component degradation; thus, $\tilde{s}^f$ preserves the shapes of component degradation signals but does not reflect the actual degradation levels of individual components. In other words, for any component $j$, the amplitude of $\tilde{s}_R^f(t_k)$ is proportional to the amplitude of component degradation signal $j$ with an unknown ratio. Consequently, we refer to $\tilde{s}_R^f(t_k)$ as the vector of “unscaled” degradation signals. However, in the application of prognostics, estimating the actual amplitude of a component degradation signal is crucial to predict its remaining useful lifetimes. Consequently, in what follows, we propose a data-driven “amplitude recovery procedure” capable of approximately recovering the amplitude information of component degradation signals.

### 4.2. Online amplitude recovery procedure

The amplitude recovery procedure is used for online estimation of the true amplitude level of component degradation signals in systems operating in the field. Our procedure is developed on the premise that individual components in the system experience what we define as a “two-phase” degradation pattern; i.e., a non-defective phase of operation followed by a phase of gradual degradation until failure. Note that this degradation pattern is typical in many mechanical applications. One example was presented by Gebraeel et al. (2005), where the authors considered the degradation of bearings and identified a similar “two-phase” pattern shown in Fig. 1. Recall that the amplitude of the defective frequency $f^*$ is typically close to zero as long as no defect exists. Thus, during the non-defective phase of an individual component, the amplitude of its degradation signal computed at $f^*$ is typically close to zero. Once a defect occurs, it marks the beginning of the subsequent degradation phase and is typically accompanied by a sudden spike in amplitude. This spike is considered as the “initial degradation level” of a component. Since we focus on the degradation of identical components, we assume that the initial degradation levels of identical opponents are equal to the same value, denoted as $Z$. The above “two-phase” phenomenon is the basis on which we develop the online amplitude recovery procedure.

![Fig. 1. Signal characteristics.](image-url)
To better understand how the online amplitude recovery works, first assume that matrix $\hat{A}$ has been evaluated by applying ICA to a historical database of training signals with characteristics discussed earlier in the article. Now consider a system consisting of at least two or more identical components and assume that real-time vibration signals are being measured by multiple sensors. Using $\hat{A}$, we can separate the degradation signals of the identical components in real-time. Next, we assume that at any time instance only one component transitions from the non-defective phase into the degradation phase. Then, we let $t_{(1),d}$ denote the transition time of the first degraded component, where subscript $d$ corresponds to the onset of the degradation phase. From a practical standpoint, $t_{(1),d}$ can be identified by applying an online change-point detection algorithm such as Cumulative Sum (CUSUM) (Basseville and Nikiforov, 1993) to the component degradation signals. Therefore, before $t_{(1),d}$, all components are in the non-defective phase, and all sensors exhibit close to zero amplitude at $f^*$. At $t_{(1),d}$, only one component exhibits a spike at frequency $f^*$, which causes all sensors to exhibit a spike at $f^*$ at this time epoch with various amplitudes. We denote the highest amplitude among all sensors at epoch $t_{(1),d}$ as $\max(x(f^*, t_{(1),d}))$. By the assumptions made in our framework, $\max(x(f^*, t_{(1),d}))$ will likely correspond to the sensor that is closest to the component that first degrades, and thus will be most sensitive to the initial degradation level $Z$. Consequently, we use $\max(x(f^*, t_{(1),d}))$ as an estimator of the initial degradation level $Z$.

For every component $j$, we expect to capture its transition time, denoted as $t_{j,d}$, using an online change-point detection algorithm such as CUSUM (Basseville and Nikiforov, 1993). At $t_{j,d}$, the corresponding amplitude of the unscaled degradation signal of component $j$ is denoted as $\hat{s}_f^j(t_{j,d})$. Recall that the true degradation signal of component $j$ at $t_{j,d}$ should be equal to $Z$, which can be estimated by $\max(x(f^*, t_{(1),d}))$. Thus, the ratio of its true degradation signal to its unscaled degradation signal (i.e., the “re-scaling ratio”), $R_j$, is defined as

$$ R_j = \frac{\max(x(f^*, t_{(1),d}))}{\hat{s}_f^j(t_{j,d})} \text{ for } j = 1, 2, \ldots, n. \quad (7) $$

Consequently, the true amplitude of the component’s degradation signal can be recovered according to

$$ \hat{s}_f^j(t_k) = R_j \times \hat{s}_f^j(t_k), \quad (8) $$

where $\hat{s}_f^j(t_k)$ is the amplitude of the unscaled component signal evaluated using ICA at time $t_k$ and $\hat{s}_f^j(t_k)$ is the corresponding “re-scaled” degradation signal.

**Remark 1.** $\max(x(f^*, t_{(1),d}))$ is maximized when dedicated sensors are placed directly on the components being monitored. That is, the highest amplitude among all sensors at the first onset of a defect will be “almost equivalent” to that of the initial degradation level. Under this condition, the amplitudes of component degradation signals will be completely recovered. However, since this article is focused on applications with physical and functional restrictions on sensor placement, the proposed amplitude recovery procedure provides only an approximate method for recovering the amplitude information.

Figure 2 summarizes the first stage of the proposed methodology described in Sections 3 and 4. The first graph...
5. Degradation modeling and prognostics

The second stage of our prognostic framework incorporates an adaptive prognostics model to characterize component degradation signals. This prognostics model is developed on the premise that the functional form of a component degradation signal can be modeled as a continuous-time continuous-state stochastic process, which is similar to the one presented in Gebrael et al. (2005). After the model is developed, the RLDs of individual components can be updated in real-time using partially observed component degradation signals. The residual lifetime of each component is defined as the time required for the trajectory of the component degradation signal to reach a pre-specified failure threshold.

The prognostics model for a specific component $j$ can be expressed as follows:

$$s_j^*(t) = \eta(t; \Theta_j, \Phi) + \epsilon(t),$$

where $s_j^*(t)$ represents the degradation signal of component $j$, $\eta(\cdot)$ represents the parametric functional forms used to model the path of the degradation signal (e.g., linear, exponential, polynomial), $\Phi$ is a vector of known coefficients that are assumed to be the same for all identical components, $\epsilon(t)$ represents the signal noise and is assumed to be a Brownian motion, and $\Theta_j$ is a vector of stochastic coefficients that takes a unique value for each specific component $j$. In another words, the values of the stochastic coefficients may differ for individual identical components, which represents the unit-to-unit variability that exists between identical components due to manufacturing differences, material inhomogeneity, etc. To capture the unit-to-unit variability, we assume that the actual value of $\Theta_j$ is unknown for component $j$ but is expected to follow a known prior distribution, whose probability distribution function (pdf) is denoted by $\pi(\Theta_j)$.

**Remark 2.** The assumption of a Brownian motion error has been widely used to characterize the stochastic error in the literature on degradation modeling (Doksum and Höylland, 1992; Gebrael et al., 2005; Wang, 2010). This assumption is only for the purpose of degradation modeling and is not related to the success of the signal separation stage, which relies on the assumption that all component increments are statistically independent and non-Gaussian. This assumption can still be satisfied if the noise term $\epsilon(t)$ follows another type of stochastic process.

We denote the residual lifetime of component $j$ as $T_j$. Before the component starts to degrade, its residual lifetime is equivalent to its lifetime. The cumulative distribution function (cdf) of its lifetime (i.e., the probability that $T_j$ is less than some time $t$), which is equivalent to the probability that the trajectory of the degradation signal reaches the pre-defined threshold (denoted by $\gamma$) before $t$, is given by the following expression:

$$Pr(T_j \leq t) = Pr\left(s_j^*(t) \geq \gamma \right) = \int_{\Theta_j} Pr\left(s_j^*(t) \geq \gamma \mid \Theta_j \right) \pi(\Theta_j) d\Theta_j.$$  \hspace{1cm} (10)

As the real-time observations of its component degradation signal become available, the posterior distribution of $\Theta_j$ can be updated given the observations using a Bayesian approach. For example, if a sequence of degradation signals $s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k)$ is observed at the following times $t_1, t_2, \ldots, t_k$, then the posterior distribution of $\Theta_j$ evaluated at observation time $t_0$ can be expressed as follows:

$$p(\Theta_j | s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k)) \propto l(s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k) | \Theta_j) \pi(\Theta_j),$$  \hspace{1cm} (11)

where $p(\Theta_j | s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k))$ is the posterior distribution of the stochastic coefficients updated at time $t_k$, $l(s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k) | \Theta_j)$ is the likelihood function, and $\pi(\Theta_j)$ is the prior distribution of $\Theta_j$.

After the posterior distribution is evaluated, the cdf of the RLD of the component satisfies a similar expression to Equation (10) as follows:

$$Pr(T_j \leq t \mid s_j^*(t_1), \ldots, s_j^*(t_k)) = Pr\left(s_j^*(t + t_k) \geq \gamma \mid s_j^*(t_1), \ldots, s_j^*(t_k) \right) = \int_{\Theta_j} Pr\left(s_j^*(t + t_k) \geq \gamma \mid \Theta_j \right) p(\Theta_j | s_j^*(t_1), \ldots, s_j^*(t_k)) d\Theta_j,$$

and the Bayes framework improves the prediction of RLD by reducing the uncertainty corresponding to the posterior distribution of $\Theta_j$.

Figure 3 provides an overall illustration of the RLD updating framework using a real-time component degradation signal. When no observations are available at time $t_0$ (as

in Fig. 2 illustrates the time-domain periodic waveform acquired by two sensors $x_1(t_k)$ and $x_2(t_k)$ as an example. Each waveform is assumed to contain a specific defective frequency $f^*$. The real parts of the DFTs of $x_1(t_k)$ and $x_2(t_k)$ are shown in the second graph, in which the values of the two sensors at $f^*$ are selected as $s_1^*(t_k)$ and $s_2^*(t_k)$. The third graph shows the two real sensor signals constructed by successive sampling. The fourth graph illustrates the signal de-trending by computing sensor increments $\Delta x_1^j$ and $\Delta x_2^j$ to estimate the mixing matrix $\mathbf{A}$, which is then utilized to estimate the unscaled degradation signals, denoted as $\hat{s}_1^j(t_k)$ and $\hat{s}_2^j(t_k)$, $k = 1, 2, \ldots, M$, illustrated in the fifth graph. The sixth graph illustrates the amplitude recovery procedure and plots the estimated amplitude of the component signals $\hat{s}_1^j(t_k)$ and $\hat{s}_2^j(t_k)$, $k = 1, 2, \ldots, M$. 

**Remark 2.** The assumption of a Brownian motion error has been widely used to characterize the stochastic error in the literature on degradation modeling (Doksum and Höylland, 1992; Gebrael et al., 2005; Wang, 2010). This assumption is only for the purpose of degradation modeling and is not related to the success of the signal separation stage, which relies on the assumption that all component increments are statistically independent and non-Gaussian. This assumption can still be satisfied if the noise term $\epsilon(t)$ follows another type of stochastic process.

We denote the residual lifetime of component $j$ as $T_j$. Before the component starts to degrade, its residual lifetime is equivalent to its lifetime. The cumulative distribution function (cdf) of its lifetime (i.e., the probability that $T_j$ is less than some time $t$), which is equivalent to the probability that the trajectory of the degradation signal reaches the pre-defined threshold (denoted by $\gamma$) before $t$, is given by the following expression:

$$Pr(T_j \leq t) = Pr\left(s_j^*(t) \geq \gamma \right) = \int_{\Theta_j} Pr\left(s_j^*(t) \geq \gamma \mid \Theta_j \right) \pi(\Theta_j) d\Theta_j.$$  \hspace{1cm} (10)

As the real-time observations of its component degradation signal become available, the posterior distribution of $\Theta_j$ can be updated given the observations using a Bayesian approach. For example, if a sequence of degradation signals $s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k)$ is observed at the following times $t_1, t_2, \ldots, t_k$, then the posterior distribution of $\Theta_j$ evaluated at observation time $t_0$ can be expressed as follows:

$$p(\Theta_j | s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k)) \propto l(s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k) | \Theta_j) \pi(\Theta_j),$$  \hspace{1cm} (11)

where $p(\Theta_j | s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k))$ is the posterior distribution of the stochastic coefficients updated at time $t_k$, $l(s_j^*(t_1), s_j^*(t_2), \ldots, s_j^*(t_k) | \Theta_j)$ is the likelihood function, and $\pi(\Theta_j)$ is the prior distribution of $\Theta_j$.

After the posterior distribution is evaluated, the cdf of the RLD of the component satisfies a similar expression to Equation (10) as follows:

$$Pr(T_j \leq t \mid s_j^*(t_1), \ldots, s_j^*(t_k)) = Pr\left(s_j^*(t + t_k) \geq \gamma \mid s_j^*(t_1), \ldots, s_j^*(t_k) \right) = \int_{\Theta_j} Pr\left(s_j^*(t + t_k) \geq \gamma \mid \Theta_j \right) p(\Theta_j | s_j^*(t_1), \ldots, s_j^*(t_k)) d\Theta_j,$$

and the Bayes framework improves the prediction of RLD by reducing the uncertainty corresponding to the posterior distribution of $\Theta_j$.

Figure 3 provides an overall illustration of the RLD updating framework using a real-time component degradation signal. When no observations are available at time $t_0$ (as
Signal separation and prognostics

Given that we have observed the following sequence of degradation signals \( s_j^r(t_1), s_j^r(t_2), \ldots, s_j^r(t_k) \), the posterior distribution of \( \beta_j \) follows a normal distribution with mean \( \mu_{j,k} \) and variance \( \sigma_{j,k}^2 \), where

\[
\mu_{j,k} = \frac{\sigma_1^2 \ln\left( s_j^r(t_k) \right) + \mu_1 \sigma^2}{\sigma_1^2 t_k + \sigma^2},
\]

\[
\sigma_{j,k}^2 = \frac{\sigma^2 \sigma_1^2}{\sigma_1^4 t_k + \sigma^2}.
\]

Consequently, the RLD of a component \( j \) evaluated at time \( t_k \) is given as

\[
Pr \left( T_j \leq t \mid s_j^r(t_1), \ldots, s_j^r(t_k) \right) = \Pr \left( \ln\left( s_j^r(t_k) + t_k \right) \geq \gamma \mid s_j^r(t_1), \ldots, s_j^r(t_k) \right) = \Phi \left( \frac{\tilde{\mu}_j(t_k + t) - D}{\tilde{\sigma}_j(t_k + t)} \right),
\]

where \( \Phi(\cdot) \) is the cdf of a standard normal random variable, \( \tilde{\mu}_j(t_k + t) = \ln(s_j^r(t_k)) + \mu_{j,k} t \) and \( \tilde{\sigma}_j(t_k + t) = \sigma_{j,k}^2 t_k + \sigma^2 \). The median of the RLD is used as a prediction of the residual lifetime.

6. A numerical example using simulated degradation signals

In this section, we use simulated degradation signals to evaluate the performance of our proposed methodology. First, we will describe the simulation setup and the procedure. Then, we will conduct sensitivity analyses on variables that may affect the performance of the signal separation stage, which include the level of noise in component degradation signals (“degradation noise” for short) and the correlation between component increments. Finally, we will investigate...
the sensitivity of the residual-life prediction to the correlation between component increments, the degradation noise, and the re-scaling ratio discussed in Section 4.2.

6.1. Simulation setup and procedure

In our numerical example, we consider a system with two identical components whose degradation processes can be monitored using three vibration sensors. We design this specific setup according to the assumption that the number of sensors, denoted as $m$, is greater than or equal to the number of components, denoted as $n$, which is a general assumption that guarantees full column rank of the mixing matrix $A$. Figure 4 shows the hypothetical system used in our simulation study. The three sensors are labeled as “Sensor 1,” “Sensor 2,” and “Sensor 3,” whereas the components are labeled as “Component A” and “Component B.” In reality, this system simulates a gearbox with identical components whose degradation processes can be monitored using three vibration sensors. We design this system initially to operate and when the system stops, respectively. $t_M$ is expected to be long enough to have two components run to failure. Specifically, $M = 300$ in this study. For $t_k < t_{j,d}$, we let the amplitude of $s_j(t_k)$ at the defective frequency $f^*$ be a small value close to zero. In this study, the amplitude is randomly chosen from the following distribution $U[4 \times 10^{-4}, 6 \times 10^{-4}]$. For $t_k \geq t_{j,d}$, we let the amplitude of $s_j(t_k)$ at the defective frequency $f^*$ increase randomly in $t_k$ according to the stochastic model expressed in Equation (13). The simulated component vibration signals for Components A and B can be represented as $s_1(t_k)$ and $s_2(t_k)$.

3. We computed the corresponding sensor vibration signals, $x_1(t_k)$, $x_2(t_k)$, and $x_3(t_k)$ based on $s_1(t_k)$ and $s_2(t_k)$ using Equation (2). We used the mixing matrix $A$ defined earlier in this section. Examples of component vibration signals $s_1(t_k)$ and $s_2(t_k)$ and sensor vibration signals $x_1(t_k)$, $x_2(t_k)$, and $x_3(t_k)$ are shown in Fig. 5. As mentioned earlier in Section 3, component degradation signals can be obtained by taking the exact amplitude of the component vibration signals at frequency $f^*$.

![Figure 4](image_url)  
**Fig. 4.** Simulation setup.

![Figure 5](image_url)  
**Fig. 5.** Sample time-domain component vibration signals (on the left) and the corresponding sensor vibration signals (on the right).
The two component degradation signals are detonated as \( s_1^{\star}(t_k) \) and \( s_2^{\star}(t_k) \), for \( k = 1, \ldots, M \).

4. The sensor vibration signals \( x_1(t_k), x_2(t_k), \) and \( x_3(t_k) \) were then used to compute the estimated degradation signals \( \hat{s}_1^{\star}(t_k) \) and \( \hat{s}_2^{\star}(t_k) \). To do this, we followed the procedure outlined in Fig. 2. The estimated degradation signals were used to predict the RLD of each component using the base case model discussed in Section 5. The median of the RLD was utilized as a prediction of the residual lifetime of each component.

5. To be able to benchmark the performance of our model, we compared the residual-life predictions computed using the estimated degradation signals \( \hat{s}_1^{\star}(t_k) \) and \( \hat{s}_2^{\star}(t_k) \) with those computed using the true component degradation signals \( s_1^{\star}(t_k) \) and \( s_2^{\star}(t_k) \). The residual-life predictions of the true degradation signals were also computed using the base case model discussed in Section 5. The accuracy of the prediction was evaluated using the following equation:

\[
\text{error} = \frac{|\text{Predicted Lifetime} - \text{Actual Lifetime}|}{\text{Actual Lifetime}},
\]

where the “Actual Lifetime” is the time point at which the component degradation signal crosses a pre-specified failure threshold, and the “Predicted Lifetime” is equal to the observation level plus the predicted residual lifetime.

Figure 6 presents an example of a pair of component degradation signals (represented by solid curves) and their corresponding estimated signals (represented by dashed curves). The horizontal line drawn at 0.025 in Fig. 6 represents the failure threshold. From Fig. 6, we observe that the shape of a component degradation signal can be estimated accurately, but the amplitude of the signal is not completely recovered. This bias is mainly due to the amplitude recovery procedure. Particularly, due to the bias/error induced by the amplitude recovery procedure, at the time epoch when the true degradation signal crosses the failure threshold (i.e., the true failure time), the estimated degradation signal has not yet crossed the failure threshold. Consequently, the estimated signal takes a slightly longer time to reach the failure threshold.

### 6.2. Sensitivity analyses on the signal separation stage

The performance of the signal separation stage relies on the effectiveness of ICA. One of the necessary assumptions for using ICA is the independence between the component increments. In this section, we study the performance of ICA when this assumption is not satisfied. This is accomplished by adjusting the correlation coefficient between the
Fig. 9. Comparing prediction errors from true and estimated component signals at three variance levels: (a) Variance = 0.0025; (b) Variance = 0.0015; and (c) Variance = 0.0005.

Brownian motion errors of the two component degradation signals. In addition, we study the impact of the degradation noise on the effectiveness of the ICA algorithm. The noise level is controlled by the variance of the Brownian motion error in a component degradation signal.

The performance of the signal separation is evaluated by the difference between the true and the estimated mixing matrices, in which the value of the true matrix has been specified earlier in this section. Since the ICA algorithm cannot preserve the norm of each column in the matrix...
(as discussed in Section 4), we focus on comparing the normalized matrices to remove any unnecessary biases. We let $A^N$ and $\hat{A}^N$ denote the normalized matrices of the true $A$ and the estimated matrix $\hat{A}$. Each column of the normalized matrix was computed by normalizing the corresponding column of the original matrix. The value of $A^N$ is $A^N = [0.6985, 0.4915; 0.4657, 0.6554; 0.5433, 0.5735]$. We compare $\hat{A}^N$ with $A^N$ using the Frobenius norm of $\delta A^N = \hat{A}^N - A^N$, which is considered as a measurement of the estimation error of the mixing matrix.

Signal separation was performed at 10 levels of correlation (correlation coefficient equal to 0, 0.1, ..., 0.9) and 10 levels of degradation noise (the variance of the Brownian motion error equal to 0.0005, 0.001, ..., 0.005). Figure 7 illustrates three sample degradation signals corresponding to three levels of the Brownian motion variance: 0.0005, 0.0015, and 0.0025. It is clear that the degradation signal with a higher Brownian motion variance is more noisy.

In total, there were $10 \times 10 = 100$ testing scenarios. For each scenario, we conducted 200 rounds of simulations. The signals from each round were utilized to compute $A$ as well as the corresponding $\| \delta A^N \|_F$. Then, we chose the median of these 200 values for $\| \delta A^N \|_F$ as a point estimator of the estimation error. Figure 8 shows the estimation errors for the 100 scenarios.

It is clear that the estimation error increases as the level of correlation increases. This is expected since increasing the correlation results in greater deviations from the assumption of independence. On the other hand, the level of Brownian motion variance does not exhibit a significant impact on the estimation error. We believe the reason for this is that the level of degradation noise mainly affects the amplitude of the component increments. Since ICA does not preserve amplitude information, changing the noise level has little effect on the performance of ICA.

### 6.3. Sensitivity analysis on the residual-life prediction

In this section, we will investigate the effects of three variables on accuracy of the residual-life prediction. The first variable is the degradation noise. Although this variable does not significantly affect the signal separation stage as we have investigated, it may affect the stage of prognostics, since a higher noise in the degradation signals may reduce the effectiveness of residual-life prediction. The second variable is the correlation between component increments. From Section 6.2, we have concluded that this correlation affects signal separation. Thus, it will also affect the residual-life prediction. The third variable is the re-scaling ratio. How well we estimate the re-scaling ratio has a direct impact on the efficiency of the amplitude recovery and thus will impact the residual-life prediction.

#### 6.3.1. The impact of degradation noise

In this study, we investigate the effect of degradation noise on residual-life prediction. Similar to Section 6.2, the noise level is controlled by tuning the variance parameter of the Brownian motion. We investigate the effects of three levels of variance: 0.0005, 0.0015, 0.0025. Three sample degradation signals corresponding to the three variance levels are already shown earlier in Fig. 7.

For each level of variance, 100 simulation runs were performed, out of which 20 were used as training signals to estimated the mixing matrix. The average of the 20 estimated matrices was utilized as the final estimator of the mixing matrix $\hat{A}$. The other 80 runs were used for validation. For each round in the validation set, the estimation $\hat{A}$ was applied to compute the estimated degradation signals of both components in this round. Next, we assumed that each estimated degradation signal was partially observable up to several lifetime percentiles (evaluated post facto): 10%, 20% until 90% of the component’s actual lifetime. For each component, given its partially observed degradation signal, we predicted its residual lifetime and calculated the corresponding prediction error. The prediction errors at different lifetime percentiles for the 80 rounds of simulation in the validation set are summarized in Fig. 9, where parts (a), (b), and (c) correspond to the three variance levels. The right column of Fig. 9 illustrates the prediction errors associated with the estimated degradation signals, whereas the left column presents the prediction errors using the true degradation signals as the benchmark.

We conducted a one-sided $t$-test, a one-sided $F$-test, and a one-sided Wilcoxon rank sum test to compare the mean, the variance, and the median of the prediction errors of the estimated versus the true degradation signals. The null hypothesis for each of the three tests was that the corresponding statistics (the mean, the variance, or the median) of the prediction errors of true and estimated degradation signals

<table>
<thead>
<tr>
<th>Variance</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Var} = 0.0025$</td>
<td>0.587</td>
<td>0.689</td>
<td>0.707</td>
<td><strong>0.0405</strong></td>
<td>0.213</td>
<td>0.222</td>
<td>0.219</td>
<td><strong>0.023</strong></td>
<td><strong>0.00226</strong></td>
</tr>
<tr>
<td>$\text{Var} = 0.0015$</td>
<td>0.192</td>
<td>0.103</td>
<td>0.106</td>
<td>0.095</td>
<td><strong>0.042</strong></td>
<td><strong>0.0041</strong></td>
<td><strong>0.006</strong></td>
<td><strong>0.0013</strong></td>
<td>$2.8 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\text{Var} = 0.0005$</td>
<td><strong>0.045</strong></td>
<td><strong>0.013</strong></td>
<td><strong>0.0049</strong></td>
<td><strong>0.0030</strong></td>
<td><strong>0.0042</strong></td>
<td><strong>0.0028</strong></td>
<td>$5.8 \times 10^{-4}$</td>
<td>$3.4 \times 10^{-5}$</td>
<td>$1.1 \times 10^{-6}$</td>
</tr>
</tbody>
</table>

$p$-values smaller than 0.05 are in bold type.
Table 2. The \(p\)-values of the \(f\)-test corresponding to the three variance groups

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Variance</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Var} = 0.0025)</td>
<td></td>
<td>0.207</td>
<td>0.628</td>
<td>0.207</td>
<td>0.469</td>
<td>0.014</td>
<td>0.018</td>
<td>0.0009</td>
<td>(1.9 \times 10^{-7})</td>
<td>(1.8 \times 10^{-12})</td>
</tr>
<tr>
<td>(\text{Var} = 0.0015)</td>
<td></td>
<td>0.0041</td>
<td>0.0003</td>
<td>0.0001</td>
<td>0.0009</td>
<td>0.0002</td>
<td>0.0009</td>
<td>0.0008</td>
<td>0.0010</td>
<td>0.0001</td>
</tr>
<tr>
<td>(\text{Var} = 0.0005)</td>
<td></td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0024</td>
<td>0.0032</td>
<td>0.0008</td>
<td>0.0002</td>
<td>0.0009</td>
<td>0.0004</td>
<td>(1.1 \times 10^{-6})</td>
</tr>
</tbody>
</table>

\(p\)-values smaller than 0.05 are in bold type.

are equal. The alternative hypothesis was that the corresponding statistics of the prediction errors of estimated degradation signals is higher than that of the prediction errors of true degradation signals. The significance of a test is represented by the \(p\)-value.

The corresponding \(p\)-values are listed in Tables 1, 2, and 3, with \(p\)-values smaller than 0.05 highlighted in boldface. According to the tables, each hypothesis test is conducted under various conditions including different variance levels of the Brownian motion error and different levels of the lifetime percentile. The \(p\)-values smaller than 0.05 indicate that we reject the null hypothesis at 5% significant level under the corresponding test conditions. In other words, under the corresponding conditions, we can conclude that the corresponding statistics of the prediction errors of estimated degradation signals is significantly higher than that of the prediction errors of true degradation signals.

From Fig. 9 and Tables 1, 2, and 3, we can summarize the following observations:

1. The prediction errors of both true and estimated degradation signals decrease as the variance of the Brownian motion decreases. This is expected because reducing the noise level usually improves predictability.

2. For each type of hypothesis test, the number of lifetime percentiles associated with \(p\)-values smaller than 0.05 increases as the noise level decreases. This observation may be due to the fact that when using true component signals, reducing the noise level directly results in significant improvements in prediction accuracy. However, this is not necessarily the case for the estimated component signals, as there is an additional error that is introduced by the signal separation step. Reducing degradation noise does not improve the performance of signal separation as discussed in Section 6.2.

Table 3. The \(p\)-values of the Wilcoxon rank sum test corresponding to the three variance groups

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Variance</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Var} = 0.0025)</td>
<td></td>
<td>0.604</td>
<td>0.538</td>
<td>0.551</td>
<td>0.092</td>
<td>0.423</td>
<td>0.377</td>
<td>0.895</td>
<td>0.089</td>
<td>0.0002</td>
</tr>
<tr>
<td>(\text{Var} = 0.0015)</td>
<td></td>
<td>0.169</td>
<td>0.152</td>
<td>0.189</td>
<td>0.506</td>
<td>0.098</td>
<td>0.010</td>
<td>0.0036</td>
<td>0.0005</td>
<td>0.0001</td>
</tr>
<tr>
<td>(\text{Var} = 0.0005)</td>
<td></td>
<td>0.547</td>
<td>0.270</td>
<td>0.216</td>
<td>(0.0003)</td>
<td>(0.017)</td>
<td>(0.047)</td>
<td>(0.0026)</td>
<td>(0.0003)</td>
<td>(1.3 \times 10^{-5})</td>
</tr>
</tbody>
</table>

\(p\)-values smaller than 0.05 are in bold type.

6.3.2. The effect of correlation between component increments

As demonstrated in Section 6.2, the correlation between component signals significantly affects the signal separation step. Here, we evaluate the accuracy of residual-life prediction under different correlation levels. Similar to Section 6.2, we control the correlation by adjusting the correlation coefficient between the Brownian motion errors of the two component degradation signals. We test three levels of correlation coefficient: 0, 0.5, and 1. Signals were simulated and prediction errors were evaluated in a similar manner to what is presented in Section 6.3.1. The prediction errors are shown in Fig. 10, in which the left and right columns correspond to the prediction errors of true and estimated degradation signals, respectively. From Fig. 10, it is clear that when the correlation coefficient is zero, there is no significant difference between the prediction errors from estimated signals versus those evaluated using the true signals. However, as the correlation coefficient is increased to 0.5 and 1, the accuracy of the residual-life prediction significantly decreases. The reason for this behavior is that increasing the correlation coefficient will contaminate the accuracy of the signal separation, which will eventually affect the accuracy of the residual-life prediction.

6.3.3. The effectiveness of the online amplitude recovery procedure

Here, we focus on the impact of the amplitude recovery procedure on the accuracy of residual-life prediction. An important factor that affects the performance of the amplitude recovery procedure is the re-scaling ratio. In Section 4.2, we highlighted that the value of the re-scaling ratio depends on how well we estimate the initial degradation level \(Z\). In this section, we test the effect of four different esti-
Fig. 10. The accuracy of the residual-life prediction under three correlation levels: (a) correlation coefficient $= 0$; (b) correlation coefficient $= 0.5$; and (c) correlation coefficient $= 1$.

Estimates of $Z$ on the accuracy of the residual-life prediction. The first estimate is equal to the true value of $Z$, which will recover 100% of the amplitude. To choose the second value, we note that $A = [0.9, 0.6; 0.6, 0.8; 0.7, 0.7]$, which indicates that $x_1(t_k) = 0.9s_1(t_k) + 0.6s_2(t_k) + e_1(t_k)$, $x_2(t_k) = 0.6s_1(t_k) + 0.8s_2(t_k) + e_2(t_k)$, $x_3(t_k) = 0.7s_1(t_k) + 0.7s_2(t_k) + e_3(t_k)$, $k = 1, 2, \ldots$. At time $t_{(1,d)}$, if Component A degrades first, we will have $\max(x^{F_1}(t_{(1,d)})) \approx 0.9Z$. On the other hand, if Component B degrades first, we will have $\max(x^{F_2}(t_{(1,d)})) \approx 0.8Z$. In general, components A and B will have an equal probability to degrade. Thus, on average, the estimate is equal to $0.85Z$. The third and fourth estimates of $Z$ were chosen to be $0.55Z$ and $0.25Z$, which corresponds to even worse estimates the re-scaling ratio.

Signals were simulated and prediction errors were evaluated in a similar manner to that presented in Section 6.3.1. However, different from the previous simulation, we did not
Fig. 11. The accuracy of the residual-life prediction under different estimates of the initial degradation level.

follow the re-scaling procedure described Section 4.2 to estimate $\max(x^T(t_{11}, d))$. Instead, we manually specified the value $\max(x^T(t_{11}, d))$ to be equal to one of the four values $(Z, 0.85Z, 0.55Z, \text{and} 0.25Z)$ and calculated the residual-life predictions for each component at different lifetime percentiles. Figure 11 illustrates the prediction errors under the four estimates of the initial degradation level. In this figure, the word “ratio” in the title of each graph represents the ratio of the estimate to the true $Z$. It can be seen that the prediction errors when ratio $= 1$ and ratio $= 0.85$ do not exhibit significant difference. This can be explained by reference Fig. 6, in which the estimated degradation signals (dashed curves) are biased from the corresponding true degradation signals (solid curves). Although this bias is observable, the failure times of the estimated degradation signals are not significantly different from those of true degradation signals. Thus, we do not expect a significant difference between the lifetime prediction errors. However, the prediction errors increase significantly as the ratio decreases to 0.55 and further to 0.25. This is expected, since a significant decrease in the ratio will eventually results in an unreliable prediction. In reality, the ratio is related to the distance between a sensor and a component. Therefore, we recommend that in order to improve the performance of the residual-life prediction for each individual component, at least one sensor should be located as close as possible to one of the components that are being monitored, especially if they are identical or generate the same signal frequency.

7. Conclusions

This article has focused on vibration monitoring of systems with identical components that generate the same defective frequencies. We have proposed a simultaneous signal separation and prognostics framework, in which we assumed that sensor vibration signals are linear combinations of component vibration signals. In the first stage, we applied ICA on pre-processed sensor signals to estimate the mixing matrix. We then proposed an amplitude recovery procedure that approximately recovers the amplitude information of component degradation signals. In the second stage, we modeled component degradation signals as continuous stochastic processes and predicted the RLDs of individual components.

To evaluate the robustness of the proposed framework, we tested the combined performance of the signal separation and the prognostics stages under different signal-to-noise ratios, different levels of correlation between the increments of individual components, and the errors in the amplitude recovery procedure. The results showed
that the signal separation is the most sensitive to the correlation between component increments, whereas the prognostics model is most sensitive to signal noise and errors in the amplitude recovery. These results can be used to provide guidelines within which our framework is expected to perform reasonably well. Since the performance of the re-scaling procedure is related to sensor locations, the results of simulation suggest that for systems with identical components or components generating similar signal frequency, it is important that at least one sensor be located as close as possible to each of these components in order to ensure a reasonable prediction of the component RLDs.

This work can be extended by relaxing the linearity assumption and perhaps employing a convolutive model to account for more realistic scenarios. In reality, the mixing process of vibration signals is not instantaneous but includes delays, which is better represented as a convolution process.

Funding

This research was sponsored by grants from the U.S. National Science Foundation (CMMI-0856192).

References


**Biography**

Li Hao received a B.S. degree in Automotive Engineering from Tsinghua University, Beijing, China, in 2009 and an M.S. degree in Statistics from the Georgia Institute of Technology in 2012. Currently, she is a Ph.D. student at the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. Her research interests are focused on the prognostics and degradation-based control of complex systems.

Nagi Gebraeel received M.S. and Ph.D. degrees in Industrial Engineering from Purdue University, West Lafayette, Indiana, in 1998 and 2003, respectively. Currently, he is an Associate Professor at the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. His research focuses on improving the accuracy of predicting unexpected failures of engineering systems by leveraging sensor-based data streams. His major research interests are in the areas of degradation modeling and sensor-based prognostics, reliability engineering, and maintenance operations and logistics. He is a member of IIE and INFORMS.

Jianjun Shi received B.S. and M.S. degrees in Electrical Engineering from the Beijing Institute of Technology, Beijing, China, in 1984 and 1987, respectively, and a Ph.D. degree in Mechanical Engineering from the University of Michigan, Ann Arbor, in 1992. Currently, he is the Carolyn J. Stewart Chair Professor in the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. His research interests include the fusion of advanced statistical and domain knowledge to develop methodologies for modeling, monitoring, diagnosis, and control of complex manufacturing systems. He is a Fellow of the Institute of Industrial Engineers, a Fellow of the American Society of Mechanical Engineers, a Fellow The Institute for Operations Research and the Management Sciences, and a member of ASQ, SME, and ASA.