

# Data-Driven Variation Source Identification for Manufacturing Process Using the Eigenspace Comparison Method

Nong Jin, Shiyu Zhou

*Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, WI 53706, USA*

Received 13 July 2005; revised 27 December 2005; accepted 4 February 2006

DOI 10.1002/nav.20150

Published online 22 March 2006 in Wiley InterScience (www.interscience.wiley.com).

**Abstract:** Variation reduction of manufacturing processes is an essential objective of process quality improvement. It is highly desirable to develop a methodology of variation source identification that helps quickly identify the variation sources, hence leading to quality improvement and cost reduction in manufacturing systems. This paper presents a variation source identification method based on the analysis of the covariance matrix of process quality measurements. The identification procedure utilizes the fact that the eigenspace of the quality measurement covariance matrix can be decomposed into a subspace due to variation sources and a subspace purely due to system noise. The former subspaces for different samples will be the same if the same variation sources dominate. A testing procedure is presented, which can determine the closeness of the subspaces under sampling uncertainty. A case study is conducted to illustrate the effectiveness of this methodology. © 2006 Wiley Periodicals, Inc. *Naval Research Logistics* 53: 383–396, 2006.

**Keywords:** eigenspace; factor analysis; principal component analysis; variation reduction; variation source identification

## 1. INTRODUCTION

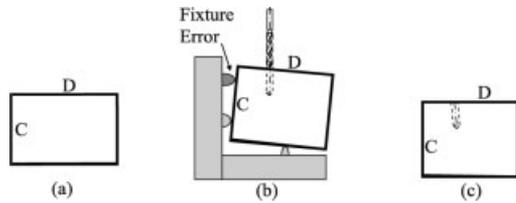
Variation reduction of manufacturing processes is an essential objective of process quality control. Effective variation reduction can improve process efficiency and product quality to gain competition advantages. Under normal working conditions, the process variation is caused by the natural variation or so-called common causes. Under abnormal working conditions, the excessive process variation is caused by process errors or so-called assignable causes. Since product quality is determined by the conditions of process tooling elements (such as a cutting tool, fixture, and welding gun) in a manufacturing system, the process errors are actually the malfunctioning tooling elements that are responsible for the defective products.

Consider the example of a simple two-step machining process (Fig. 1). The workpiece is a cube of metal (only front view is shown). Surface C of the workpiece is milled in the first step (Fig. 1a). In the second step, a hole is drilled on surface D (Fig. 1b) while the workpiece is located in the fixture through the left and bottom surfaces. Clearly, the resulting hole is not perpendicular to surface D (Fig. 1c) due to the fixture error. The fixture error could be a mean shift

of the locating pin height due to an error in setup or a variance increase of the locating pin height due to the loosening of the pin. In most cases, the fixed mean-shift error can be easily handled by compensation even without knowing the error sources. For example, if the drilled hole always deviated from 90° by a fixed value, the angle between the drill and the fixture can be adjusted to compensate for this deviation without removing the fixture error. The error of variation increase is much more difficult to remove. The sources of this variation increase, so-called “variation sources” in this paper, often need to be identified before being eliminated. One point needing to be clarified here is that there are many sources in a process to cause variation. However, only those sources that cause excessive variation are defined as “variation sources” in this paper. Therefore, variation sources are also called “process fault” in this paper. It might be easy to figure out the variation source for a simple two-step manufacturing process. However, for a large complex manufacturing system, it will be very time-consuming and expensive to inspect every stage of the manufacturing system until we eventually locate the variation source whenever an abnormal working condition is detected.

Statistical process control (SPC) [29] is a popular technique used in practice for quality improvement. SPC technique can be used to detect the quality change. After a

*Correspondence to:* S. Zhou (szhou@engr.wisc.edu)



**Figure 1.** Effect of fixture error on product dimensional quality. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

process change is detected, it is critical to quickly determine the variation sources so that they can be eliminated and the process can be brought back to its normal condition. However, SPC possesses limited diagnosis capabilities to identify the variation sources—the diagnosis of malfunctioning tooling elements is often left to human operators. The variation source identification has been identified as an important research direction in SPC [36].

Due to the fast development of information and sensing technology, massive information regarding the manufacturing process (e.g., product design, process planning, in-process sensing information of process status, product quality information) is now readily available. The 100% inspection for discrete processes and very high sampling rate for continuous processes are not rare in practice now. For example, in the autobody assembly process, 100% dimensional inspection has been achieved through in-line optical coordinate measurement machines [9]. This huge amount of process/product information provides great opportunities to develop variation source identification methodologies.

Recent research has advanced toward this direction. The reported variation source diagnosis methods can be put into two categories: (1) diagnosis based on known physical models linking product quality measurements to process variation sources and (2) diagnosis only based on the product quality measurements.

Linear fault-quality models have been developed for particular processes based on the specific physical process knowledge. Jin and Shi [23], Mantripragada and Whitney [28], Ding, Ceglarek, and Shi [14], and Camelio, Hu, and Ceglarek [8] developed models that link the variation sources such as the fixture error with the product quality measurements for multistage assembly processes. Huang, Zhou, and Shi [22], Djurdjanovic and Ni [18], and Zhou, Huang, and Shi [39] provided linear fault-quality diagnostic models for multistage machining processes. Barton and Gonzalez-Barreto [6] proposed a process-oriented basis representation for multivariate process diagnostics.

These linear physical models can be put in the generic form

$$\mathbf{y} = \mathbf{A}\mathbf{f} + \boldsymbol{\varepsilon}, \quad (1)$$

where  $\mathbf{y}$  is an  $n$  by 1 vector consisting of product quality measurements ( $n$  is the dimension of the quality measurements),  $\mathbf{A}$  is a constant coefficient matrix determined by process/product design,  $\mathbf{f}$  is a vector representing the process variation sources (also called process faults), and  $\boldsymbol{\varepsilon}$  includes the measurement noise and un-modeled variation.

The variation source diagnostic methods based on physical models can be classified as the pattern matching method and the direct estimation method. The pattern matching method is based on the fact that if only one fault happens (only one nonzero component in  $\mathbf{f}$ ) and the covariance of  $\boldsymbol{\varepsilon}$  is in the form of  $\sigma^2\mathbf{I}$ ,  $\mathbf{I}$  is the identity matrix, the eigenvector associated with the largest eigenvalue of the covariance matrix of  $\mathbf{y}$  is the same as the column of  $\mathbf{A}$  that corresponds to the nonzero component of  $\mathbf{f}$ . This method was first proposed by Ceglarek and Shi [10] to identify fixture error in a rigid assembly process. Rong, Ceglarek, and Shi [31] extended this method to compliant assembly process. Ding, Ceglarek, and Shi [15] further relaxed the constraint on the covariance of  $\boldsymbol{\varepsilon}$  and provided a boundary of the eigenvector when the covariance of  $\boldsymbol{\varepsilon}$  is of general structure. The limitation of this method is obvious: this method is only for single fault identification. There are two typical methods in direct estimation category: (i) least squares estimation method. In this method,  $\mathbf{f}$  is first estimated using least squares method [3, 11]. Then, based on the estimation results  $\hat{\mathbf{f}}$ , the variance of  $\mathbf{f}$  is estimated and a hypothesis testing procedure is setup to identify the variation sources [3]. This method can handle the identification of multiple variation sources. A diagnosability study for this method is conducted by Ding, Shi, and Ceglarek [16]. (ii) Variance component analysis method. This method is based on the fact that Model (1) is actually a linear mixed model. The variation estimation methods such as Maximum Likelihood, Restricted Maximum Likelihood, and Minimum Norm Quadratic Unbiased Estimation (MINQUE) can be used to identify the variation sources. Different estimation methods are compared in [17]. A diagnosability study for this method is conducted by Zhou et al. [38]. A hypothesis testing procedure is developed and experimentally validated by Zhou, Chen, and Shi [37]. This method can also handle the identification of multiple variation sources.

The above-mentioned variation source identification methods are based on a physical fault-quality model. The physics of the process need to be thoroughly studied to build up the process model. For a large-scale system, it will be very difficult, if not impossible, to build up such a comprehensive physical model [12]. Limited research has been done on the variation source identification without knowing the process fault-quality model. Apley and Shi [4] presented a statistical technique for estimating the coefficient matrix  $\mathbf{A}$  in Model (1) only based on the quality measurements. A ragged lower triangular form is assumed for  $\mathbf{A}$  to remove

the indeterminacy in the estimation. The columns of  $\mathbf{A}$  are called “fault geometry vectors.” The physical interpretations of the faults are pursued after  $\mathbf{A}$  is estimated. Therefore, this method is more of a descriptive method for extracting and interpreting the information from the quality data. Most recently, Apley and Lee [5] proposed a blind separation approach to identify spatial variation patterns in manufacturing data. Their method focuses on separating the simultaneous variation sources in the present system and does not deal with the variation source identification issue. In other words, their method does not discuss the variation pattern matching issue.

In this article, we proposed a variation source identification technique that does not require any priori information on  $\mathbf{A}$  matrix in the process fault-quality model and the developed model can be further utilized for fault diagnosis. This method first estimates the column vectors of  $\mathbf{A}$  matrix using a gradually learning procedure. These columns can be viewed as signatures of corresponding variation sources. A library of the estimates of the column vectors can be built. Then, the variation source identification can be realized by decomposing and comparing the eigenspaces of the covariance matrix of the quality measurement according to this library.

This paper is structured as follows. In Section 2, the problem formulation and the method of variation source identification are presented. A complete procedure is developed to quickly identify the process variation sources. A case study is illustrated in Section 3 to demonstrate the effectiveness of this technique. Conclusions are presented in Section 4.

## 2. VARIATION SOURCE IDENTIFICATION USING COMMON EIGENSPACE

### 2.1. Problem Formulation

In this paper, we adopt the linear relationship between the process faults and product quality, as shown in (1). In general, the impact of the process fault on the product quality is nonlinear, which is represented by a general function  $g(\cdot)$  in (2),

$$\mathbf{y} = g(\mathbf{f}) + \boldsymbol{\varepsilon}, \quad (2)$$

where  $\mathbf{y}$  is the quality measurements (e.g., the dimensional measurements of a workpiece),  $\mathbf{f}$  represents the process variation sources (e.g., the locations of the pins of the fixture), and  $\boldsymbol{\varepsilon}$  is the measurement noise. Although the relationship between  $\mathbf{y}$  and  $\mathbf{f}$  is in general nonlinear, a linear relationship can provide a fairly good approximation because the process faults are often small in magnitude. This

linear approximation can be obtained through Taylor series expansion around the nominal value of  $\mathbf{f}_0$  as

$$\mathbf{y} \approx g(\mathbf{f}_0) + \left. \frac{\partial g(\mathbf{f})}{\partial \mathbf{f}} \right|_{\mathbf{f}=\mathbf{f}_0} \Delta \mathbf{f} + \boldsymbol{\varepsilon}. \quad (3)$$

If  $\mathbf{f}$  is  $\mathbf{f}_0$ , then no process fault happens and  $g(\mathbf{f}_0)$  is the nominal quality measurement value, which is denoted  $\mathbf{y}_0$ . If we denote  $\Delta \mathbf{y}$  as  $\mathbf{y} - \mathbf{y}_0$ ,  $\left. \frac{\partial g(\mathbf{f})}{\partial \mathbf{f}} \right|_{\mathbf{f}=\mathbf{f}_0}$  as  $\mathbf{A}$ , then

$$\Delta \mathbf{y} = \mathbf{A} \cdot \Delta \mathbf{f} + \boldsymbol{\varepsilon}. \quad (4)$$

Without causing confusion,  $\boldsymbol{\varepsilon}$  in Eq. (4) includes both measurement noise and linearization error. The linear fault quality model as in (1) can be obtained by replacing  $\Delta \mathbf{y}$  by  $\mathbf{y}$  and  $\Delta \mathbf{f}$  by  $\mathbf{f}$ . In this article, the following assumptions are made on Model (1) or (4):

- (1)  $\mathbf{A}$  is an unknown  $n$  by  $p$  matrix. The columns of  $\mathbf{A}$  are linearly independent.
- (2)  $\mathbf{f}$  is a  $p$  by 1 vector that follows a multivariate normal distribution  $N(\mathbf{0}, \mathbf{D})$ , where  $\mathbf{D}$  is a diagonal matrix. The components of  $\mathbf{f}$  are assumed independent because the process faults are often independent of each other. In this article, we focus on the fault of variation increase as opposed to mean shift. Hence, we further assume  $\mathbf{f}$  has zero mean.
- (3)  $\boldsymbol{\varepsilon}$  is a  $n$  by 1 vector that follows a multivariate normal distribution  $N(\mathbf{0}, \sigma^2 \mathbf{I})$ , where  $\sigma^2$  is a scalar and  $\mathbf{I}$  is the identity matrix. Further,  $\boldsymbol{\varepsilon}$  is independent with  $\mathbf{f}$ . Because the measurement noise is dominating in  $\boldsymbol{\varepsilon}$ , this assumption is reasonable if the same measurement device is used to measure all the quality characteristics. This is not rare in practice. For example, the Coordinate Measurement Machine is often used in practice to measure all the dimensions of a workpiece.

This model is similar to the factor analysis model that is widely used in psychometrics [1]. In the conventional factor analysis model, the covariance matrix of  $\boldsymbol{\varepsilon}$  is assumed to be diagonal, not in the form of  $\sigma^2 \mathbf{I}$ , and the covariance matrix of  $\mathbf{f}$  is assumed to be identity, not only a diagonal matrix. In this model, each fault is associated with the corresponding column vector of  $\mathbf{A}$ . In [4], this column vector is called “fault geometry vector.” We will use the same terminology here. The magnitude of a variation source is captured by the variance of corresponding elements of  $\mathbf{f}$ , i.e., the corresponding diagonal element of  $\mathbf{D}$ .

Based on this linear model, the problem of variation source identification can be formulated as follows: Given multiple observations of  $\mathbf{y}$ , how to identify which faults

happen? What are the magnitudes?, i.e., how to identify which fault geometry vectors exist in  $\mathbf{A}$  and what is the variance of  $\mathbf{f}$ ?

One thing that needs to be pointed out is that during the sampling period, not all potential process faults happen simultaneously. Assume there are  $m$  potential process variation sources in total. It is unreasonable to assume that the  $m$  process faults always happen simultaneously. Therefore, we have to allow the number of components of  $\mathbf{f}$  vector and the number of columns of  $\mathbf{A}$  to be changeable from sample to sample.

**2.2. Estimation of the Fault Geometry Vector**

Because  $\mathbf{A}$  in (1) is assumed unknown, we need to obtain the fault geometry vectors first.

First, since  $\mathbf{D}$  is a diagonal covariance matrix of  $\mathbf{f}$ ,  $\mathbf{D}^{1/2}$  and  $\mathbf{D}^{-1/2}$  are diagonal and real. Hence, the variable transformation  $\mathbf{f} = \mathbf{D}^{1/2}\mathbf{f}^*$  and  $\mathbf{A} = \mathbf{A}^*\mathbf{D}^{-1/2}$  can be substituted into (1) to get

$$\mathbf{y} = \mathbf{A}^*\mathbf{f}^* + \boldsymbol{\varepsilon}, \tag{5}$$

where  $\mathbf{f}^* = \mathbf{D}^{-1/2}\mathbf{f}$  and  $\mathbf{A}^* = \mathbf{A}\mathbf{D}^{1/2}$ . The covariance of  $\mathbf{f}^*$ ,  $\boldsymbol{\Sigma}_{\mathbf{f}^*}$ , is an identity matrix because  $\boldsymbol{\Sigma}_{\mathbf{f}^*} = \mathbf{D}^{-1/2}\mathbf{D}(\mathbf{D}^{-1/2})^T = \mathbf{D}^{-1/2}\mathbf{D}\mathbf{D}^{-1/2} = \mathbf{I}$ .

A close look at (5) reveals that it is exactly the model used in probabilistic principal component analysis [34]. It has been shown by Tipping and Bishop [34] that the maximum likelihood estimation of  $\mathbf{A}^*$  is given as

$$\mathbf{A}_{ML}^* = \mathbf{U}_p(\boldsymbol{\Gamma}_p - \sigma_{ML}^2\mathbf{I})^{1/2}\mathbf{R}. \tag{6}$$

In the above equation,  $\boldsymbol{\Gamma}_p$  is a  $p$  by  $p$  diagonal matrix with diagonal elements of  $\lambda_1 \sim \lambda_p$ , where  $\lambda_i$ 's ( $i = 1, \dots, n$ ,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ) are the eigenvalues of  $\mathbf{S}_y$ , the sample covariance matrix of  $\mathbf{y}$ .  $\mathbf{U}_p$  is a  $n$  by  $p$  matrix consisting of the  $p$  eigenvectors that are associated with  $\lambda_1 \sim \lambda_p$ .  $\sigma_{ML}^2$  is the maximum likelihood estimation of  $\sigma^2$  and is given by  $\sigma_{ML}^2 = \frac{1}{n-p} \sum_{j=p+1}^n \lambda_j$ , where  $\sigma^2$  is the variance of the measurement noise. Matrix  $\mathbf{R}$  is an arbitrary  $p$  by  $p$  orthogonal matrix. Clearly, the estimation of  $\mathbf{A}^*$  is not unique. It possesses rotational indeterminacy. Based on the result in (6), we can get the corresponding estimation result for  $\mathbf{A}$ ,

$$\mathbf{A}_{ML}\mathbf{D}^{1/2} = \mathbf{U}_p(\boldsymbol{\Gamma}_p - \sigma_{ML}^2\mathbf{I})^{1/2}\mathbf{R}, \tag{7}$$

where  $\mathbf{D}$  is the covariance matrix of  $\mathbf{f}$ , representing the magnitude of the process faults. If multiple faults happen simultaneously (i.e.,  $p > 1$ ), Eq. (7) can only provide limited information of  $\mathbf{A}_{ML}$  because of the rotational inde-

terminacy. However, the rotational indeterminacy does not exist if only one fault happens (i.e.,  $p = 1$ ). In this case,  $\mathbf{R} = 1$  is a scalar,  $\boldsymbol{\Gamma}_p = \lambda_1$  is a scalar,  $\mathbf{U}_p$  is the eigenvector  $\mathbf{v}_1$  that is associated with the largest eigenvalue,  $\mathbf{D}^{1/2}$  is a scalar, and  $\mathbf{A}_{ML}$  is the maximum likelihood estimation of the geometry vector ( $\mathbf{a}_f$ ) of the current fault. Clearly, if only one fault exists, the direction of the fault geometry vector is uniquely determined by the direction of  $\mathbf{v}_1$ . However, the indeterminacy of the scale of  $\mathbf{a}_f$  still exists: we cannot uniquely identify the length of  $\mathbf{a}_f$  and the magnitude ( $\mathbf{D}$ ) of the fault. This indeterminacy can be easily removed by putting a constraint on the length of  $\mathbf{a}_f$ . In this paper, we assume  $\|\mathbf{a}_f\| = 1$ , where  $\|\cdot\|$  is the Euclidean norm. With this constraint, we have

$$\tilde{\mathbf{a}}_f = \mathbf{v}_1 \quad \text{and} \quad \tilde{\sigma}_f^2 = \lambda_1 - \sigma_{ML}^2, \tag{8}$$

where  $\tilde{\mathbf{a}}_f$  and  $\tilde{\sigma}_f^2$  are the estimates of the geometry vector and the magnitude (variance) of current single fault, respectively.

Because Eq. (8) can uniquely identify the fault geometry vector and the fault magnitude, it sheds light on the fault identification without knowing the coefficient matrix  $\mathbf{A}$ . First, based on Eq. (8), we can estimate the fault geometry vector when a single fault happens. Then fault geometry vectors can be collected to form a library. Finally, the fault diagnosis can be achieved through a matching between current variation patterns and the fault geometry vectors in the library. The detailed procedure is listed in the following section.

**2.3. The Procedure for Data-Driven Variation Source Identification**

The steps of the variation source identification are listed as follows and summarized in Fig. 2.

- S1: The multivariate product quality measurements are obtained in the data collection step. The sample size is assumed to be  $N$  and the dimension of the measurement is  $n$ .
- S2: Based on the quality measurement, we can calculate the sample covariance matrix. From this covariance matrix, the number of significant variation sources can be estimated and tested. The testing procedure will be discussed in Section 2.4.1.
- S3: From Step 2, if only one variation source exists in the system, then the eigenvector associated with the largest eigenvalue of  $\mathbf{S}_y$  is an estimate of the fault geometry vector associated with current variation source. If there are already some fault geometry vectors in the library, then we need to test if the current fault geometry vector is the same as one of

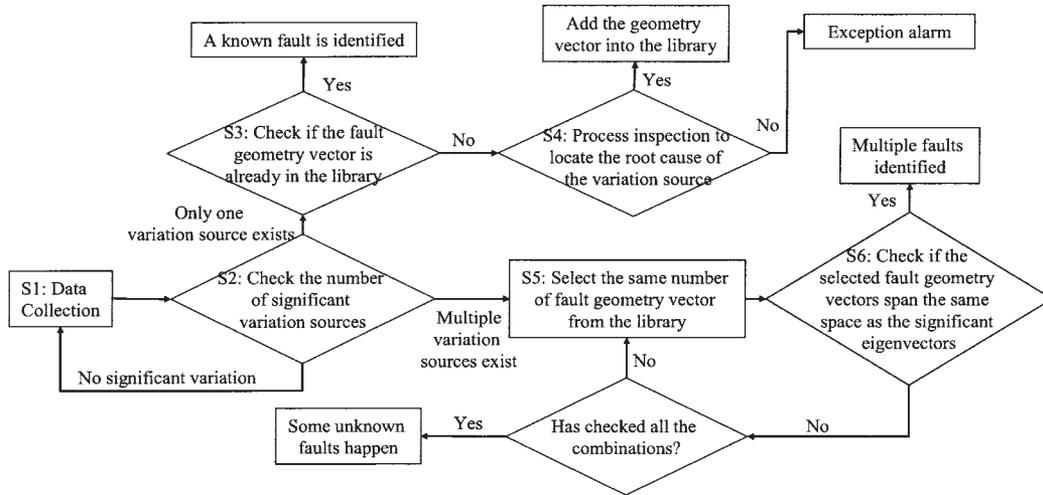


Figure 2. The procedure of variation source identification.

them (the testing procedure will be discussed in Section 2.4.2). If yes, we can determine that the current fault is the same as some previous fault.

- S4: From Step 3, if the current fault is different from any previous fault, we need to inspect the process and locate the physical root cause of the new fault. If the new fault is located, we will put the eigenvector associated with the largest eigenvalue of  $S_y$  into the library as the fault geometry vector of the new fault. If the new fault cannot be located, we will only generate an exception alarm and will not put the eigenvector into the library. This step is very critical to guarantee that each fault geometry vector in the library is corresponding to a real variation source. In practice, the eigenvector should not be put into the library if the physical root cause of the corresponding new fault is not identified.
- S5: From Step 2, if  $p$  ( $p > 1$ ) faults exist in the system, we can use the fault geometry vectors to identify which faults have happened. First we can randomly select  $p$  fault geometry vectors in the fault geometry library. If the total number of fault geometry vectors in the current library is smaller than  $p$ , then we can only say that some unknown faults have happened.
- S6: We determine which process faults exist in the current case by matching the  $p$  selected fault geometry vectors and the current variation patterns (the details are presented in Section 2.4.2). If there is a match, the same faults associated with the selected fault geometry vectors have happened. If no, we can select another set of fault geometry vectors from the library and check it again. If all the potential combinations of fault geometry vectors have been tested and no one spans the same space as current signifi-

cant eigenvectors, we cannot identify the individual faults in this case and can only claim some unknown faults happened.

- S7: After the variation sources have been identified, we can further estimate the variation magnitude of the process faults. One straightforward method is based on the equation  $\Sigma_y = \mathbf{A}\mathbf{D}\mathbf{A}^T + \sigma^2\mathbf{I}$ . It can be rearranged as  $\mathbf{A}\mathbf{D}\mathbf{A}^T = \Sigma_y - \sigma^2\mathbf{I}$ . After left multiplying  $\mathbf{A}^T$  and right multiplying  $\mathbf{A}$  on both sides of the equation, we have  $\mathbf{A}^T\mathbf{A}\mathbf{D}\mathbf{A}^T\mathbf{A} = \mathbf{A}^T(\Sigma_y - \sigma^2\mathbf{I})\mathbf{A}$ . Assuming  $\mathbf{A}^T\mathbf{A}$  is invertible and estimating  $\Sigma_y$ ,  $\mathbf{A}$  with  $S_y$  and  $\mathbf{A}_{ML}$ , respectively ( $\mathbf{A}_{ML}$  is the maximal likelihood estimator of  $\mathbf{A}$ ), then  $\mathbf{D}$  can be obtained as

$$\mathbf{D} \approx (\mathbf{A}_{ML}^T \mathbf{A}_{ML})^{-1} \mathbf{A}_{ML}^T (\Sigma_y - \sigma_{ML}^2 \mathbf{I}) \mathbf{A}_{ML} (\mathbf{A}_{ML}^T \mathbf{A}_{ML})^{-1}. \tag{9}$$

In this variation source identification procedure, two important testing procedures are required: (1) the number of significant eigenvalues of a covariance matrix (i.e., the number of faults in the system) and (2) determination of current faults based on the selected fault geometry vectors. The testing methods are discussed in Section 2.4.

REMARKS:

- It is common that only one fault happens in the manufacturing process [10, 15, 31]. In this paper, we will only store the fault geometry vector into the library only when one fault happens in the system. After a short learning period, the library that is accumulated from only single fault cases will be sufficient for future fault detection.

- The proposed procedure also requires process inspection before a fault geometry vector can be added to the library. Process inspection is very important in the proposed variation source identification procedure because the whole procedure is built up on statistical testing results. Due to the inherent uncertainty of statistical tests, there will always be errors in the conclusions of the statistical tests. We could have a false alarm or miss a detection in the statistical tests. These errors are inherent: we can only try to reduce them by increasing the sample size or developing more efficient test statistics, but we can never eliminate them. Thus, the process inspection is a necessary step in the procedure to eliminate the influence of statistical testing error. For example, we can depend on process inspection for decision making when a multiple match occurs; i.e., if multiple matches are found, process inspection is needed to make the final decision on which faults have occurred. In practice, we also need process inspection to identify the physical mechanism of the new fault and then eliminate it. After all, it is meaningless to store a fault geometry vector of an unknown fault into the library. One point that needs to be emphasized is that the proposed procedure only needs process inspection when a new fault (that has not occurred before) occurs. Compared with traditional process control techniques such as SPC, in which process inspection is needed for every fault, the process inspection efforts are significantly reduced.
- In this variation source identification procedure, we do not put the fault geometry vectors from multiple fault case in the library due to their rotational indeterminacy. According to Eq. (7), when multiple faults occur, the estimated fault geometry vector could be of an arbitrary rotation within the linear space spanned by the corresponding multiple column vectors of  $\mathbf{A}$  matrix. It cannot be utilized for fault diagnosis.

## 2.4. Testing Procedures Used in Variation Source Identification

### 2.4.1. Estimation of the Number of Faults

Different methods to solve this problem have been studied and compared through simulation by Apley and Shi [4]. This problem can be formulized as a hypothesis testing problem. Denote  $\mathbf{S}_y$  as the sample covariance matrix of the  $n$ -dimensional quality measurements. Then, to determine whether  $p$  faults exist in the system, we can test the hypothesis that  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq \sigma^2 = \lambda_{p+1} = \dots = \lambda_n$ , where  $\lambda_i$ ,  $i = 1 \dots n$ , are the eigenvalues of  $\mathbf{\Sigma}_y$ . Several

asymptotical testing procedures are available. In [4], the authors recommended the Akaike (AIC) and minimum description length (MDL) information criteria to estimate the number of fault. The AIC and MDL criteria are given as

$$AIC(l) = N(n-l)\log(a_l/g_l) + l(2n-l) \quad (10)$$

and

$$MDL(l) = N(n-l)\log(a_l/g_l) + l(2n-l)\log(N)/2, \quad (11)$$

where  $N$  is the sample size,  $n$  is the dimension of the quality measurement,  $a_l$  and  $g_l$  are the arithmetic mean and the geometric mean of the  $n-l$  smallest eigenvalues of  $\mathbf{S}_y$ , respectively. To use this criteria,  $AIC(l)$  and  $MDL(l)$  are evaluated for  $l = 0 \dots n-1$ . The estimated fault number  $p$  is chosen as the  $l$  that minimizes  $AIC(l)$  or  $MDL(l)$ , respectively. In this paper, we select  $MDL$  criterion to identify the number of faults.

### 2.4.2. Variation Source Identification Based on Matching of Fault Geometry Vectors

The variation source identification problem in Steps 3 and 5 can be formulated as follows. Consider the  $i$ th sample of quality measurements with sample size  $N$  and dimension of measurement  $n$ . The  $n$  by  $n$  sample covariance is denoted  $\mathbf{S}_i$ . We assume  $p$  ( $p \geq 1$  and  $p$  is less than the total number of potential faults in the system) faults exist in the process during this sampling period. The underlying process model during this period is denoted  $\mathbf{y} = \mathbf{A}_i \mathbf{f}_i + \boldsymbol{\varepsilon}$ . Clearly, the existing variation sources during this period are manifested by the columns of  $\mathbf{A}_i$ .

Without loss of generality, we can assume  $p$  ( $p \geq 1$ ) fault geometry vectors, denoted  $\{\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \dots, \tilde{\mathbf{a}}_p\}$ , are selected from the library in Step 4. These vectors are the estimates of the true fault geometry vectors, denoted  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$ . The matching of geometry vectors is actually to test if  $\mathbf{A}_i$  contains the same columns as  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$  regardless of the sequence of the columns (e.g.,  $\mathbf{A}_i = [\mathbf{a}_1 \ \mathbf{a}_2]$  and  $\mathbf{A}_i = [\mathbf{a}_2 \ \mathbf{a}_1]$  are considered to have the same columns). If the test result is yes, then we can claim the same faults associated with  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$  happened during the  $i$ th sampling period.

To conduct the test, one useful fact is that both the columns of  $\mathbf{A}_i$  and  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$  are subsets of the complete geometry vector set. Because the geometry vectors are assumed independent of each other,  $\mathbf{A}_i$  containing the same columns as  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$  is equivalent to  $\mathcal{R}(\mathbf{A}_i) = \mathcal{R}([\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p])$ , where  $\mathcal{R}(\cdot)$  is the range space of a matrix. Furthermore, it is known that for the model  $\mathbf{y} = \mathbf{A}_i \mathbf{f}_i + \boldsymbol{\varepsilon}$ , the  $p$  eigenvectors of  $\mathbf{\Sigma}_y$  that associ-

ated the  $p$  largest eigenvalues span the same space as  $\mathcal{R}(\mathbf{A}_i)$  if the covariance of  $\mathbf{f}_i$  is a diagonal matrix and covariance of  $\boldsymbol{\varepsilon}$  is in the form of  $\sigma^2\mathbf{I}$  [4, 24]. Therefore, the testing if  $\mathbf{A}_i$  containing the same columns as  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$  is equivalent to test if  $\mathcal{R}([\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_p]) = \mathcal{R}([\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_p])$ , where  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$  are the  $p$  eigenvectors of  $\boldsymbol{\Sigma}_y$  associated with the  $p$  largest eigenvalues.

In practice, both the true fault geometry vectors  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\}$  and the population eigenvectors  $\{\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_p\}$  are unknown. A statistical testing procedure using the estimates  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  and the eigenvectors of the sample covariance matrix  $\mathbf{S}_i$  need to be developed. Although not the same, the current testing problem is related to the statistical testing of common eigenspace, which is formulated as follows.

Given two  $n$  by  $n$  population covariance matrices  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$  and their samples  $\mathbf{S}_1$  and  $\mathbf{S}_2$ , we assume the eigenvalue/eigenvector pairs of  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$  are  $(\lambda_{11}, \mathbf{v}_{11}), (\lambda_{12}, \mathbf{v}_{12}) \dots (\lambda_{1n}, \mathbf{v}_{1n}), \lambda_{11} \geq \lambda_{12} \geq \dots \geq \lambda_{1n}$ , and  $(\lambda_{21}, \mathbf{v}_{21}), (\lambda_{22}, \mathbf{v}_{22}), \dots, (\lambda_{2n}, \mathbf{v}_{2n}), \lambda_{21} \geq \lambda_{22} \geq \dots \geq \lambda_{2n}$  respectively. The testing of common eigenspace is to test if the eigenvectors of  $\boldsymbol{\Sigma}_1$  and  $\boldsymbol{\Sigma}_2$  associated with first  $p$  largest eigenvalues span the same subspace, i.e., to test if  $\mathcal{R}(\mathbf{L}) = \mathcal{R}(\mathbf{M})$ , where  $\mathbf{L} \equiv [\mathbf{v}_{11} \ \mathbf{v}_{12} \ \dots \ \mathbf{v}_{1p}]$  and  $\mathbf{M} \equiv [\mathbf{v}_{21} \ \mathbf{v}_{22} \ \dots \ \mathbf{v}_{2p}]$ . Several techniques have been developed for this testing problem [7, 19, 32, 33]. These testing procedures are generic and powerful. To conduct these tests, a sample from each of the two populations is required. However, in the variation source identification problem, the sample from the first population is not always available, particularly for the multiple faults case. Instead, only an estimate of the columns of the coefficient matrix  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  is available. Although we might artificially generate a sample covariance from the selected  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  and the estimated measurement noise, the impact of this operation on the available testing procedures is not clear. Some authors developed analytical procedure for testing if certain eigenvectors of a covariance matrix span a subspace of the range space of a known matrix [2, 13, 35]. However, in our problem, the true values of fault geometry vectors are unknown. Therefore, those techniques cannot be directly applied in our case.

In this paper, we use an intuitive test statistics that can be easily accepted by practitioners. The basic idea is to use the angle between two subspaces as a measurement of the closeness of two subspaces. A similar test statistic has also been adopted by Krzanowski [25, 26]. The following proposition gives the definition and the calculation method of this angle.

**PROPOSITION:** Given an arbitrary vector  $\mathbf{v} \in \mathcal{R}(\mathbf{L})$ , then the largest angle between  $\mathbf{v}$  and its closest vector in  $\mathcal{R}(\mathbf{M})$ , where  $\mathbf{L}$  and  $\mathbf{M}$  are sets of  $p$  orthonormal bases, is given by  $\cos^{-1}(\sqrt{\lambda_p})$ ,  $\lambda_p$  is the smallest eigenvalue of  $\mathbf{L}^T\mathbf{M}\mathbf{M}^T\mathbf{L}$ .

**PROOF:** The proof is similar to the proof of theorem 1 in [25]. A vector of arbitrary direction in  $\mathcal{R}(\mathbf{L})$  can be expressed as  $\mathbf{b} = \mathbf{L} \cdot \mathbf{a}$ , where  $\mathbf{a}$  is an arbitrary unit vector. The closest vector to  $\mathbf{b}$  in  $\mathcal{R}(\mathbf{M})$  is its projection onto  $\mathcal{R}(\mathbf{M})$ , i.e.,  $\mathbf{M}\mathbf{M}^T\mathbf{b}$ . Therefore, the angle between these two are  $\cos^2\theta = \mathbf{b}^T\mathbf{M}\mathbf{M}^T\mathbf{b} = \mathbf{a}^T\mathbf{L}^T\mathbf{M}\mathbf{M}^T\mathbf{L}\mathbf{a}$ . Since  $\mathbf{L}^T\mathbf{M}\mathbf{M}^T\mathbf{L}$  is symmetric, the smallest  $\cos^2\theta$  (i.e., the largest  $\theta$ ) is given by the smallest eigenvalue of  $\mathbf{L}^T\mathbf{M}\mathbf{M}^T\mathbf{L}$  [21]. This proves the result.  $\square$

Based on this proposition, this largest angle can be used to represent the difference between the subspaces spanned by  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  and the first  $p$  eigenvectors of  $\mathbf{S}_i$ . The major advantage of this test statistic is that it is computationally simple and possesses clear geometric explanation. Practitioners can easily visualize this test statistic and link it with physical explanations. The disadvantage of this test statistic is that the analytical information about the sampling behavior of this statistic is absent. Hence, no analytical expressions for critical values of this statistical testing are available. However, similar to Krzanowski [26], the Monte Carlo simulation method can be used to provide approximations of the critical values of this test.

Sample uncertainties exist in both the selected geometry vectors from the library  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  and the eigenvectors of  $\mathbf{S}_y$ . In the simulation, we use two loops to simulate and characterize the sample uncertainties. The inner loop starts with a fixed set of fault geometry vectors  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$ . For this set of vector, replications of  $\mathbf{S}_i$  under the null case (i.e., the same  $p$  faults associated with the current  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  happen in the process when  $\mathbf{S}_i$  is simulated) are generated and the angle between the current  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  and the first  $p$  eigenvectors of  $\mathbf{S}_i$  is calculated. If a sufficient number of replications is produced, the main features of the sampling distribution of the angle under given simulation conditions can be captured. The percentile values (such as 90, 95, and 99% points) of the sample distribution of the angle are then obtained. Clearly, these percentile values are obtained conditioning on a fixed set of  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$ . In the outer loop, replications of the fault geometry vectors  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  are simulated (the vector  $\hat{\mathbf{a}}_i$  is actually the eigenvector associated with the largest eigenvalue of  $\mathbf{S}_y$  when only the  $i$ th fault exists in the system). For each replicate of  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$ , percentile values of the angle are obtained from the inner loop. The final percentile values of the sample distribution of the angle are obtained by averaging the percentile values associated with the replicates of  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$ . For example, if denote  $\theta_{95\%,i}$  is the 95th percentile value of the angle for the  $i$ th replicate of  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$ , then the final 95th percentile value of the angle is taken as  $\theta_{95\%} = \frac{1}{M} \sum_{i=1}^M \theta_{95\%,i}$ , where  $M$  is the total number of replications of  $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_p\}$  in the simulation.

In the simulation, the selection of  $\mathbf{A}$  is a difficult problem because the choices and settings of the coefficient matrix  $\mathbf{A}$ , and hence the population covariance matrix  $\Sigma_{\mathbf{y}}$ , are infinite. However, because we can always transform  $\Sigma_{\mathbf{y}}$  into a diagonal form through a linear variable transformation,  $\Sigma_{\mathbf{y}}$  can simply put in a diagonal form with decreasing diagonal elements. Krzanowski [26] adopts the same simplification in his simulation.

With this simplification, the parameters required in the simulation are the sample size  $N$ , dimension of the measurement  $n$ , number of the fault  $p$ , and the diagonal elements of  $\Sigma_{\mathbf{y}}$ . Without loss of generality, we assume the  $n$  diagonal elements of  $\Sigma_{\mathbf{y}}$  are given by  $\{\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2, \sigma_e^2, \dots, \sigma_e^2\}$ . It has been pointed out by some authors [2, 19] that if the ratio between  $\sigma_p^2$  and  $\sigma_e^2$  is small, the associated eigenvectors of the sample covariance matrix are not stable. Considering this fact, we take  $\sigma_p^2/\sigma_e^2$  as a parameter and call it the “variation ratio (VR)” of the fault in the simulation. When this ratio is large, the contrast between the process fault and the measurement noise is large. Hence, we should expect good diagnostic accuracy. On the other hand, when this ratio is small, the process fault is “blurred” by the measurement noise and the diagnosis accuracy will be deteriorated. In practice, because the variance of measurement noise ( $\sigma_e^2$ ) can be obtained through a gauge capability analysis and  $\sigma_p^2$  can be viewed as “tolerance” level on the variation sources that are often specified by design requirements (i.e., the minimum variation level of a variation source to be treated as significant), the boundaries of the VR ratio can often be obtained. Besides the VR ratio, it is also pointed out by Krzanowski [26] that the ratio between  $\sigma_1^2$  and the trace of  $\Sigma_{\mathbf{y}}$  will also impact on the sample uncertainty of the eigenvectors. In this paper, we denote  $C \equiv \sigma_1^2/\text{tr}(\Sigma_{\mathbf{y}})$ . When only one fault happens ( $p = 1$ ),  $C$  is automatically determined by our definition of  $\sigma_1^2 = \text{VR} \times$

$\sigma_e^2$  since  $C = \frac{\sigma_1^2}{\sigma_1^2 + (n-1) \times \sigma_e^2}$ . Thus the  $C$  parameter is excluded in the single fault case. When multiple faults happen ( $p \geq 2$ ), we will include the  $C$  parameter to show how  $C$  will impact the critical angle. We will present three cases ( $C = 75, 50$ , and  $33\%$ ) in this paper.

In summary, the selected parameters for the simulation include  $N, n, p, \text{VR}, \sigma_e^2$ , and  $C$  (when  $p \geq 2$ ). For each loop, 1000 replicates are generated. The results are summarized in the Appendix.  $\bar{\theta}$  and  $\sigma_{\theta}$  are the mean and standard deviation of the angle, respectively. The last three columns present the 90, 95, and 99% points of this angle. In the simulation, the values of the parameters are selected for some typical engineering applications. For parameter values that are not provided in the tables, a similar simulation can be easily conducted to estimate the corresponding critical values. From the results listed in the Appendix and some further simulations, the following observations can be made.

- (1) Not surprisingly, the sample size ( $N$ ) and the VR ratio ( $\sigma_p^2/\sigma_e^2$ ) impact the angle significantly. Our simulation results showed that this method worked well when the sample size is at least 50 and the VR ratio is greater than 50. This VR ratio is in a reasonable range, which can be justified as follows. According to Montgomery [30], a  $\mathbf{P/T}$  (precision to tolerance) ratio of 0.1 or less often is adequate gage capability. For common process capability ratio (PCR)  $C_p = 1$ , a typical process will have  $\sigma_p/\sigma_e \geq 10$ , which corresponds to  $\text{VR} \geq 100$ .
- (2) Another significant factor impacting the angle is the number of faults in the system. A large number of variation sources in the system will increase this angle and hence reduce the accuracy of this identification method. Therefore, we need to be cautious when the number of faults in the system is large. However, it is worth noting that this is not a serious limitation of this method in practice because even the system has quite a few potential faults; the number of faults that happen simultaneously is often limited.
- (3) The increase of the quality measurement ( $n$ ) will usually lead to a larger critical angle. The magnitude of the noise ( $\sigma_e^2$ ) also influences this angle. However, the influence seems insignificant.
- (4) The parameter  $C$  did not show significant impact under normal engineering assumptions ( $N \geq 50$  and  $\text{VR} \geq 50$ ). This effect is slightly larger when  $N$  is close to 50 and  $p = 2$ . However, the maximal variation of the critical angle caused by  $C$  is still less than 2 degrees, which can be ignored in many cases in practice. The impact of  $C$  will be “diluted” as  $p$  increases.

## 2.5. False Alarm Rate of the Proposed Variation Source Identification Procedure

For the proposed procedure in Section 2.4, we will study the false alarm rate for the following two cases: (a) when there is no fault in the system, our procedure finds at least one fault; (b) when there is one fault in the system, we incorrectly identify it as a different fault.

In case (a), the false alarm rate is the probability that our method will falsely claim that at least one fault occurred when the system is actually under normal working conditions (no fault occurred). In this case, the false alarm rate is determined by the MDL test. If the MDL test falsely identified a fault that actually does not exist, then this is a false alarm. Monte Carlo simulation is used to evaluate the false alarm rate. We used the linear model  $\mathbf{y} = \mathbf{A}\mathbf{f} + \boldsymbol{\varepsilon}$  to generate the sample covariance matrices and use MDL test to estimate the number of faults in the system. A typical

**Table 1.** False alarm rate for the fault identification procedure.

$N$	$n$	VR	Type I error	$N$	$n$	VR	Type I error
50	10	50	0.1016	100	10	50	0.1163
		100	0.0503			100	0.0552
		150	0.0207			150	0.0221
		200	0.0132			200	0.0127
	20	50	0.1419		20	50	0.1804
		100	0.0635			100	0.0673
		150	0.0244			150	0.0275
		200	0.0102			200	0.008
	30	50	0.188		30	50	0.2489
		100	0.0658			100	0.0783
		150	0.0246			150	0.0285
		200	0.0089			200	0.0064

parameter setting in practice is selected in the simulation: the maximal number of faults in the system is five faults; the dimension of  $\mathbf{y}$  is 10; the variance of the noise is  $\sigma_{\varepsilon}^2 = 0.01^2$ , and all the eigenvalues of  $\Sigma_{\mathbf{r}}$  are equal to  $0.005^2$ , which corresponds to the situation when no fault occurred in the system. Without loss of generality, we also assume that all the columns of matrix  $\mathbf{A}$  are orthogonal to each other. In total, 10,000 cases are simulated. For sample size  $N = 50$ , the false alarm rate is very small (0.0005), indicating that the MDL test is quite effective in fault detection.

In case (b), we already know there is one fault in the system from which the new sample is collected. The false alarm rate is the probability of incorrectly identifying the fault type. The overall false alarm rate of the diagnosis includes both the errors in the MDL tests (the probability that we falsely estimate the total number of faults in the system) and the errors in the hypothesis testing for fault geometry vector matching (the probability that we did not correctly identify the fault when the fault actually occurred). It is very difficult to obtain analytical results of the overall false alarm rate. To demonstrate the effectiveness of the proposed procedure, Monte Carlo simulation is used again. We also include the VR ratio in the simulation. For the sake of simplification, we also assume that all the sample covariance matrices and the fault signature matrices share the

sample size  $N$  and VR in each case. However, this is not a limitation of the testing procedure. In total, 10,000 cases are simulated. In each case, two samples for the same fault will be generated. One sample will be treated as the fault geometry vector while the other one will be treated as the new sample to be identified. If the testing procedure cannot correctly estimate the number of faults in either one of these two samples, it will be marked as an error case. If the testing procedure estimates an angle that is greater than the corresponding 99% critical angle given in the Appendix, it will also be marked as an error case. The false alarm rate for the whole testing procedure is calculated by the total number of error cases divided by 10,000. The false alarm rate for the whole fault identification procedure is presented in Table 1.

#### REMARKS:

- A larger fault magnitude (represented by VR given in Table 1) generally will have a smaller false alarm rate in the fault diagnosis. This is reasonable since larger VR value tends to make the variation sources stand out from the noises.
- As the dimension of quality measurements  $n$  increases, the false alarm rate increases as well when VR value is not extremely large.

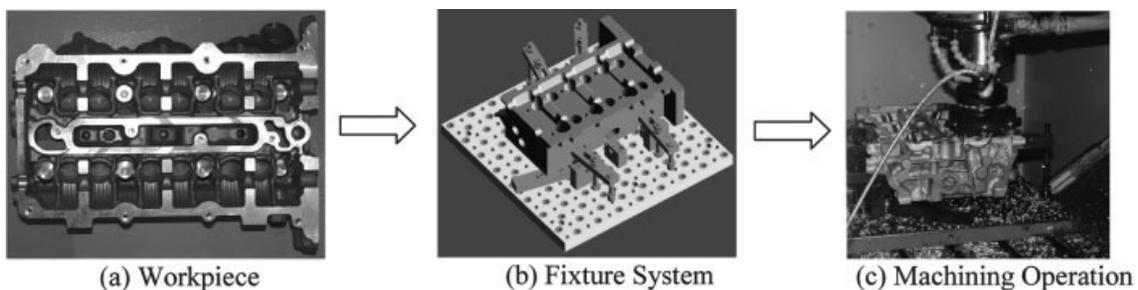
It should be pointed out that it is unsafe to draw very general conclusions from one simulation case. However, because the simulation is a typical case in practice, it is reasonable to conclude that the proposed method is an effective method in many engineering applications.

The procedure of variation source identification is presented in this section. A case study is presented in the next section to demonstrate the application of this procedure and its effectiveness.

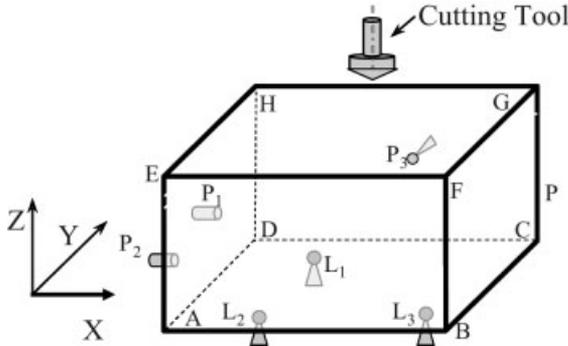
### 3. CASE STUDY

#### 3.1. Introduction to the Process

In this case study, a machining process is adopted to illustrate the effectiveness of this method. To machine a workpiece, we need first to locate the workpiece in a fixture



**Figure 3.** Illustration of a machining process. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]



**Figure 4.** A typical 3-2-1 fixturing configuration. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

system and then mount the fixture system on the working table of the machining center. The position of the cutting tool during cutting is calibrated with respect to the working table. This procedure is shown in Fig. 3.

In Fig. 3, the workpiece is an automotive engine component, the fixture system is of 3-2-1 configuration, and the machine tool is a vertical machining center (only the cutting tool and the working table of this machining center are shown in Fig. 3). The 3-2-1 fixturing setup is widely used in practice, in which there are  $3 + 2 + 1 = 6$  locating pins. The position and orientation of the workpiece are fixed in space with respect to the fixture if the workpiece touches these six pins.

Figure 4 illustrates a typical 3-2-1 fixturing system.  $L_1 \sim L_3$  and  $P_1 \sim P_3$  represent the six locating pins. By requiring surface ABCD in Fig. 4 to touch  $L_1 \sim L_3$ , the translational motion in the Z direction and the rotational motion in the X and Y directions are restrained. Similarly, surface ADHE constrains the translational motion in the X direction and the rotational motion in the Z direction by touching  $P_1$  and  $P_2$ . Surface DCGH constrains the translational motion

in the Y direction by touching  $P_3$ . Therefore, all six degrees of freedom associated with the workpiece are constrained by these three surfaces and the corresponding locating pins.

The cutting tool path is calibrated with respect to the machine coordinate system XYZ. Clearly, an error in the position of locating pins will cause a geometric error in the machined feature. Suppose that we mill a slot on surface EFGH in Fig. 4. If  $L_1$  is higher than its nominal position, the workpiece will be tilted with respect to XYZ. However, the cutting tool path is still determined with respect to XYZ. Hence, the bottom surface of the finished slot will not be parallel to the surface (ABCD). By measuring the position and orientation of the resulting surface with respect to surface ABCD, the faulty locating pin can be identified. Further, the magnitude of the fault can be estimated.

For the sake of simplicity, we only consider the identification of locating pin faults of a simple machining operation in this case study. The machining operation is illustrated in Fig. 3c. In this operation, the cover face of the engine component is milled. The quality measurements for this operation are the deviations of 15 points on the cover face from their nominal values. The relationship between the quality measurements and the errors in the locating pins can be described by  $\mathbf{y} = \mathbf{A}\mathbf{f} + \boldsymbol{\varepsilon}$ . In this model,  $\mathbf{y}$  is a 15 by 1 vector that includes the measurement of 15 points. The unit of  $\mathbf{y}$  is millimeter.  $\mathbf{f}$  is a 3 by 1 vector that represents the error in three locating pins. The reason that only three locating pin is included in  $\mathbf{f}$  is that only three locating pins will influence the accuracy of the resulting surface in this simple operation.  $\boldsymbol{\varepsilon}$  is a 15 by 1 vector that represents the measurement noise. Coordinate measurement machine is the most common device in practice. Based on the normal accuracy of CMM, we select  $\sigma_{\boldsymbol{\varepsilon}} = 0.01$  mm. The process normal variation  $\sigma_{\boldsymbol{\varepsilon}}^2$  is around  $0.05^2$  mm<sup>2</sup>. Furthermore, for this particular cover face machining operation,  $\mathbf{A}$  can be obtained as

$$\mathbf{A} = \begin{bmatrix} -0.1968 & -0.2352 & -0.0971 & 0.0249 & 0.1879 & 0.3117 & 0.3607 & 0.4322 \\ 0.3933 & 0.2432 & 0.0894 & -0.0463 & -0.2277 & -0.3655 & -0.2864 & -0.1715 \\ 0.2154 & 0.3947 & 0.3946 & 0.3945 & 0.3943 & 0.3942 & 0.2699 & 0.0891 \\ & & & 0.4775 & 0.3708 & 0.2329 & 0.1240 & 0.0193 & -0.0299 & -0.1011 \\ & & & -0.0985 & 0.0427 & 0.1866 & 0.3079 & 0.4243 & 0.3325 & 0.2178 \\ & & & -0.0256 & -0.0464 & -0.0374 & -0.0373 & -0.0372 & 0.0991 & 0.2795 \end{bmatrix}^T$$

Please note that the columns of  $\mathbf{A}$  have been normalized. This model is validated in [39] and has been utilized for process fault identification in [37].

In this case study, however, this model is assumed unknown. Instead, we use this model to generate different cases to simulate the machining operation. The proposed variation source identification procedure is applied to these generated cases to illustrate its effectiveness.

### 3.2. Application of the Variation Source Identification Technique

To demonstrate the proposed variation source identification technique, the following cases in Table 2 are considered. In each case, the sample size  $N = 50$ , the dimension of the quality measurement  $n = 15$ , and  $\sigma_{\boldsymbol{\varepsilon}} = 0.01$ . In Table 2,  $\sigma_{f_1}$ ,  $\sigma_{f_2}$ , and  $\sigma_{f_3}$  are the standard deviation of  $\mathbf{f}(1)$ ,  $\mathbf{f}(2)$ , and  $\mathbf{f}(3)$ , respectively. From this data, we can see that

**Table 2.** Different cases considered.

Case no.	1	2	3	4	5	6	7	8
$\sigma_1$	0.005	0.08	0.005	0.005	0.005	0.09	0.005	0.1
$\sigma_2$	0.005	0.005	0.1	0.08	0.1	0.08	0.005	0.005
$\sigma_3$	0.005	0.005	0.005	0.005	0.11	0.005	0.1	0.08

in case 1, all the standard deviations are quite small. Hence, case 1 represents the normal condition without fixture error. This result is confirmed by MDL test results later in Table 4. The rest of the cases represent different faulty conditions.

Table 3 shows the eigenvalues of  $S_y$  for each case and Table 4 lists the MDL test results ( $MDL(l), l = 0, 1, \dots, 5$ ) for cases 1–8. The  $l$  that minimizes MDL for each case is the estimated result of the number of significant variation sources. The result is listed in the last row. It can be seen that the number of significant variation sources is correctly identified for each case.

Based on the number of faults identified, the associated eigenvectors for each case are listed in Table 5. If  $p$  ( $p = 1, 2$ ) significant variation sources exist in the system, then  $p$  eigenvectors associated with the first  $p$  largest eigenvalues are listed.

Since no previously identified fault geometry vector is available, in Case 2 we put the eigenvector ( $v_1$ ) that is associated with the largest eigenvalue into the fault geometry library and denote it  $\tilde{v}_1$ . Please note that  $\tilde{v}_1$  is similar to the first column of  $A$ . The magnitude of the variation can also be estimated by substituting  $A$  by  $\tilde{v}_1$  in (9). The estimated  $\sqrt{D}$  is 0.0714, which is close to the true value of 0.08.

In Case 3, we can obtain  $v_2$ . Since  $\tilde{v}_1$  is already in the library of fault geometry vector, we need first to check whether  $\tilde{v}_1$  and  $v_2$  represent the same fault. The angle between  $\tilde{v}_1$  and  $v_2$  can be calculated using the result of Theorem 2 by simply substituting  $L$  as  $\tilde{v}_1$  and  $M$  as  $v_2$ . The result is  $\theta = 61.42^\circ$ , which is much larger than the critical value. Therefore, we can claim  $\tilde{v}_1$  and  $v_2$  are different. It should also be noted that  $v_2$  is similar to the second column of  $A$ . Hence,  $v_2$  is also put into the library and is denoted  $\tilde{v}_2$ . Similar to Case 2, the magnitude of the fault can be estimated as  $\sqrt{D} = 0.0943$ .

In Case 4, again, a single fault is identified. The angles between the first eigenvector and  $\tilde{v}_1, \tilde{v}_2$  are obtained as

$\theta_{v_1} = 62.87^\circ$  and  $\theta_{v_2} = 3.85^\circ$ , respectively. Clearly,  $\theta_{v_1}$  is much larger than the critical value and  $\theta_{v_2}$  is within in the acceptance region (in this case,  $N = 50, n = 15, p = 1$ . By looking at Table A1 in the Appendix, the 99% critical value should be between  $5.84^\circ$  and  $7.63^\circ$ , depending on the actual VR value). Therefore, we can claim that the same fault as that happened in Case 3 happens in this case, which is consistent with the true situation. The magnitude of the fault can be estimated as  $\sqrt{D} = 0.09$ .

In Case 5, two faults exist. Since there are two fault geometry vectors ( $\tilde{v}_1$  and  $\tilde{v}_2$ ) in the library, we can check whether the same two faults happened. First we can use  $\tilde{v}_1$  and  $\tilde{v}_2$  to form a matrix  $\tilde{V} = [\tilde{v}_1 \ \tilde{v}_2]$ . By using the simple Gram–Schmidt procedure [20,27], two orthonormal bases for  $\mathcal{R}(\tilde{V})$ , denoted  $\tilde{v}'_1$  and  $\tilde{v}'_2$ , can be obtained. Using  $[\tilde{v}'_1 \ \tilde{v}'_2]$  as  $L$  and the matrix that consists of the first two eigenvectors of  $S_y$  as  $M$ , we can calculate the angle between the two subspaces based on the result of Theorem 2. The result is  $\theta = 86.94^\circ$ . Obviously, they do not span the same space. In this case, we can only say that some unknown faults happened.

For Case 6, following the same procedure as that in Case 5, we can obtain  $\theta = 7.06^\circ$ . Based on the data in Table A2 in the Appendix, the 99% critical value should be within  $7.87^\circ$  and  $11.06^\circ$  regardless of the actual VR ratio (as long as the actual VR satisfies  $50 < C < 100$ ) and  $C$  values. Therefore, we can claim that the same faults happened in this case as that happened in Case 2 and Case 3. Using  $A = [\tilde{v}_1 \ \tilde{v}_2]$ , the magnitude of the faults can be estimated based on Eq. (9). In this case,  $D$  will be a 2 by 2 matrix. If only the diagonal elements are used, we have  $\sqrt{D_{11}} = 0.15$  and  $\sqrt{D_{22}} = 0.08$ .

In Case 7, only one fault is detected. The angle between the eigenvector that is associated with the largest eigenvalue of  $S_y$  and the fault geometry vectors  $\tilde{v}_1$  and  $\tilde{v}_2$  are  $\theta_{v_1} = 84.53^\circ$  and  $\theta_{v_2} = 85.22^\circ$ , respectively. Clearly, a new fault happened. This eigenvector is put into the library and is denoted as  $\tilde{v}_3$ . The magnitude of the fault is estimated as  $\sqrt{D} = 0.1008$ .

In Case 8, two faults are detected. Since there are three fault geometry vectors in the library, we need to check the angles between the space spanned by the first two eigenvectors in this case and the space spanned by  $[\tilde{v}_1 \ \tilde{v}_2], [\tilde{v}_1 \ \tilde{v}_3]$ , and  $[\tilde{v}_2 \ \tilde{v}_3]$ . Using the same procedure as that used in Case 6, these angles can be obtained as  $\theta_1 = 89.68^\circ, \theta_2 = 5.25^\circ$ , and  $\theta_3 = 60.64^\circ$ ,

**Table 3.** The eigenvalues of  $S_y$  for each case ( $10^{-3} \text{ mm}^2$ ).

Cases	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$	$\lambda_7$	$\lambda_8$	$\lambda_9$	$\lambda_{10}$	$\lambda_{11}$	$\lambda_{12}$	$\lambda_{13}$	$\lambda_{14}$	$\lambda_{15}$
1	0.2423	0.1995	0.1729	0.1534	0.1266	0.1181	0.0959	0.0905	0.0824	0.0663	0.0595	0.0543	0.0472	0.0448	0.0302
2	5.5265	0.2165	0.1824	0.1464	0.1406	0.1289	0.1007	0.0945	0.0845	0.0679	0.0538	0.0460	0.0439	0.0343	0.0262
3	12.4105	0.1758	0.1713	0.1419	0.1242	0.1204	0.1102	0.1006	0.0901	0.0762	0.0694	0.0544	0.0445	0.0388	0.0327
4	6.2765	0.2100	0.1791	0.1547	0.1462	0.1093	0.1012	0.0936	0.0862	0.0695	0.0620	0.0521	0.0395	0.0347	0.0298
5	12.6120	11.4840	0.2248	0.2026	0.1502	0.1358	0.1230	0.1056	0.1000	0.0789	0.0695	0.0563	0.0503	0.0444	0.0278
6	23.6044	8.8845	0.1866	0.1742	0.1485	0.1231	0.1199	0.0881	0.0803	0.0706	0.0628	0.0549	0.0500	0.0353	0.0219
7	7.2109	0.1933	0.1524	0.1452	0.1299	0.1200	0.1018	0.0967	0.0794	0.0693	0.0663	0.0515	0.0481	0.0411	0.0227
8	11.4467	6.6512	0.2322	0.1802	0.1515	0.1370	0.1158	0.0918	0.0765	0.0718	0.0666	0.0517	0.0464	0.0408	0.0274

**Table 4.** The MDL testing results.

Cases	1	2	3	4	5	6	7	8
MDL(0)	123.3794	1025.1	1344.5	1291.4	2006.5	1941.4	1354	1515.9
MDL(1)	163.9793	176.7	151.1	185.8	1630.6	1231.7	172.4	986.4
MDL(2)	200.8788	209.1	190.9	219.9	197.3	235.7	205.7	199.8
MDL(3)	236.4687	238.1	226.7	251.5	230.4	262.3	240	231.9
MDL(4)	270.5744	267.9	260.1	281.3	263.1	296.5	272.2	261.5
MDL(5)	302.0241	299.2	291.4	308.2	290.9	325.1	302.3	287.2
No. of faults	0	1	1	1	2	2	1	2

respectively. Thus, we can conclude that the same faults that happened in Case 2 and Case 7 are found in this case. Similar to Case 6, the magnitude of the fault can be estimated as  $\sqrt{D_{11}} = 0.1045$  and  $\sqrt{D_{22}} = 0.0822$ .

From the above numerical studies, we can see that the variation source identification method developed in this paper is effective. One point that needs to be mentioned is that in most practical situations, pinpointing the faults is much more important than estimating the magnitude of the fault. In many cases, people only need to know what the fault is to eliminate it.

**4. CONCLUDING REMARKS**

This paper presented a variation source identification methodology. The major characteristic of this method is that it does not require a predefined model that links the system measurements and the variation sources. Instead, this method identifies the unique geometry vector for each fault only based on quality measurement data. The variation source identification is conducted based on the key fact of eigenspace decomposition of  $S_y$ . The procedure of testing the closeness of two spaces is also presented. A numerical case study is conducted to illustrate the effectiveness of this method. This method can be used in quick root cause identification of a manufacturing process, which will lead to product quality improvement and production downtime reduction and hence a remarkable cost reduction in manufacturing systems.

There are several open issues related to this method. The first issue is the assumption on the system noise term  $\epsilon$ . In this paper, we assume its variance is in the form of  $\sigma^2 I$ . Although many practical systems can satisfy this requirement (e.g., a

system is very close to linear and the same measurement device is used to measure all the system output), some systems (particularly the systems with severe nonlinearity) do not have this property. In that case, we have to assume a general form for the covariance matrix of  $\epsilon$ . How this general form of the covariance matrix structure will impact on the eigenspace of  $\Sigma_\epsilon$  is an interesting problem. Although it is known that the perturbation caused by  $\Sigma_\epsilon$  in the eigenvectors that associated with the significant eigenvalues of  $\Sigma_y$  is small [15], no rigorous sampling results are available. This problem is currently under investigation. The second issue is the diagnosability of this method. The angle between two eigenspaces is not only determined by the sampling uncertainty, but also determined by the nominal angle between them. If the nominal angle is too small, we might not be able distinguish them due to the sampling uncertainty. What condition  $A$  should satisfy so that we can uniquely identify each variance source by this method is an important and interesting problem. The third issue is related to the gradual building of the library of fault geometry vector library. In this paper, we only add the geometry vector associated with single fault in the library. Although it cannot be directly extended from single fault case due to the rotational indeterminacy, it might be interesting to extract more fault information if possible. Furthermore, when several fault geometry vectors are obtained for the same fault, how do we choose the fault geometry vector that may provide the best detection power? A straightforward way might be to combine all the fault geometry vectors for the same fault together and average them. The efficiency of this method is currently under investigation and the results will be reported in the near future.

**Table 5.** Important eigenvectors of  $S_y$  for each case.

Cases	Important eigenvectors														
2*	0.1923	0.2424	0.1008	-0.0389	-0.1768	-0.2985	-0.3708	-0.4325	-0.4872	-0.3383	-0.2712	-0.1198	0.0041	0.0598	0.0653
3*	0.4096	0.2611	0.0949	-0.0589	-0.2116	-0.3814	-0.2802	-0.1823	-0.1108	0.0460	0.1665	0.3108	0.4063	0.3231	0.2032
4	0.4048	0.2335	0.1038	0.0065	-0.1841	-0.4091	-0.2927	-0.1517	-0.0926	0.0264	0.1704	0.2854	0.4331	0.3223	0.2228
5	0.4365	0.3169	0.1545	0.0305	-0.1333	-0.2986	-0.2391	-0.1490	-0.1195	0.0452	0.1918	0.3065	0.4000	0.3508	0.2606
	0.1581	0.3659	0.3792	0.3976	0.4022	0.4358	0.3053	0.1077	-0.0224	-0.0686	-0.0659	-0.0705	-0.0763	0.0534	0.2441
6	0.1507	-0.0081	0.0111	-0.0155	-0.0243	-0.0028	0.0726	0.2199	0.3912	0.3876	0.4001	0.4578	0.4023	0.2871	0.0734
	0.3651	0.2780	0.1109	-0.0327	-0.2477	-0.4050	-0.3701	-0.3508	-0.3116	-0.1678	-0.0166	0.1239	0.2548	0.2181	0.1929
7*	0.2236	0.3940	0.3915	0.4014	0.3958	0.3729	0.2814	0.0843	-0.0308	-0.0517	-0.0531	-0.0328	-0.0306	0.0898	0.2861
8	0.2272	0.3952	0.4202	0.3863	0.4069	0.3828	0.2165	0.0681	-0.0647	-0.0773	-0.0331	-0.0263	-0.0263	0.1022	0.2824
	0.1939	0.2781	0.1137	-0.0139	-0.1642	-0.3140	-0.3823	-0.4242	-0.4587	-0.3660	-0.2268	-0.1086	-0.0149	0.0327	0.1070

\* The eigenvectors of Case 2, 3, and 7 are denoted  $\tilde{v}_1$ ,  $\tilde{v}_2$ , and  $\tilde{v}_3$ , respectively.

APPENDIX

**Table A1.** Critical angles (degree) and summary statistics ( $\sigma_\varepsilon^2 = 0.01^2, p = 1$ ).

$N$	$n$	VR	$\bar{\theta}$	$\sigma_\theta$	90%	95%	99%	$N$	$n$	VR	$\bar{\theta}$	$\sigma_\theta$	90%	95%	99%	
50	10	50	4.01	0.28	5.29	5.71	6.53	100	10	50	2.79	0.17	3.65	3.93	4.45	
		100	3.14	0.33	4.08	4.38	4.95			100	100	2.20	0.22	2.84	3.04	3.41
		150	2.83	0.36	3.62	3.86	4.34			150	150	1.97	0.24	2.51	2.67	2.98
		200	2.61	0.38	3.30	3.51	3.93			200	200	1.83	0.26	2.30	2.44	2.71
	15	50	5.02	1.01	6.34	6.78	7.63		15	50	3.51	0.68	4.39	4.67	5.20	
		100	3.97	0.74	4.93	5.23	5.84			100	100	2.76	0.50	3.41	3.61	3.99
		150	3.55	0.62	4.34	4.59	5.08			150	150	2.47	0.42	3.01	3.18	3.49
		200	3.32	0.54	4.01	4.23	4.66			200	200	2.32	0.37	2.79	2.94	3.21
	20	50	5.90	0.29	7.26	7.70	8.59		20	50	4.11	0.19	5.01	5.29	5.82	
		100	4.63	0.36	5.60	5.92	6.54			100	100	3.25	0.23	3.90	4.10	4.49
		150	4.15	0.41	4.96	5.22	5.72			150	150	2.90	0.27	3.45	3.61	3.93
		200	3.88	0.42	4.59	4.81	5.24			200	200	2.71	0.29	3.19	3.33	3.61
30	50	7.31	0.31	8.74	9.21	10.15	30	50	5.10	0.19	6.03	6.31	6.87			
	100	5.77	0.40	6.79	7.12	7.77		100	100	4.03	0.24	4.70	4.91	5.30		
	150	5.14	0.44	5.98	6.25	6.78		150	150	3.58	0.26	4.14	4.31	4.63		
	200	4.80	0.46	5.53	5.76	6.22		200	200	3.37	0.30	3.85	4.00	4.28		

**Table A2.** Critical angles (degree) and summary statistics ( $\sigma_\varepsilon^2 = 0.01^2, p = 2$ ).

$N$	$n$	VR	C	$\bar{\theta}$	$\sigma_\theta$	90%	95%	99%	$N$	$n$	VR	C	$\bar{\theta}$	$\sigma_\theta$	90%	95%	99%		
50	10	50	75%	7.04	2.02	7.67	7.91	8.42	100	10	50	75%	5.82	1.88	6.18	6.32	6.60		
			50%	7.41	2.04	8.14	8.41	8.98				50%	6.00	1.82	6.43	6.59	6.91		
			33%	7.76	1.86	8.76	9.15	9.94				33%	6.10	1.74	6.71	6.93	7.38		
		100	75%	6.87	2.21	7.25	7.40	7.70			100	75%	5.49	1.96	5.72	5.81	5.98		
			50%	7.03	2.31	7.50	7.67	8.04				50%	5.66	1.93	5.94	6.04	6.24		
			33%	7.22	2.22	7.86	8.10	8.60				33%	5.70	1.87	6.09	6.23	6.52		
		150	75%	6.75	2.33	7.04	7.15	7.37			150	75%	5.46	1.95	5.63	5.70	5.82		
			50%	6.86	2.24	7.22	7.35	7.63				50%	5.56	1.96	5.78	5.86	6.01		
			33%	6.99	2.20	7.47	7.65	8.03				33%	5.66	1.90	5.95	6.05	6.27		
		15	50	75%	7.68	1.87	8.43	8.71			9.29	15	50	75%	6.08	1.80	6.52	6.68	7.00
				50%	8.13	1.89	8.96	9.26			9.89			50%	6.18	1.74	6.68	6.86	7.22
				33%	8.73	1.91	9.82	10.23			11.06			33%	6.58	1.68	7.25	7.48	7.95
	100			75%	7.24	2.09	7.70	7.87		8.21	100		75%	5.73	1.88	6.01	6.10	6.30	
				50%	7.40	2.11	7.96	8.16		8.56			50%	5.82	1.78	6.15	6.26	6.49	
				33%	7.64	2.04	8.39	8.66		9.22			33%	6.12	1.72	6.55	6.70	7.02	
	150		75%	7.15	2.18	7.50	7.63	7.88		150	75%	5.64	1.89	5.86	5.93	6.08			
			50%	7.33	2.14	7.76	7.91	8.22			50%	5.75	1.90	6.01	6.10	6.28			
			33%	7.41	2.08	7.98	8.19	8.62			33%	5.93	1.84	6.27	6.39	6.63			
	20		50	75%	8.24	1.77	9.08	9.39		10.03	20	50	75%	6.49	1.72	6.98	7.15	7.50	
				50%	8.62	1.75	9.53	9.86		10.54			50%	6.50	1.57	7.05	7.24	7.62	
				33%	9.28	1.58	10.45	10.87		11.75			33%	7.09	1.52	7.78	8.02	8.50	
		100		75%	7.59	2.01	8.11	8.29		8.68		100	75%	6.02	1.77	6.32	6.43	6.64	
				50%	7.92	2.03	8.52	8.73		9.16			50%	6.14	1.72	6.51	6.63	6.87	
				33%	8.36	1.90	9.15	9.43		10.02			33%	6.38	1.63	6.86	7.02	7.35	
150		75%	7.42	2.15	7.82	7.97	8.26	150	75%	5.92	1.84	6.16	6.24	6.39					
		50%	7.61	2.11	8.08	8.25	8.58		50%	5.94	1.76	6.23	6.32	6.51					
		33%	7.92	1.95	8.54	8.77	9.23		33%	6.19	1.76	6.56	6.69	6.95					
30		50	75%	9.12	1.58	10.13	10.49	11.23	30	50	75%	7.00	1.45	7.57	7.77	8.16			
			50%	9.55	1.58	10.58	10.95	11.70			50%	7.15	1.39	7.77	7.98	8.38			
			33%	10.28	1.40	11.54	11.98	12.90			33%	7.72	1.36	8.45	8.70	9.19			
	100		75%	8.24	1.85	8.86	9.08	9.53		100	75%	6.53	1.71	6.88	7.00	7.24			
			50%	8.60	1.82	9.28	9.52	10.00			50%	6.54	1.58	6.95	7.09	7.36			
			33%	8.99	1.65	9.89	10.21	10.86			33%	6.88	1.55	7.41	7.59	7.94			
	150	75%	7.91	1.93	8.38	8.54	8.88	150	75%	6.17	1.72	6.45	6.54	6.72					
		50%	8.17	1.88	8.71	8.90	9.28		50%	6.34	1.68	6.67	6.78	6.99					
		33%	8.58	1.75	9.29	9.54	10.05		33%	6.60	1.62	7.02	7.16	7.44					

## ACKNOWLEDGMENT

The financial support of this work is provided by NSF Award DMI-0322147. The authors appreciate the editor's and referees' valuable comments and suggestions.

## REFERENCES

- [1] R.D. Anderson and H. Rubin, "Statistical inference in factor analysis," in Proceedings of the Third Berkeley Symposium of Mathematical Statistics and Probability, Vol. 5, 1956, pp. 111–150.
- [2] T.W. Anderson, Asymptotic theory for principal component analysis, *Ann Math Stat* 34 (1963), 122–148.
- [3] D.W. Apley and J. Shi, Diagnosis of multiple fixture faults in panel assembly, *ASME J Manufact Sci Eng* 120 (1998), 793–801.
- [4] D.W. Apley and J. Shi, A factor-analysis methods for diagnosing variability in multivariate manufacturing processes, *Technometrics* 43 (2001), 84–95.
- [5] D.W. Apley and H.Y. Lee, Identifying spatial variation patterns in multivariate manufacturing processes: A blind separation approach, *Technometrics* 45(3) (2003), 220–234.
- [6] R.R. Barton and D.R. Gonzalez-Barreto, Process-oriented basis representations for multivariate process diagnostics, *Qual Eng* 9(1) (1996), 107–118.
- [7] R.J. Boik, Spectral models for covariance matrices, *Biometrika* 89 (2002), 159–182.
- [8] A.J. Camelio, S.J. Hu, and D.J. Ceglarek, "Modeling variation propagation of multi-station assembly systems with compliant parts," in Proceedings of the 2001 ASME Design Engineering Technical Conference, Pittsburgh, PA, September 9–12, 2001.
- [9] D. Ceglarek and J. Shi, Dimensional variation reduction for automotive body assembly, *J Manufact Rev* 8 (1995), 139–154.
- [10] D. Ceglarek and J. Shi, Fixture failure diagnosis for autobody assembly using pattern recognition, *ASME J Eng Ind* 188 (1996), 55–65.
- [11] M. Chang and D.C. Gossard, Computational method for diagnosis of variation-related assembly problem, *Int J Prod Res* 36 (1998), 2985–2995.
- [12] L.H. Chiang, E.L. Russell, and R.D. Braatz, *Fault detection and diagnosis in industrial systems*, Springer, London, 2001.
- [13] K. Choochaow, Dimension reduction in PCA: Likelihood-based methods. Ph.D. dissertation, Montana State University, Bozeman, MT, 2002.
- [14] Y. Ding, D. Ceglarek, and J. Shi, "Modeling and diagnosis of multistage manufacturing processes: Part I state space model," in Proceedings of the 2000 Japan/USA Symposium on Flexible Automation, July 23–26, 2000, Ann Arbor, MI, 2000JUSFA-13146.
- [15] Y. Ding, D. Ceglarek, and J. Shi, "Fault diagnosis of multistage manufacturing processes by using state space approach," in ASME Transactions, Journal of Manufacturing Science and Engineering, 2002, pp. 313–322.
- [16] Y. Ding, J. Shi, and D. Ceglarek, Diagnosability analysis of multistage manufacturing processes, *ASME J Dynam Syst Measure Contr* 124 (2002), 1–13.
- [17] Y. Ding, S. Zhou, and Y. Chen, A comparison of process variation estimators for in-process dimensional measurements and control, *ASME Trans J Dynam Syst Measure Contr* 127 (2005), 69–79.
- [18] D. Djurdjanovic and J. Ni, Linear state space modeling of dimensional machining errors, *Trans NAMRI/SME*, XXIX (2001), 541–548.
- [19] B.N. Flury, Two generalizations of the common principal component model, *Biometrika* 74(1) (1987), 59–69.
- [20] G.H. Golub and C.F. Van Loan, *Matrix Computation*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
- [21] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [22] Q. Huang, N. Zhou, and J. Shi, "Stream of variation modeling and diagnosis of multi-station machining processes," in Proceedings of IMECE 2000, MED-Vol. 11, pp. 81–88, November 5–10, 2000, Orlando, FL.
- [23] J. Jin and J. Shi, State space modeling of sheet metal assembly for dimensional control, *ASME J Manufact Sci Eng* 121 (1999), 756–762.
- [24] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, 5th ed., Prentice Hall, Upper Saddle River, NJ, 2002.
- [25] W.J. Krzanoski, Between-groups comparison of principal components, *J Am Stat Assoc* 74(367) (1979), 703–707.
- [26] W.J. Krzanoski, Between-groups comparison of principal components—Some sampling results, *J Stat Comp Simul* 15 (1982), 141–154.
- [27] D.C. Lay, *Linear Algebra and Its Applications*, 2nd ed., New York, Addison-Wesley, 1997.
- [28] R. Mantripragada and D.E. Whitney, Modeling and controlling variation propagation in mechanical assemblies using state transition models, *IEEE Trans Robot Automat* 15 (1999), 124–140.
- [29] D.C. Montgomery and W.H. Woodall, A discussion on statistically-based process monitoring and control, *J Qual Technol* 29 (1997), 121–162.
- [30] D.C. Montgomery, *Introduction to Statistical Quality Control*, 4th ed., John Wiley & Sons, 2001.
- [31] Q. Rong, D. Ceglarek, and J. Shi, Dimensional fault diagnosis for compliant beam structure assemblies, *ASME J Manufact Sci Eng* 122 (2000), 773–780.
- [32] J.R. Schott, Common principal component subspaces in two groups, *Biometrika* 75(2) (1988), 229–236.
- [33] J.R. Schott, Some tests for common principal component subspaces in several groups, *Biometrika* 78(4) (1991), 771–777.
- [34] M.E. Tipping and C.M. Bishop, Probabilistic principal component analysis, *J R Stat Soc B* 61(3) (1999), 611–622.
- [35] D.E. Tyler, Asymptotic inference for eigenvectors, *Ann Stat* 9 (1981), 725–736.
- [36] W.H. Woodall and D.C. Montgomery, Research issues and ideas in statistical process control, *J Qual Technol* 31 (1999), 376–386.
- [37] S. Zhou, Y. Chen, and J. Shi, Statistical estimation and testing for variation root cause determination of multistage manufacturing processes, *IEEE Trans Robot Automat* 1(1) (2004), 73–83.
- [38] S. Zhou, Y. Ding, Y. Chen, and J. Shi, Diagnosability study of multistage manufacturing processes based on linear mixed-effects models, *Technometrics* 45(4) (2003), 312–325.
- [39] S. Zhou, Q. Huang, and J. Shi, State space modeling of dimensional variation propagation in multistage machining process using differential motion vectors, *IEEE Trans Robot Automat* 19(2) (2003), 296–309.