IIE Transactions (2008) **40**, 919–930

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170801961404

Multiple fault signature integration and enhancing for variation source identification in manufacturing processes

LI ZENG, NONG JIN and SHIYU ZHOU*

Department of Industrial and Systems Engineering, University of Wisconsin, Madison, WI 53706, USA
E-mail: szhou@engr.wisc.edu

Received April 2006 and accepted November 2007

Signature matching is an increasingly popular method for variation source identification in manufacturing processes. In this method, the variation source is identified through matching the variation patterns of specific process faults, also called fault signatures, with the variation patterns in the newly collected quality data. There are situations in which a fault has occurred several times and consequently several signatures exist for the same fault. A technique is proposed that is able to integrate these multiple signatures together to enhance the accuracy of variation source identification. A linearly combined fault signature is constructed to increase the detection power of the fault identification. A numerical study is also presented to validate the effectiveness and robustness of the proposed method.

Keywords: Eigenprojection, fault signature, fault diagnosis, integrated signature

1. Introduction

Variation reduction during the manufacturing phase can play an essential role in the success of a manufacturing enterprise in today's globally competitive marketplace. As a critical step in variation reduction, variation source (also called process fault) identification has received significant attention in recent years. The available quantitative variation source identification methods can be roughly classified as analytical methods, which employ off-line physical models from engineering analyses, and data-driven techniques, which employ statistical models using only process data.

Most current analytical methods are based on linear models. Several linear models have been proposed to link the quality measurement data and the variation sources for complex multi-stage assembly processes (e.g., Jin and Shi (1999), Mantripragada and Whitney (1999), Ding *et al*. (2000), Camelio *et al*. (2003)) and machining processes (e.g., Djurdjanovic and Ni (2001), Zhou *et al*. (2003b), Loose *et al*. (2007)). These models can be put in the following generic form:

$$\mathbf{y} = \mathbf{A}\mathbf{f} + \varepsilon, \tag{1}$$

where $\mathbf{y}$ is the product quality measurement which is a vector that consists of the deviation of each quality characteristic from its nominal value, $\mathbf{A}$ is a constant coefficient matrix determined by process/product design, $\mathbf{f}$ is a vector that represents the process variation sources, and ε represents the system background noise including the process natural variation, unmodeled variation and measurement noises. Based on this linear model, direct estimation methods (Apley and Shi, 1998; Zhou *et al*., 2003a; Zhou *et al*., 2004; Ding *et al*., 2005) and pattern matching methods (Ceglarek and Shi, 1996; Rong *et al*., 2000; Ding *et al*., 2002; Li and Zhou, 2006; Li *et al*., 2006) were developed to identify the variation sources in the system. These analytical methods need a thorough understanding of the physics of the process to build the process model, which is usually very difficult, if not impossible, for a complex system (Chiang *et al*., 2001).

Unlike analytical methods, data-driven techniques focus on investigating the patterns in the extensive historical quality data sets and thus do not require comprehensive *a priori* knowledge of the manufacturing process. Factor analysis (Apley and Shi, 2001) and blind source separation techniques (Apley and Lee, 2003) have been used toward this purpose. Most recently, Jin and Zhou proposed a self-improving data-driven variation source identification procedure (Jin and Zhou 2006a, 2006b). In this procedure, the symptom of the current variation source(s), which is defined to be the sample covariance matrix of current quality measurement, is first extracted and then compared with the existing fault signatures in the fault library. If a match is found, then the current fault is identified as the fault associated with the matched signature. If no match is found, the process is inspected to identify the new process fault, and the associated symptom is added into the library as the signature of the new fault. The self-improving method is a data-driven method that can significantly reduce the

*Corresponding author

process inspection number in variation source identification: one needs to inspect the process only when new faults occur in the system.

This article focuses on an open issue in data-driven variation source identification, i.e., which signature we should use if multiple signatures exist in the fault library for the same fault? In practice, the library will be increasing as new faults occur. It is quite common that the same fault recurs multiple times and thus multiple fault signatures of the same fault will be found in the fault library as time goes on. These fault signatures provide an opportunity to improve the detection power of the corresponding signature matching. However, the refining of the fault signatures based on multiple signatures of the same fault has not been studied in the literature. In this article, we propose a signature refining method that linearly combines all the fault signatures of the same fault and then uses the integrated signature in signature matching. Optimal weights are identified that can minimize the variance involved in the statistical testing and thus maximize the detection power of the matching. The performance of the refined signature is evaluated in terms of type I error and type II error of the variation source identification procedure. The robustness of the proposed method with respect to the critical assumptions is investigated as well.

This paper is structured as follows. Some technical details of the self-improving variation source identification procedure are reviewed in Section 2 and followed by problem formulation of this study. The signature refining method is presented in Section 3. Section 4 demonstrates the effectiveness and robustness of the proposed method using numerical examples. Finally, the paper is concluded in Section 5.

2. Problem formulation

2.1. *Brief review of the data-driven variation source identification procedure*

Before presenting the problem formation of signature integration, we provide some background information regarding the data-driven variation source identification procedure. First, a linear relationship between the process faults and product quality, as shown in Equation (1), is assumed. Furthermore, the following assumptions are made.

(A1) $\mathbf{A}$ is an *unknown* $m \times k$ matrix. The columns of $\mathbf{A}$ are linearly independent.

(A2) $\mathbf{f}$ is a $k \times 1$ vector that follows multivariate normal distribution $N(\mathbf{0}, \mathbf{\Sigma_f})$, where the covariance matrix $\mathbf{\Sigma_f}$ is diagonal. The components of $\mathbf{f}$ are assumed independent because the process faults are often independent of each other. The fault occurrence is often unrelated with system background noises and so we also assume $\mathbf{f}$ is independent of ε. Moreover, the ith fault is said to occur if the (i, i)th element of $\mathbf{\Sigma_f}$ is non-zero.

(A3) ε is an $m \times 1$ vector following multivariate normal distribution $N(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{I}_m)$, where σ_ε^2 is a scalar and $\mathbf{I}_m$ is an $m \times m$ identity matrix.

These assumptions are not very restrictive and quite common in the existing variation source identification literature (Ceglarek and Shi, 1996; Apley and Shi, 1998; Rong *et al*., 2000). It should be noted that assumption (A3) is mainly for the purpose of simplifying the analytical derivation. It is noted that although (A3) is reasonable when the same measurement device is used for all the quality measurements and the measurement noise is the dominant part of the background noise, it does not precisely hold in some practical situations. In this article, (A3) is assumed to make the problem analytically tractable. Then, an extensive numerical study is conducted to investigate the robustness of the derived results when (A3) is violated.

We take covariance on both sides of Equation (1) and with assumptions (A1) to (A3), get

$$\mathbf{\Sigma_y} = \mathbf{A}\mathbf{\Sigma_f}\mathbf{A}' + \mathbf{\Sigma}_\varepsilon, \tag{2}$$

where $\mathbf{\Sigma_y}$ is the population covariance matrix of $\mathbf{y}$ and the superscript "′" means transposition. It is known that if *one* fault exists in the system, $\lambda_1 \gg \lambda_2 = \ldots = \lambda_m$, where λ_i, $i = 1, \ldots, m$, are eigenvalues of $\mathbf{\Sigma_y}$. That is, we will have a very large eigenvalue due to the fault and the rest comparatively small and equal eigenvalues (called small eigenvalues later) due to the system background noise. A straightforward analysis can show that the eigenvector associated with the largest eigenvalue of $\mathbf{\Sigma_y}$ will be equal to the column vector (if it is normalized) of $\mathbf{A}$ corresponding to the occurring fault (Apley and Shi, 1998; Johnson and Wichern, 2002; Ding *et al*., 2002). Based on this property, $\mathbf{S_y}$, the sample covariance matrix under a fault condition can be treated as a fault signature and is used for fault identification when the same fault recurs. Although the signature matching technique can be used to identify multiple faults simultaneously, we will focus on single fault cases (i.e., only one fault occurs in the system at a time) in this paper. The rationale is that the probability of simultaneous occurrence of multiple independent process faults is often very small compared with the single fault case.

The signature matching can be achieved through testing the coincidence of the eigenprojections of the symptom and the signature (Jin and Zhou, 2006b). Given a fault signature $\mathbf{S}_1$ and a symptom $\mathbf{S}_2$, assume their corresponding population covariance matrices are $\mathbf{\Sigma}_1$ (with eigenvalues being $\lambda_{11} \gg \lambda_{12} = \ldots = \lambda_{1m}$ and corresponding orthonormal eigenvectors being $\mathbf{q}_{11}, \mathbf{q}_{12}, \ldots, \mathbf{q}_{1m}$) and $\mathbf{\Sigma}_2$ (with eigenvalues being $\lambda_{21} \gg \lambda_{22} = \ldots = \lambda_{2m}$ and eigenvectors being $\mathbf{q}_{21}, \mathbf{q}_{22}, \ldots, \mathbf{q}_{2m}$) respectively. Accordingly, the (population) principal eigenprojections, which are associated with the largest eigenvalues, are $\mathbf{P}_1 = \mathbf{q}_{11} + \mathbf{q}'_{11}$ and $\mathbf{P}_2 = \mathbf{q}_{21}\mathbf{q}'_{21}$. Also, define $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$, and $\mathbf{P}_0$ is the principal

eigenprojection of $\mathbf{P}$. It's easy to get that, under the null case where $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ associate the same process fault, $\mathbf{P}_0 = \mathbf{P}_1 = \mathbf{P}_2$ and, equivalently, $(\mathbf{I}_m - \mathbf{P}_0)\mathbf{P}_i = \mathbf{0}$, $i = 1, 2$. With sample uncertainty, it is shown that the residual $\mathbf{v} = [\mathbf{v}'_1\mathbf{v}'_2]'$, where $\mathbf{v}_i = \text{vec}\{(\mathbf{I}_m - \hat{\mathbf{P}}_0)\hat{\mathbf{P}}_i\}$, $i = 1, 2$, has an asymptotic normal distribution with zero mean vector and covariance matrix $\mathbf{\Phi}$. Here, vec(.) is a stack operator that transforms a matrix to a column vector, and $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{P}}_i$ are estimates of $\mathbf{P}_0$ and $\mathbf{P}_i$ calculated based on $\mathbf{S}_1$ and $\mathbf{S}_2$. $\mathbf{\Phi}$ has the expression of (Schott, 1999):

$$\mathbf{\Phi} = \text{diag}(\mathbf{\Delta}_{11}, \mathbf{\Delta}_{22}) + \mathbf{C}, \tag{3}$$

where diag(.) places the elements in the parentheses along the diagonals sequentially. The terms involved in Equation (3) are expressed as follows:

1. $\mathbf{\Delta}_{ii} = \{\mathbf{I}_m \otimes (\mathbf{I}_m - \mathbf{P}_0)\}\mathbf{\Psi}_i\{\mathbf{I}_m \otimes (\mathbf{I}_m - \mathbf{P}_0)\}$, $i = 1, 2$, and
$$\mathbf{\Psi}_i = \mathbf{H}_i \,\text{var}\{\text{vec}(\mathbf{\Lambda}_i)\}\mathbf{H}_i, \tag{4}$$
where
$$\mathbf{H}_i = \sum_{l=2}^{m} (\lambda_{i1} - \lambda_{il})^{-1}(\mathbf{q}_{i1}\mathbf{q}'_{i1} \otimes \mathbf{q}_{il}\mathbf{q}'_{il}),$$
$$\mathbf{\Lambda}_i = \mathbf{S}_i - \mathbf{\Sigma}_i, \tag{5}$$
and var{vec($\mathbf{\Lambda}_i$)} is the covariance matrix of vec($\mathbf{\Lambda}_i$). When the sample size is large, vec ($\mathbf{\Lambda}_i$) has an asymptotic normal distribution and the specific expression of var{vec($\mathbf{\Lambda}_i$)} is $(N_i - 1)^{-1}(\mathbf{I}_{m^2} + \mathbf{K}_{mm})(\mathbf{\Sigma}_i \otimes \mathbf{\Sigma}_i)$, where N_i is the sample size of $\mathbf{S}_i$, "$\otimes$" denotes the Kronecker product and $\mathbf{K}_{mm}$ is a commutation matrix (Magnus and Neudecker, 1979).
2. $\mathbf{C}$ is a partitioned matrix, $[\mathbf{C}_{hi}]$, $h = 1, 2$, $i = 1, 2$, defined by
$$\mathbf{C}_{hi} = \{\mathbf{I}_m \otimes (\mathbf{I}_m - \mathbf{P}_0)\}\Bigg\{\sum_{f=1}^{2}(\mathbf{P}_h\mathbf{P}^+ \otimes \mathbf{I}_m)\mathbf{\Psi}_f \times (\mathbf{P}^+\mathbf{P}_i \otimes \mathbf{I}_m) - (\mathbf{P}_h\mathbf{P}^+ \otimes \mathbf{I}_m)\mathbf{\Psi}_i - \mathbf{\Psi}_h(\mathbf{P}^+\mathbf{P}_i \otimes \mathbf{I}_m)\Bigg\}\{\mathbf{I}_m \otimes (\mathbf{I}_m - \mathbf{P}_0)\}.$$

Based on this result, a generalized Wald statistic:

$$T = \mathbf{v}'\hat{\mathbf{\Phi}}^+\mathbf{v}, \tag{6}$$

can be developed that follows an asymptotic chi-squared distribution with degrees of freedom $v = m - 1$ under the null case (Schott, 1999). In Equation (6), $\hat{\mathbf{\Phi}}^+$ is a consistent estimator of $\mathbf{\Phi}^+$, the Moore–Penrose generalized inverse of $\mathbf{\Phi}$. To construct $\hat{\mathbf{\Phi}}^+$, $\hat{\mathbf{\Phi}}$ should first be obtained and $\hat{\mathbf{\Phi}}^+$ is just the Moore–Penrose generalized inverse of $\hat{\mathbf{\Phi}}$. One way to obtain $\hat{\mathbf{\Phi}}$ is deriving the expression of $\mathbf{\Phi}$ by population quantities ($\mathbf{\Sigma}$, $\mathbf{P}$, etc.) and then replacing the population quantities with their sample counterparts. The resulting formula to calculate T is (Schott, 1999):

$$T = \mathbf{u}'(\mathbf{F}\hat{\mathbf{\Theta}}\mathbf{F})^+\mathbf{u}. \tag{7}$$

The terms involved in Equation (7) are defined as follows:

1. Let $\hat{\mathbf{q}}_{1j}$ and $\hat{\mathbf{q}}_{2j}$, $j = 1, \ldots, m$, be the eigenvectors of $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively, corresponding to their jth largest eigenvalues, $\hat{\lambda}_{1j}$ and $\hat{\lambda}_{2j}$. $\hat{\mathbf{P}}_1 = \hat{\mathbf{q}}_{11}\hat{\mathbf{q}}'_{11}$, $\hat{\mathbf{P}}_2 = \hat{\mathbf{q}}_{21}\hat{\mathbf{q}}'_{21}$, $\hat{\mathbf{P}} = \hat{\mathbf{P}}_1^+\hat{\mathbf{P}}_2$. Let $\hat{\mathbf{q}}_j$ be the eigenvector of $\hat{\mathbf{P}}$ corresponding to its jth largest eigenvalue. $\hat{\mathbf{\Gamma}}_0 = (\hat{\mathbf{q}}_2, \hat{\mathbf{q}}_3, \ldots, \hat{\mathbf{q}}_m)$ and $\hat{\mathbf{\Gamma}}_1 = \hat{\mathbf{q}}_{11}$, $\hat{\mathbf{\Gamma}}_2 = \hat{\mathbf{q}}_{21}$. Then $\mathbf{u} = (\text{vec}(\hat{\mathbf{\Gamma}}'_0\hat{\mathbf{\Gamma}}_1)', \text{vec}(\hat{\mathbf{\Gamma}}'_0\hat{\mathbf{\Gamma}}_2)')'$.
2. $\hat{\mathbf{\Theta}} = \text{diag}(\hat{\mathbf{\Xi}}_1, \hat{\mathbf{\Xi}}_2) + \mathbf{V}$, where:
$$\hat{\mathbf{\Xi}}_i = \sum_{l=2}^{m} \frac{\hat{\lambda}_{i1}\hat{\lambda}_{il}}{n_i(\hat{\lambda}_{i1} - \hat{\lambda}_{il})^2}\hat{\mathbf{\Gamma}}'_0\hat{\mathbf{q}}_{il}\hat{\mathbf{q}}'_{il}\hat{\mathbf{\Gamma}}_0, \tag{8}$$
where $n_i = N_i - 1$, $i = 1, 2$. $\mathbf{V}$ is a partitioned matrix $[\mathbf{V}_{hi}]$, $h = 1, 2$, $i = 1, 2$, and $\mathbf{V}_{hi} = \sum_{f=1}^{2}(\hat{\mathbf{\Gamma}}'_h\hat{\mathbf{K}}^+\hat{\mathbf{\Gamma}}_1 \otimes \mathbf{I}_{m-1})\,\hat{\mathbf{\Xi}}_f(\hat{\mathbf{\Gamma}}'_f\hat{\mathbf{K}}^+\hat{\mathbf{\Gamma}}_i \otimes \mathbf{I}_{m-1}) - (\hat{\mathbf{\Gamma}}'_h\hat{\mathbf{K}}^+\hat{\mathbf{\Gamma}}_i \otimes \mathbf{I}_{m-1})\hat{\mathbf{\Xi}}_i - \hat{\mathbf{\Xi}}_h(\hat{\mathbf{\Gamma}}'_h\hat{\mathbf{K}}^+\hat{\mathbf{\Gamma}}_i \otimes \mathbf{I}_{m-1})$ where $\hat{\mathbf{K}} = \hat{\lambda}_1\hat{\mathbf{q}}_1\hat{\mathbf{q}}'_1$.
3. $\mathbf{F}$ is the eigenprojection matrix of $\hat{\mathbf{\Theta}}$ corresponding to its $m - 1$ largest eigenvalues.

This result provides a statistical testing to identify if a fault signature and a fault symptom match each other.

2.2. *Mathematical formulation of signature integration*

The above mentioned method can only match a symptom with one single signature. Due to the common existence of cases in practice where multiple signatures of the same fault are available, it is necessary to consider the following problem: given w fault signatures, $\mathbf{S}_{11}, \mathbf{S}_{12}, \ldots, \mathbf{S}_{1w}$ (with $\mathbf{\Sigma}_{1t}$, $t = 1, \ldots, w$, being the corresponding population covariance matrices) which associate the same process fault, and a fault symptom $\mathbf{S}_2$ (with $\mathbf{\Sigma}_2$ being its population covariance matrix), how to conduct the signature matching?

To efficiently utilize the available signatures, a natural idea is to construct a linearly combined signature:

$$\mathbf{S}_1^c = \sum_{t=1}^{w} \gamma_t \mathbf{S}_{1t}, \tag{9}$$

called the integrated signature, where the weights γ_t, $t = 1, \ldots, w$, s.t., $\sum_{t=1}^{w} \gamma_t = 1$, and extend the signature matching method in Section 2.1 to match $\mathbf{S}_2$ and $\mathbf{S}_1^c$. This requires that the statistic in Equation (6) be adjusted when $\mathbf{S}_1^c$ is used in place of $\mathbf{S}_1$.

Furthermore, it is easy to see that when $\mathbf{S}_1^c$ is used, the variance of the resulted asymptotic normal distribution (i.e., an adjusted version of $\mathbf{\Phi}$ as defined in Equation (3); without causing confusion, later $\mathbf{\Phi}$ is still used to represent the variance of the distribution when $\mathbf{S}_1^c$ is used) will be a function of γ_t, $t = 1, \ldots, w$. Since the detection power of a Wald test is in general inversely decided by the variance of

the asymptotic distribution, we can minimize $\mathbf{\Phi}$ and consequently maximize the detection power of the test through identifying the optimal set of weights. Here minimizing the matrix $\mathbf{\Phi}$ means minimizing an appropriately chosen scalar coefficient as a measure of $\mathbf{\Phi}$.

Specifically, the following questions need to be answered:

1. Are the statistical properties of $\mathbf{S}_1^c$ (e.g., its principal eigenprojection, its statistical distribution properties, etc.) equivalent to those of an individual signature (e.g., $\mathbf{S}_1$) and thus the statistic in Equation (6) can be adjusted using $\mathbf{S}_1^c$ in place of $\mathbf{S}_1$? The answer to this question is not obvious because an individual signature is a sample from a single population, while $\mathbf{S}_1^c$ is a combination of multiple samples from *different* populations (although they share the same eigenprojection associated with the largest eigenvalues).
2. If $\mathbf{S}_1^c$ is statistically equivalent to an individual $\mathbf{S}_1$, what are the optimal weights γ_t^*, $t = 1, \ldots, w$, such that $\mathbf{\Phi}$ is minimized, and what is the expression of the adjusted chi-squared statistic T in Equation (6) when the integrated signature constructed by the optimal weights, $\mathbf{S}_1^*$, is used?
3. Can we put the integration procedure into a simple iterative updating process? In other words, assume currently we have an integrated signature $\mathbf{S}_{1\ldots k}^*$ that is an integration of k individual signatures and in order to integrate these k individual signatures with a new signature, say, $\mathbf{S}_{1(k+1)}$, can we simply integrate $\mathbf{S}_{1\ldots k}^*$ with this new signature instead of considering all k individual signatures? If the answer is yes, then the signature integration process can be greatly simplified.

In the following section, these questions are discussed and answered.

3. Multiple fault signature integration and enhancing

3.1. *Equivalence of integrated and individual signatures in the hypothesis testing*

A close examination of the proof of Equation (6) (Schott, 1999) reveals that for the establishment of the asymptotic normality of $\mathbf{v}$, which is the basis of the testing, the only requirements for $\mathbf{S}_1$ and $\mathbf{S}_2$ are their independence and normality of $\text{vec}(\mathbf{\Lambda}_i)$, $i = 1, 2$. In that proof, first $\hat{\mathbf{P}}_i$ and consequently $\hat{\mathbf{P}}_0$ are expanded using the first-order Taylor formula which holds for all positive definite symmetric matrices (Tyler, 1981), and then, the expression of $\mathbf{v}_i$, $i = 1, 2$ is reached as a linear combination of $\text{vec}(\mathbf{\Lambda}_1)$ and $\text{vec}(\mathbf{\Lambda}_2)$. Because of the independence (following the fact that $\mathbf{S}_1$ and $\mathbf{S}_2$ are independent of each other) and normality of $\text{vec}(\mathbf{\Lambda}_i)$, $i = 1, 2$, it is decided that $\mathbf{v}_i$ and thus $\mathbf{v}$ are asymptotically normally distributed, and the variance, as defined in Equation (3), is obtained as a linear combination of $\text{var}\{\text{vec}(\mathbf{\Lambda}_1)\}$ and $\text{var}\{\text{vec}(\mathbf{\Lambda}_2)\}$. This implies that $\mathbf{S}_1$ can be replaced by any matrix $\mathbf{S}$ constructed from samples, which satisfies the following.

1. A positive definite and symmetric matrix $\mathbf{\Sigma}$, the nominal population covariance matrix corresponding to $\mathbf{S}$, can be constructed accordingly such that $\mathbf{S}$ is a consistent estimator of $\mathbf{\Sigma}$ and $\text{vec}(\mathbf{S} - \mathbf{\Sigma})$ has an asymptotic normal distribution.
2. $\mathbf{S}$ is independent of $\mathbf{S}_2$.
3. The principal eigenprojection of $\mathbf{\Sigma}$ is equal to that of $\mathbf{\Sigma}_1$.

Now let us turn to the properties of $\mathbf{S}_1^c$ and see if it satisfies the above listed conditions. First, from assumptions (A1) to (A3) listed in Section 2.1, we have the following.

1. All the population covariance matrices $\mathbf{\Sigma}_{1t}$, $t = 1, \ldots, w$ and $\mathbf{\Sigma}_2$, are positive definite.
2. All the sample covariance matrices $\mathbf{S}_{1t}$, $t = 1, \ldots, w$ and $\mathbf{S}_2$, are from independent samples generated by model (2) and thus they are statistically independent of each other.
3. Because we are looking at the multiple signatures of the *same* fault, we have $\mathbf{P}_{11} = \mathbf{P}_{12} = \ldots = \mathbf{P}_{1w}$, where $\mathbf{P}_{1t}$, $t = 1, \ldots, w$, is the principal eigenprojection of $\mathbf{\Sigma}_{1t}$.
4. Let $\lambda_{1t1}, \lambda_{1t2}, \ldots \lambda_{1tm}$ be the eigenvalues of $\mathbf{\Sigma}_{1t}$, $t = 1, \ldots, w$, and $\mathbf{q}_{1t1}, \ldots, \mathbf{q}_{1tm}$ are the corresponding orthonormal eigenvectors. Since only a single fault is considered, we have $\lambda_{1t1} \gg \lambda_{1t2} = \ldots = \lambda_{1tm} = \alpha_t$, i.e., the largest eigenvalue of each signature should be predominant among its eigenvalues and the small eigenvalues are equal to each other. These two properties are based on assumption (A3). The robustness of the developed method with respect to these two properties will be investigated in Section 4.2.

Let $\mathbf{\Sigma}_1^c = \sum_{t=1}^w \gamma_t \sum_{1t}$ be the nominal population covariance matrix corresponding to $\mathbf{S}_1^c$, and $\lambda_{11}^c, \ldots, \lambda_{1m}^c$ and $\mathbf{q}_{11}^c, \ldots, \mathbf{q}_{1m}^c$ are eigenvalues and eigenvectors of $\mathbf{\Sigma}_1^c$. Based on the above results, the following properties of $\mathbf{S}_1^c$ can be established.

(P_1) Since $\mathbf{\Sigma}_{1t}$, $t = 1, \ldots, w$, are all positive definite and symmetric, $\mathbf{\Sigma}_1^c$ is also positive definite and symmetric.

(P_2) Because $\mathbf{S}_{1t}$ is a consistent estimator of $\mathbf{\Sigma}_{1t}$, i.e., $\mathbf{S}_{1t} \to \mathbf{\Sigma}_{1t}$, as $N_{1t} \to \infty$ (Schott, 2005), where N_{1t} is the sample size of $\mathbf{S}_{1t}$, $\mathbf{S}_1^c = \sum_{t=1}^w \gamma_t \mathbf{S}_{1t} \to \sum_{t=1}^w \gamma_t \mathbf{\Sigma}_{1t} = \mathbf{\Sigma}_1^c$, meaning that $\mathbf{S}_1^c$ is also a consistent estimator of $\mathbf{\Sigma}_1^c$.

(P_3) According to the property of the "vec" operator, $\text{vec}(\mathbf{S}_1^c - \mathbf{\Sigma}_1^c) = \sum_{t=1}^w \gamma_t \text{vec}(\mathbf{S}_{1t} - \mathbf{\Sigma}_{1t})$. Because $\text{vec}(\mathbf{S}_{1t} - \mathbf{\Sigma}_{1t})$, $t = 1, \ldots, w$, are independent of each other (following the fact that $\mathbf{S}_{1t}$ are independent) and $\text{vec}(\mathbf{S}_{1t} - \mathbf{\Sigma}_{1t})$ has an asymptotic normal distribution, $\text{vec}(\mathbf{S}_1^c - \mathbf{\Sigma}_1^c)$ is also asymptotically normally distributed (Rencher, 2002).

(P_4) Because $\mathbf{S}_{1t}$, $t = 1, \ldots, w$, are all independent of $\mathbf{S}_2$, $\mathbf{S}_1^c$ is also independent of $\mathbf{S}_2$.

(P_5) $\mathbf{q}_{11}^c = \mathbf{q}_{111} = \mathbf{q}_{121} = \ldots = \mathbf{q}_{1w1}$. It follows that $\mathbf{P}_1^c = \mathbf{q}_{11}^c \mathbf{q}_{11}^{c'} = \mathbf{q}_{1t1}\mathbf{q}_{1t1}' = \mathbf{P}_{1t}$, $t = 1, \ldots, w$, where $\mathbf{P}_1^c$ and $\mathbf{P}_{1t}$ are the principal eigenprojections of $\mathbf{\Sigma}_1^c$ and $\mathbf{\Sigma}_{1t}$ respectively. This means that the principal eigenprojection of $\mathbf{\Sigma}_1^c$ is the same as that of every individual signature.

(P_6) $\lambda_{11}^c = \sum_{t=1}^{w} \gamma_t \lambda_{1t1}$.

Clearly, from these properties, we can see that $\mathbf{S}_1^c$ satisfies all the conditions to replace the individual signature $\mathbf{S}_1$ in the test statistics in Equation (6). It is also worth pointing out that $\mathbf{S}_1^c = \mathbf{S}_{1r}$, $r \in \{1, \ldots, w\}$, as $\gamma_r = 1$ and $\gamma_t = 0$, for $t \neq r$, indicating that the test using $\mathbf{S}_1^c$ covers all the tests involving any individual fault signature. Thus, optimal integration of signatures will lead to better or at least the equivalent performance compared with individual signatures.

3.2. *Optimal integration of individual signatures*

The previous section validates that the integrated signature can be used as a fault signature and compared with the fault symptom through rigorous hypothesis testing. This section will give the optimal weights for the integrated signature such that the variance of the corresponding asymptotic distribution, $\mathbf{\Phi}$, is minimized, and the adjusted chi-squared statistic when the integrated signature constructed by the optimal weights is used.

Clearly, when $\mathbf{S}_1^c$ is used in place of $\mathbf{S}_1$ in the statistical testing described in Section 2.1, all the expressions about $\mathbf{\Phi}$ have the same forms as defined in Section 2.1 except that $\mathbf{S}_1$ is replaced by $\mathbf{S}_1^c$, $\mathbf{P}_1$ is replaced by $\mathbf{P}_1^c$, etc. Thus, it is easy to see that $\mathbf{\Phi}$ is determined by $\mathbf{P}_1^c$, $\mathbf{P}_2$, $\mathbf{P}$, $\mathbf{P}_0$ and $\mathbf{\Psi}_i$, $i = 1, 2$. (Note that now $\mathbf{P}_2$, $\mathbf{P}$, $\mathbf{P}_0$ and $\mathbf{\Psi}_i$ refer to the adjusted version since $\mathbf{S}_1^c$ is used in place of $\mathbf{S}_1$.) Because $\mathbf{P}_2$ and $\mathbf{\Psi}_2$ are only related with the symptom $\mathbf{S}_2$ and $\mathbf{\Sigma}_2$ which cannot be controlled beforehand, and $\mathbf{P}_1^c$ and thus $\mathbf{P}$ and $\mathbf{P}_0$ are all fixed for given signatures, these quantities are not considered here. Consequently, $\mathbf{\Phi}$ is only related with $\mathbf{\Psi}_1$ and minimizing $\mathbf{\Phi}$ is equivalent to minimizing $\mathbf{\Psi}_1$ or more precisely, an appropriately chosen scalar measure of $\mathbf{\Psi}_1$. The following theorem gives the specific expression of $\mathbf{\Psi}_1$ when $\mathbf{S}_1^c$ is used and the optimal weights γ_t^*, $t = 1, \ldots, w$, such that $\mathbf{\Psi}_1$ is minimized.

Theorem 1. *When $\mathbf{S}_1^c$ given by Equation (9) is used in place of $\mathbf{S}_1$ in the statistical testing in Section 2.1, $\mathbf{\Psi}_1$ defined by Equation (4) has the following expression:*

$$\mathbf{\Psi}_1 \approx \frac{\eta}{\left(\lambda_{11}^c\right)^2} \mathbf{P}_1^c \otimes \left(\mathbf{I}_m - \mathbf{P}_1^c\right), \tag{10}$$

where $\eta = \sum_{t=1}^{w} \gamma_t^2 \varsigma_t$ and $\varsigma_t = \lambda_{1t1}\alpha_t / n_{1t}$, $n_{1t} = N_{1t} - 1$. $\mathbf{\Psi}_1$ is minimized as

$$\gamma_t^* = \frac{\omega_t}{\gamma_{1\ldots w}}, \tag{11}$$

where $\omega_t = n_{1t}/\alpha_t$ and $\gamma_{1\ldots w} = \sum_{r=1}^{w} \omega_r$.

Equation (10) shows that actually $\mathbf{\Psi}_1$ consists of a scalar coefficient part, $\eta/(\lambda_{11}^c)^2$, and a matrix part, $\mathbf{P}_1^c \otimes (\mathbf{I}_m - \mathbf{P}_1^c)$. According to property (P_5), the matrix part is fixed for given signatures under the null case. Consequently, minimizing $\mathbf{\Psi}_1$ is equivalent to minimizing the scalar coefficient. That is how Equation (11) is obtained. The proof is given in Appendix I. In practice, the small eigenvalue α_t is estimated by

$$\hat{\alpha}_t = \frac{1}{m-1} \sum_{j=2}^{m} \hat{\lambda}_{1tj}, \tag{12}$$

and accordingly, the optimal weights are estimated by $\hat{\gamma}_t^* = \hat{\omega}_t / \hat{\gamma}_{1\ldots w}$, $t = 1, \ldots, w$, where $\hat{\omega}_t = n_{1t}/\hat{\alpha}_t$ and $\hat{\gamma}_{1\ldots w} = \sum_{r=1}^{w} \hat{\omega}_r$.

From the theorem, we have the following observations.

1. To maximize the detection power of the statistical testing, the importance of each individual fault signature $\mathbf{S}_{1t}$, $t \in \{1, \ldots, w\}$ is determined by its sample size and small eigenvalues, as indicated by Equation (11). Since it has been pointed out that $\hat{\alpha}_t$ actually represents the average noise level of $\mathbf{S}_{1t}$, this result is intuitive because a larger sample size and lower noise level correspond with greater accuracy of the estimate of eigenprojection from a fault signature.
2. When $\alpha_1 \approx \alpha_2 \approx \ldots \approx \alpha_w$, $\gamma_t^* \approx n_{1t} / \sum_{r=1}^{w} n_{1r}$, meaning that the optimal weights are only related with sample sizes of signatures. In other words, the standard method of setting weights based solely on sample size works only when the noise level is similar for all the fault signatures. However, this is, obviously, a restrictive case because in practice it is common that the noise level changes from sample to sample due to different operators, measuring instruments and other physical conditions in which a sample is obtained. Thus, the weights we find are a more general solution for optimally integrating the multiple signatures.

Accordingly, the integrated signature constructed by the optimal weights is

$$\mathbf{S}_1^* = \sum_{t=1}^{w} \hat{\gamma}_t^* \mathbf{S}_{1t}. \tag{13}$$

Given the integrated signature $\mathbf{S}_1^*$ as defined in Equation (13) and the symptom $\mathbf{S}_2$, the hypothesis testing for matching between them becomes $\mathbf{H}_0$: $\mathbf{S}_1^*$ and $\mathbf{S}_2$ associate the same fault or $\mathbf{P}_1^* = \mathbf{P}_2$ vs. $\mathbf{H}_1$: $\mathbf{S}_1^*$ and $\mathbf{S}_2$ associate different faults or $\mathbf{P}_1^* \neq \mathbf{P}_2$. The corresponding formula to calculate the chi-squared statistic follows the same form as Equation (7),

that is

$$T^* = \mathbf{u}^*(\mathbf{F}^*\hat{\mathbf{\Theta}}^*\mathbf{F}^*)^+\mathbf{u}^*, \tag{14}$$

where $\mathbf{u}^*$, $\mathbf{F}^*$ and $\hat{\mathbf{\Theta}}^*$ are the counterparts of $\mathbf{u}$, $\mathbf{F}$ and $\hat{\mathbf{\Theta}}$, respectively, in Equation (7) with terms associated with $\mathbf{S}_1$ being replaced by those associated with $\mathbf{S}_1^*$. For example, $\hat{\mathbf{P}}_1^* = \hat{\mathbf{q}}_{11}^*\hat{\mathbf{q}}_{11}^*$, $\hat{\mathbf{P}}^* = \hat{\mathbf{P}}_1^* + \hat{\mathbf{P}}_2$ and $\hat{\mathbf{\Gamma}}_0^* = (\hat{\mathbf{q}}_2^*, \hat{\mathbf{q}}_3^*, \ldots, \hat{\mathbf{q}}_m^*)$, where $\hat{\mathbf{q}}_j^*$ is the eigenvector of $\hat{\mathbf{P}}^*$ corresponding to its jth largest eigenvalue. The only exception is that $\hat{\mathbf{\Xi}}_1^*$, which satisfies $\hat{\mathbf{\Theta}}^* = \text{diag}(\hat{\mathbf{\Xi}}_1^*, \hat{\mathbf{\Xi}}_2) + \mathbf{V}^*$, has an expression different from Equation (8):

$$\hat{\mathbf{\Xi}}_1^* = \sum_{l=2}^{m} \frac{\hat{\eta}^*}{\hat{\lambda}_{11}^{*2}} \hat{\mathbf{\Gamma}}_0^* \hat{\mathbf{q}}_{1l}^* \hat{\mathbf{q}}_{1l}^* \hat{\mathbf{\Gamma}}_0^*, \tag{15}$$

where $\hat{\eta}^* = \sum_{t=1}^{w} \hat{\gamma}_t^{*2} \hat{\varsigma}_t$ and $\hat{\varsigma}_t = \hat{\lambda}_{1t1}\hat{\alpha}_t / n_{t1}$. Note that, in the same way, we can calculate the chi-squared statistic for any linearly combined signature $\mathbf{S}_1^c$.

3.3. *Iterative updating procedure for signature integration*

To calculate the optimal integrated signature in Equation (13) and the statistic in Equation (14), all the w individual signatures from $\mathbf{S}_{11}$ to $\mathbf{S}_{1w}$ are needed. However, this integration can be significantly simplified through an iterative updating process. Specifically, denote $\mathbf{S}_{1\ldots k}^*$ and $\mathbf{S}_{1\ldots k+1}^*$ as the optimal integrated signatures as defined in Equation (13) for $\{\mathbf{S}_{11}, \mathbf{S}_{12}, \ldots, \mathbf{S}_{1k}\}$ and for $\{\mathbf{S}_{11}, \mathbf{S}_{12}, \ldots, \mathbf{S}_{1k}, \mathbf{S}_{1(k+1)}\}$, $k = 1, \ldots, w$, respectively, then we have the following iterative relationship:

$$\mathbf{S}_{1\ldots k+1}^* = \frac{\hat{\gamma}_{1\ldots k}}{\hat{\gamma}_{1\ldots k+1}} \mathbf{S}_{1\ldots k}^* + \frac{\hat{\omega}_{k+1}}{\hat{\gamma}_{1\ldots k+1}} \mathbf{S}_{1(k+1)}, \tag{16}$$

where $\hat{\gamma}_{1\ldots k} = \sum_{t=1}^{k} \hat{\omega}_t$, $\hat{\gamma}_{1\ldots k+1} = \hat{\gamma}_{1\ldots k} + \hat{\omega}_{k+1}$ and $\hat{\omega}_t$ is as defined in Equation (11).

Also,

$$\hat{\eta}_{1\ldots k+1} = \left(\frac{\hat{\gamma}_{1\ldots k}}{\hat{\gamma}_{1\ldots k+1}}\right)^2 \hat{\eta}_{1\ldots k} + \left(\frac{\hat{\omega}_{k+1}}{\hat{\gamma}_{1\ldots k+1}}\right)^2 \hat{\varsigma}_{k+1}, \tag{17}$$

where $\hat{\eta}_{1\ldots k} = \sum_{t=1}^{k} \hat{\gamma}_t^{*2} \hat{\varsigma}_t$, $\hat{\eta}_{1\ldots k+1} = \hat{\eta}_{1\ldots k} + \hat{\gamma}_{k+1}^{*2} \hat{\varsigma}_{k+1}$ and $\hat{\varsigma}_t$ is as defined in Equation (15). The proof is given in Appendix II.

This clearly implies that the three summary terms $\mathbf{S}_{1\ldots k}^*$, $\hat{\eta}_{1\ldots k}$ and $\hat{\gamma}_{1\ldots k}$ are sufficient to substitute all the k individual fault signatures $\mathbf{S}_{11}, \mathbf{S}_{12}, \ldots, \mathbf{S}_{1k}$ in terms of signature integration. Thus, we can discard all the fault signatures of the same fault from the fault library and only keep these three summary terms extracted from them. When a new fault signature, $\mathbf{S}_{1(k+1)}$ is obtained, its effect can be easily combined with $\mathbf{S}_{1\ldots k}^*$ to form an integrated signature which is equivalent to that constructed based on all the $k+1$ individual fault signatures. This makes it possible to continuously update the fault library as the manufacturing process proceeds.

With this iterative signature integration, the procedure of the self-improving variation source identification can be summarized as follows. Assume an integrated signature $\mathbf{S}_1^*$ exists in the fault library, and a symptom $\mathbf{S}_2$ is newly obtained. First, we need to perform the hypothesis testing given in Section 3.2 based on $\mathbf{S}_1^*$, $\hat{\eta}^*$, $\mathbf{S}_2$ and N_2. Specifically, get the value of T^* according to Equation (14). Then compare T^* with the critical value of the test, $\chi_{1-\alpha,v}^2$, where α is a selected significance level such as 0.05, 0.01, etc., and v is the degrees of freedom as defined for Equation (6). If $T^*$is larger than the critical value, then we claim that there is significant statistical evidence against $\mathbf{H}_0$ and thus a different fault occurs. Otherwise we claim that it is from the same fault.

In the second step, if $\mathbf{S}_2$ is decided to be from the same fault as $\mathbf{S}_1^*$, then we can view $\mathbf{S}_1^*$ as $\mathbf{S}_{1\ldots k}^*$ and $\mathbf{S}_2$ as $\mathbf{S}_{1(k+1)}$ and then use Equations (16) and (17) to update $\mathbf{S}_1^*$ and take the resulting $\mathbf{S}_{1\ldots k+1}^*$ as the new $\mathbf{S}_1^*$. If $\mathbf{S}_2$ is decided to be from a different fault, then other signatures in the library will be used to match with $\mathbf{S}_2$ following similar steps as above. If $\mathbf{S}_2$ does not match any existing signatures in the library, then the process should be inspected to decide if $\mathbf{S}_2$ indicates a novel fault. If yes, it is entered in the library as a new signature.

In the next section, a numerical study is conducted to demonstrate the effectiveness and robustness of the proposed method.

4. Numerical study

4.1. *Effectiveness of the proposed method*

To show the effectiveness of the proposed method, the performances of individual fault signatures, combined signature weighted only by sample sizes, and the proposed integrated signature are compared in 18 scenarios. The performance measures are type I error and type II error probabilities of the fault identification procedure. The type I error probability in variation source identification is defined as the probability that a symptom of the same fault will be wrongly rejected and be claimed as a new fault, which gives a "false alarm" of the occurrence of a novel fault. The type II error probability is the probability that a symptom of a different fault will be falsely accepted as the signature of the same fault, which means that we miss the opportunity of detecting a new fault.

In this study, we will use the linear model in Equation (1) to generate the sample covariance matrices (signatures and symptoms). The variable $\mathbf{y}$ is a 5×1 quality measurement vector and $\mathbf{f}$ is a 3×1 vector that represents the variation sources. ε is a 5×1 vector that models the system background noise with a covariance matrix $\sigma_\varepsilon^2\mathbf{I}$. The matrix $\mathbf{A}$

is assumed to be

$$\mathbf{A} = \begin{bmatrix} \sin\theta & \cos\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}',$$

where the angle θ is used to model the relative variation direction between the first two faults. Obviously, if θ is close to zero, the variation direction of these two faults will be very close and thus we will have a large type II error probability in variation source identification. If θ is close to 90°, then the impacts of these two process faults are quite different and thus we will have small type II error probability. To better demonstrate the advantage of the integrated signature, we choose $\theta = 4.5^\circ$ in this study.

Assume there are three fault signatures of the first fault available in the fault library, with population covariance matrices $\mathbf{\Sigma}_{11}$, $\mathbf{\Sigma}_{12}$ and $\mathbf{\Sigma}_{13}$. The diagonal components of $\mathbf{\Sigma}_{1t}$, $t = 1, 2, 3$, are σ^2_{1tk}, $k = 1, 2, 3$, representing, respectively, the magnitudes of the three faults corresponding to $\mathbf{\Sigma}_{1t}$. Since only the first fault happens, σ_{1tk}, $k = 2, 3$, is set to be zero. The noise variances of the three signatures, $\sigma^2_{\varepsilon_1}$, $\sigma^2_{\varepsilon_2}$ and $\sigma^2_{\varepsilon_3}$, are 0.006^2, 0.012^2 and 0.016^2 respectively.

A total of 18 scenarios are generated in the simulation. $(\sigma_{111}, \sigma_{121}, \sigma_{131})$ takes values of (0.05, 0.05, 0.05) and (0.05, 0.1, 0.2) to simulate the spread of the fault magnitude among the fault signatures. (N_{11}, N_{12}, N_{13}) takes values of nine combinations to simulate the different sample sizes among the fault signatures. 2000 Monte Carlo cases are generated for each scenario.

In each case, three fault signatures, $\mathbf{S}_{11}$, $\mathbf{S}_{12}$ and $\mathbf{S}_{13}$, of the first process fault are randomly generated based on the selected parameters. The combined signature weighted only by sample sizes, $\mathbf{S}_1^N$, can be calculated by Equation (9) with the weights $\hat{\gamma}_t = N_{1t} / \sum_{i=1}^{3} N_{1i}$, $t = 1,2,3$. The integrated fault signature $\mathbf{S}_1^*$ can be obtained by Equation (13) or Equation (16). Furthermore, two symptom covariance matrices $\mathbf{SI}_{\text{new}}$ and $\mathbf{SII}_{\text{new}}$(corresponding to the first and second process faults respectively) are generated and compared with $\mathbf{S}_{11}$, $\mathbf{S}_{12}$, $\mathbf{S}_{13}$, $\mathbf{S}_1^N$ and $\mathbf{S}_1^*$ separately. The generation of the two symptoms follows: (i) the variance of the fault is uniformly distributed between 0.05 and 0.5; (ii) the sample size is uniformly distributed between 25 and 100; and (iii) the noise variance is fixed at 0.01^2. Moreover, the hypothesis testing given in Section 2.1 is applied to match the individual signatures and fault symptoms, while that given in Section 3.2 is used to match $\mathbf{S}_1^N$ and $\mathbf{S}_1^*$ with symptoms.

To calculate the type I error and the type II error of the matching with a fault signature, say, $\mathbf{S}_{11}$, the test statistic T based on $\mathbf{S}_{11}$ and a fault symptom (either $\mathbf{SI}_{\text{new}}$ or $\mathbf{SII}_{\text{new}}$) will be calculated and compared with the critical value of the test, $\chi^2_{1-\alpha,v}$, where the degrees of freedom of the chi-square test $v = 4$ and the significance level α is set at 0.05. If T is larger than the critical value, then the null hypothesis should be rejected, and we decide that the chosen fault signature and the fault symptom represent two different faults. Otherwise we decide that the chosen fault signature and the fault symptom associate the same fault. Two variables, $CT1$ and $CT2$, are used to count the number of cases where the null hypothesis is rejected when $\mathbf{SI}_{\text{new}}$ and $\mathbf{SII}_{\text{new}}$ are the fault symptoms, respectively. Thus, the final type I and type II error probabilities can be calculated by $CT1/NC$ and $1 - CT2/NC$ respectively.

The simulation results are listed in Table 1, where α_1 and β_1 denote the smallest type I and type II error probabilities of the signature matching when individual signatures are used, α^N and β^Ndenote those when $\mathbf{S}_1^N$ is used and α^*

Table 1. Performance evaluation of the fault signatures.

$\sigma_1, \sigma_2, \sigma_3$	N_1, N_2, N_3	α_1	β_1	α^N	β^N	α^*	β^*
0.05, 0.05, 0.05	25,25,25	0.1220(1)	0.2890(1)	0.0925	0.3855	0.0800	0.1725
	25,50,75	0.0770(3)	0.2660(1)	0.0855	0.2155	0.0860	0.0845
	25,75,125	0.0610(3)	0.2610(1)	0.0820	0.1045	0.0900	0.0440
	50,50,50	0.0750(2)	0.0810(1)	0.0770	0.1370	0.0730	0.0360
	50,75,100	0.0565(2)	0.0785(1)	0.0785	0.0865	0.0675	0.0305
	50,100,150	0.0520(3)	0.0930(1)	0.0810	0.0500	0.0740	0.0285
	100,100,100	0.0650(1)	0.0265(1)	0.0840	0.0360	0.0695	0.0135
	100,125,150	0.0640(1)	0.0250(1)	0.0915	0.0285	0.0770	0.0210
	100,150,200	0.0600(2)	0.0240(1)	0.0775	0.0250	0.0810	0.0205
0.05, 0.1, 0.2	25,25,25	0.1035(1)	0.0850(3)	0.0730	0.0395	0.0645	0.0315
	25,50,75	0.0590(3)	0.0210(3)	0.0600	0.0175	0.0575	0.0170
	25,75,125	0.0675(2)	0.0140(3)	0.0725	0.0120	0.0705	0.0115
	50,50,50	0.0740(1)	0.0180(3)	0.0715	0.0130	0.0630	0.0110
	50,75,100	0.0660(3)	0.0125(3)	0.0665	0.0125	0.0650	0.0130
	50,100,150	0.0580(3)	0.0165(3)	0.0660	0.0155	0.0725	0.0145
	100,100,100	0.0615(3)	0.0120(3)	0.0575	0.0105	0.0595	0.0110
	100,125,150	0.0545(3)	0.0160(3)	0.0585	0.0140	0.0650	0.0130
	100,150,200	0.0615(2)	0.0090(3)	0.0680	0.0085	0.0745	0.0085

and β^* are those when $\mathbf{S}_1^*$ is used. One point needed to be mentioned is that there is no individual signature giving uniformly better results than the rest of the individual signatures. Thus, the values of α_1 and β_1 in the table are obtained from different individual signatures which are indicated in the parentheses.

From Table 1, we can see that all the tests produce reasonable type I error probabilities that are around 0.10 or below. Comparatively, individual signatures lead to relatively higher type I error than the two combined signatures when the sample size is small (e.g., $N_{11} = N_{12} = N_{13} = 25$), but in most cases, the three exhibit similar performances. The type II errors resulted from $\mathbf{S}_1^*$ are uniformly the smallest in all scenarios and in particular, impressively smaller than those from individual signatures and $\mathbf{S}_1^N$ when sample sizes and fault magnitudes are small. The type II errors resulted from $\mathbf{S}_1^N$ are slightly smaller than those from individual signatures when the fault magnitudes are big, but larger in two thirds of the scenarios when fault magnitudes are small. It means that combining the signatures based solely on sample sizes cannot necessarily lead to a better performance than using individual signatures; instead, the performance of the test may even be worse under some conditions. Since a smaller Type II error corresponds with higher detection power, this result validates that the proposed integrated signature is the desired one to improve the detection power of the signature matching.

Moreover, we can also find that the individual signature that gives the smallest errors is changing with scenarios. This implies that in practice, to achieve the best performance of individual signatures, we cannot just casually pick one from the available fault signatures or choose one by some simple criterion such as sample size and keep using it in signature matching. Rather, a selection procedure is needed to identify the optimal signature in each case and this will, obviously, bring many difficulties. Thus, actually the good performance, if there is one, of individual signatures in the simulation is hard to be realized. This adds to the strength of the proposed integrated signature.

4.2. *Robustness of the proposed method to assumptions*

There are two critical conditions leading to Theorem 1 as mentioned in Section 3.1: C1**)** the largest eigenvalue of each signature should be predominant among its eigenvalues; and C2**)** the small eigenvalues of each signature are equal to each other. Accordingly, two measures about the eigenvalues of the signatures are created to represent how these two conditions are satisfied, that is, the ratio of the largest to the second largest eigenvalue of each signature (for simplicity, later we call this the first/second eigenvalue ratio) and the Coefficient Of Variation (COV) of the small eigenvalues, which is defined as the ratio of the standard deviation of the small eigenvalues to their mean. In this study, the robustness of the proposed method is evaluated by the difference between the true significance level (estimated from simulations) and the nominal significance level ($\alpha = 0.05$) under different values of the first/second eigenvalue ratio and COV. Using the difference between the true and nominal significance level as a measure of the robustness of a hypothesis test has been adopted by many studies (e.g., Schott (1999)) because the difference represents the effectiveness of the asymptotic distribution as an approximation to the actual null distribution of the statistic in Equation (14).

The same linear model and $\mathbf{A}$ matrix as defined in Section 4.1 are used to generate the sample covariance matrices in the simulation. Still, three fault signatures of the first process fault are assumed, with fault magnitude $\sigma_{111} = \sigma_{121} = \sigma_{131} = 0.05$ and sample size $N_{11} = 200$, $N_{12} = 100$, $N_{13} = 300$. The noise covariance matrix of the signatures is now assumed to bear the form $\mathbf{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \times \text{diag}[\kappa, 1, 1, 1, 1]$ (the noise covariance matrix of the ith, $i = 1, 2, 3$, signature is $\sigma_{\varepsilon_i}^2 \times \text{diag}[\kappa_i, 1, 1, 1, 1]$). The first/second eigenvalue ratio can be controlled by both σ_ε^2 and κ, whereas the COV is only determined by κ.

Since the signatures have the same structure, their eigenvalues will exhibit similar characteristics. Thus, the first/second eigenvalue ratio and the COV of one signature, e.g., $\mathbf{S}_{13}$, are used to generate the cases. The following two scenarios are considered: Scenario 1 is designed to show the robustness of the proposed method to C1 under different levels of COV. Specifically, the COV of $\mathbf{S}_{13}$ is fixed by setting $\kappa_1 = \kappa_2 = \kappa_3 = 1$, 3 or 10 and under each setting, $(\sigma_{\varepsilon_1}^2, \sigma_{\varepsilon_2}^2, \sigma_{\varepsilon_3}^2) = \xi \times (0.006^2, 0.012^2, 0.016^2)$, where ξ is a coefficient whose values are chosen to make the first/second eigenvalue ratio of $\mathbf{S}_{13}$ be a series of values between one and 40. Correspondingly, Scenario 2 is designed to show the robustness of the proposed method to C2 under different levels of the first/second eigenvalue ratio. The way to realize this is a little bit complicated: for a certain level of the first/second eigenvalue ratio of $\mathbf{S}_{13}$, a series of values of the associated COV which are between 0.1 and 1.6 are obtained by setting $\kappa_1 = \kappa_2 = \kappa_3 = 1, 2.5, 4, \ldots, 16$ and for each of these values, the coefficient ξ is chosen to maintain the first/second eigenvalue ratio at the given level.

In each case of the two scenarios, three fault signatures are first generated and the integrated signature $\mathbf{S}_1^*$ values are calculated. Then, a symptom of the first process fault is generated with fault magnitude uniformly distributed between 0.05 and 0.5 and sample size being 200. The noise covariance matrix of the symptom is still $\sigma_\varepsilon^2 \mathbf{I}$ with $\sigma_\varepsilon^2 = 0.01$. The symptom will be matched with $\mathbf{S}_1^*$through the hypothesis testing in Section 3.2 and the corresponding type I error probability is obtained through 2000 simulations.

The resulted type I error probabilities of signature matching for scenarios 1 and 2 are shown in Figs. 1(a) and 1(b) respectively. In scenario 1, the three levels of COV are 0.14, 0.67 and 1.38, whereas in scenario 2, the levels of first/second eigenvalue ratio are three, five and ten. According to Fig. 1(a), under each level of COV, the type I error diminishes rapidly to be around 0.05 as the first/second

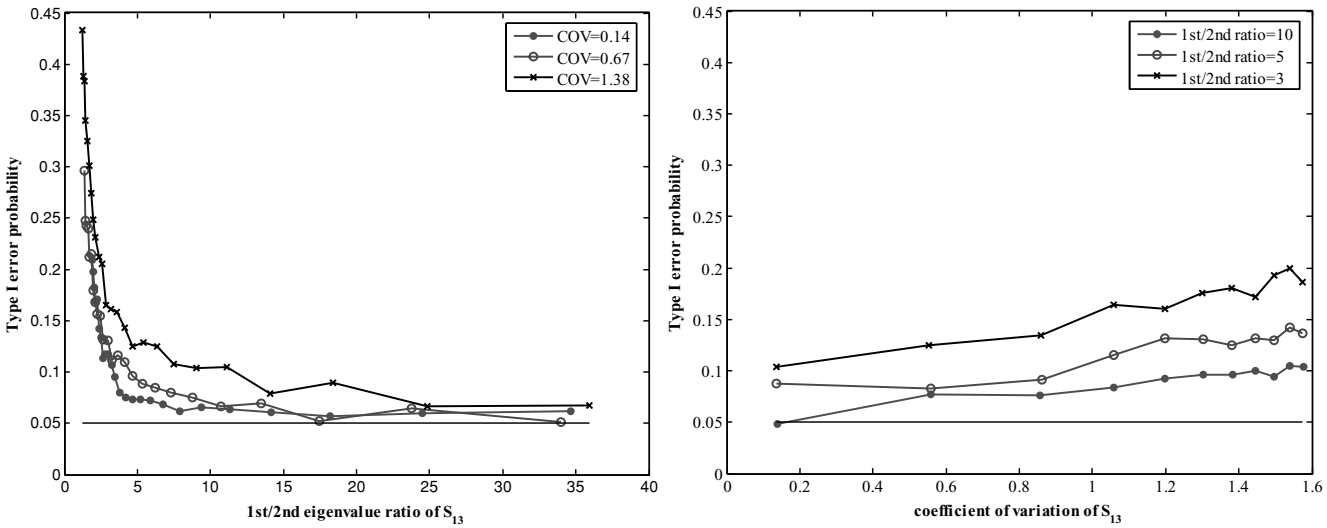

Fig. 1. Type I error probabilities in the simulation: (a) scenario 1; and (b) scenario 2.

eigenvalue ratio increases, and a smaller COV level corresponds with a faster degradation of the error, indicating that the smaller the COV, the stronger the robustness of the proposed method to C1. Moreover, even under the largest level of COV, the type I error falls within 0.1 when the first/second eigenvalue ratio is as large as ten. From Fig. 1(b), we can see that the type I error increases and deviates from 0.05 more and more with the increases of COV, and a larger first/second eigenvalue ratio leads to a smaller type I error, meaning that the larger the first/second eigenvalue ratio, the stronger the robustness of the proposed method to C2. It is also clear that even in the worst case where the first/second eigenvalue ratio is as small as three, the type I error is within 0.2.

Essentially, the first/second eigenvalue ratio represents the Signal-to-Noise Ratio (SNR) in a signature, whereas the COV of small eigenvalues indicates the differences in noise impact on different measurement dimensions. It is very difficult, if not impossible, to derive an analytical criteria on the robustness of the proposed method with respect to the SNR and the noise difference. However, based on this study, we can conclude that in common cases where the SNR is not very small and the noise difference is not very big, the asymptotic distribution of the statistic in Equation (14) will approximate its actual distribution reasonably and thus the proposed method will work.

5. Conclusions and future work

The data-driven variation source identification methodology is an important technique for production quality improvement in large-scale complex manufacturing process. This paper developed a linearly weighted integration procedure to combine multiple fault signatures of the same fault to improve the detection power of the fault identification procedure. It is found that the optimal weights for the fault signatures are related with their sample sizes and small eigenvalues. A numerical study is also presented to demonstrate the effectiveness of the integrated signature in terms of type I and type II errors of the identification procedure. Moreover, the robustness of the proposed method with respect to two critical assumptions is studied and it is shown that the proposed method can be commonly used in practice.

This research furnished the existing variation source identification methodology with a powerful tool to improve the accuracy of the fault signatures. There are some interesting open issues related with this work. First, in this study, we assume all the fault signatures are resulted from one single fault. However, the proposed method can be extended to multiple faults case under certain conditions. Actually we can prove that if the **A** matrix in model (1) is orthogonal, the proposed method can be easily extended to multiple fault cases without loss of strength. The fact that is critical to make the extension possible is that the significant eigenvectors of the fault signatures resulted from an orthogonal (and also normalized) **A** will be the same and thus any combined signature will have the similar simple properties as described in Section 3.1. However, if **A** is not orthogonal, the relationship between a combined signature and the individual fault signatures on which it is constructed may become very complex and hard to formulate. Consequently, the optimal weights will be difficult to define and obtain. New strategies are, obviously, needed to solve these problems to integrate the multiple fault signatures produced in multiple fault cases. Another issue is that, as we pointed out in the paper, the developed method

essentially requires a large SNR *or* small noise difference. Thus, it might be interesting to study comprehensively how small the SNR and how large the noise difference can go without deteriorating the performance of the tests to an unacceptable level. The numerical study presented in this paper is a rough analysis and more details need to be explored to form some rule-of-thumb guidelines for practice. This seems a challenging problem considering the large number of parameters that need to be considered. Finally, because errors in the testing for signature matching are unavoidable, it is also interesting and useful to study the appropriate way to handle the testing errors and corresponding signatures.

Acknowledgements

The authors gratefully acknowledge financial support from the NSF under grants CMMI-0322147, CMMI-0529327 and CMMI-0545600. The authors also thank the editor and the referee for their valuable comments and suggestions.

References

Apley, D.W. and Lee, H.Y. (2003) Identifying spatial variation patterns in multivariate manufacturing processes: a blind separation approach. *Technometrics*, **45**(3), 220–234.

Apley, D.W. and Shi, J. (1998) Diagnosis of multiple fixture faults in panel assembly. *Journal of Manufacturing Science and Engineering-Transactions of the ASME*, **120**(4), 793–801.

Apley, D.W. and Shi, J. (2001) A factor-analysis method for diagnosing variability in multivariate manufacturing processes. *Technometrics*, **43**(1), 84–95.

Camelio, J., Hu, S.J. and Ceglarek, D. (2003) Modeling variation propagation of multi-station assembly systems with compliant parts. *Journal of Mechanical Design*, **125**(4), 673–681.

Ceglarek, D. and Shi, J. (1996) Fixture failure diagnosis for autobody assembly using pattern recognition. *Journal of Engineering for Industry–Transactions of the ASME*, **118**(1), 55–66.

Chiang, L.H., Russell, E.L. and Braatz, R.D. (2001) *Fault Detection and Diagnosis in Industrial Systems*, Springer.

Ding, Y., Ceglarek, D. and Shi, J. (2000) *Modeling and diagnosis of multistage manufacturing processes: part I: state space model*. in Proceedings of the 2000 Japan/USA Symposium on Flexible Automation, Ann Arbor, MI, July 23–26, paper 2000JUSFA-13146.

Ding, Y., Ceglarek, D. and Shi, J. (2002) Fault diagnosis of multistage manufacturing processes by using state space approach. *Journal of Manufacturing Science and Engineering–Transactions of the ASME*, **124**(2), 313–322.

Ding, Y., Zhou, S. and Chen, Y. (2005) A comparison of process variation estimators for in-process dimensional measurements and control. *Journal of Dynamic Systems Measurement and Control–Transactions of the ASME*, **127**(1), 69–79.

Djurdjanovic, D. and Ni, J. (2001) Linear state space modeling of dimensional machining errors. *Transactions of NAMRI/SME*, **XXIX**, 541–548.

Jin, J. and Shi, J. (1999) State space modeling of sheet metal assembly for dimensional control. *Journal of Manufacturing Science and Engineering–Transactions of the ASME*, **121**(4), 756–762.

Jin, N. and Zhou, S. (2006a) Data-driven variation source identification of manufacturing processes based on eigenspace comparison. *Naval Research Logistics*, **53**(5), 383–396.

Jin, N. and Zhou, S. (2006b) Signature construction and matching for fault diagnosis in manufacturing processes through fault space analysis. *IIE Transactions*, **38**, 341–354.

Johnson, R.A. and Wichern, D.W. (2002) *Applied Multivariate Statistical Analysis*, Prentice Hall, Upper Saddle River, NJ.

Li, Z. and Zhou, S. (2006) Robust method of multiple variation sources identification in manufacturing processes for quality improvement. *Journal of Manufacturing Science and Engineering Transactions of the ASME*, **128**, 326–336.

Li, Z., Zhou, S. and Ding, Y. (2006) Pattern matching for root cause identification of manufacturing processes with consideration of general structured noise. *IIE Transactions*, **39**, 251–263.

Loose, J.-P., Zhou, S. and Ceglarek, D. (2007) Kinematic analysis of dimensional variation propagation for multistage machining processes with general fixture layout. *IEEE Transactions on Automation Science and Engineering*, **4**(2), 141–152.

Magnus, J.R. and Neudecker, H. (1979) Commutation matrix – some properties and applications. *Annals of Statistics*, **7**(2), 381–394.

Mantripragada, R. and Whitney, D.E. (1999) Modeling and controlling variation propagation in mechanical assemblies using state transition models. *IEEE Transactions on Robotics and Automation*, **15**(1), 124–140.

Rencher, A.C. (2002) *Methods of Multivariate Analysis*, Wiley, New York, NY.

Rong, Q., Ceglarek, D. and Shi, J. (2000) Dimensional fault diagnosis for compliant beam structure assemblies. *Journal of Manufacturing Science and Engineering Transactions of the ASME*, **122**(4), 773–780.

Schott, J.R. (1999) Partial common principal component subspaces. *Biometrika*, **86**(4), 899–908.

Schott, J.R. (2005) *Matrix Analysis for Statistics*, Wiley, Hoboken, NJ.

Tyler, D.E. (1981) Asymptotic inference for eigenvectors. *The Annals of Statistics*, **9**(4), 725–736.

Zhou, S., Chen, Y. and Shi, J. (2004) Statistical estimation and testing for variation root-cause identification of multistage manufacturing processes. *IEEE Transactions on Automation Science and Engineering*, **1**(1), 73–83.

Zhou, S., Ding, Y., Chen, Y. and Shi, J. (2003) Diagnosability study of multistage manufacturing processes based on linear mixed-effects models. *Technometrics*, **45**(4), 312–325.

Zhou, S., Huang, Q. and Shi, J. (2003) State space modeling of dimensional variation propagation in multistage machining process using differential motion vectors. *IEEE Transactions on Robotics and Automation*, **19**(2), 296–309.

Appendix

Appendix I

Proof of Theorem 1.

The following derivations are based on the properties of $\mathbf{S}_1^c$ stated in Section 3.1. Let us first derive the specific expression of Ψ_1 for any weight $\gamma_t, t = 1, \ldots, w$. By Equation (5):

$$\mathbf{H}_1 = \sum_{l=2}^{m} \left(\lambda_{11}^c - \lambda_{1l}^c\right)^{-1} \left(\mathbf{q}_{11}^c \mathbf{q}_{11}^{c'} \otimes \mathbf{q}_{1l}^c \mathbf{q}_{1l}^{c'}\right).$$

Because $\lambda_{1t1} \gg \lambda_{1tl}$, $t = 1, \ldots, w$, $l = 2, \ldots, m$, it is easy to get that $\lambda_{11}^{c} \gg \lambda_{1l}^{c}$ and thus:

$$\mathbf{H}_1 \approx \left(\lambda_{11}^{c}\right)^{-1} \left\{ \mathbf{q}_{11}^{c} \mathbf{q}_{11}^{c'} \otimes \sum_{l=2}^{m} \mathbf{q}_{1l}^{c} \mathbf{q}_{1l}^{c'} \right\} = \left(\lambda_{11}^{c}\right)^{-1} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\}. \tag{A1}$$

Let $\mathbf{\Lambda}_1 = \mathbf{S}_1^{c} - \mathbf{\Sigma}_1^{c}$, then:

$$\begin{aligned}
&\operatorname{var}\{\operatorname{vec}(\mathbf{\Lambda}_1)\} \\
&\quad = \sum_{t=1}^{w} \gamma_t^2 \operatorname{var}\{\operatorname{vec}(\mathbf{S}_{1t} - \mathbf{\Sigma}_{1t})\} \\
&\quad = \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t}} (\mathbf{I}_{m^2} + \mathbf{K}_{mm})(\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) \\
&\quad = \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t}} (\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) + \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t}} \mathbf{K}_{mm} (\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) \\
&\quad = \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t}} (\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) + \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t}} (\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) \mathbf{K}_{mm}.
\end{aligned} \tag{A2}$$

Let

$$\sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t}} (\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) = \mathbf{\Delta},$$

then $\operatorname{var}\{\operatorname{vec}(\mathbf{\Lambda}_1)\} = \mathbf{\Delta} + \mathbf{\Delta}\mathbf{K}_{mm}$ and

$$\begin{aligned}
\mathbf{H}_1 \mathbf{\Delta} &= \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t} \lambda_{11}^{c}} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} (\mathbf{\Sigma}_{1t} \otimes \mathbf{\Sigma}_{1t}) \\
&= \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t} \lambda_{11}^{c}} \left\{ \mathbf{P}_1^{c} \mathbf{\Sigma}_{1t} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \mathbf{\Sigma}_{1t} \right\}.
\end{aligned}$$

Since $\mathbf{\Sigma}_{1t} \approx \lambda_{1t1} \mathbf{P}_{1t} + \alpha_t (\mathbf{I}_m - \mathbf{P}_{1t})$, $\mathbf{P}_{1t}^2 = \mathbf{P}_{1t}$, $\mathbf{P}_{1t}(\mathbf{I}_m - \mathbf{P}_{1t}) = 0$ and $\mathbf{P}_1^{c} = \mathbf{P}_{1t}$, $t = 1, \ldots w$:

$$\mathbf{P}_1^{c} \mathbf{\Sigma}_{1t} = \lambda_{1t1} \mathbf{P}_1^{c} \text{ and } \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \mathbf{\Sigma}_{1t} = \alpha_t \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right).$$

Thus,

$$\begin{aligned}
\mathbf{H}_1 \mathbf{\Delta} &= \sum_{t=1}^{w} \frac{\gamma_t^2}{n_{1t} \lambda_{11}^{c}} \left\{ \lambda_{1t1} \mathbf{P}_1^{c} \otimes \alpha_t \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \\
&= \sum_{t=1}^{w} \frac{\gamma_t^2 \lambda_{1t1} \alpha_t}{n_{1t} \lambda_{11}^{c}} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \\
\mathbf{H}_1 \mathbf{\Delta} \mathbf{H}_1 &= \sum_{t=1}^{w} \frac{\gamma_t^2 \lambda_{1t1} \alpha_t}{n_{1t} \left(\lambda_{11}^{c}\right)^2} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \\
&= \sum_{t=1}^{w} \frac{\gamma_t^2 \lambda_{1t1} \alpha_t}{n_{1t} \left(\lambda_{11}^{c}\right)^2} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \\
&= \frac{\eta}{\left(\lambda_{11}^{c}\right)^2} \left\{ \mathbHowever, \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\}.
\end{aligned} \tag{A3}$$

$$\begin{aligned}
&\mathbf{H}_1 \Delta \mathbf{K}_{mm} \mathbf{H}_1 \\
&\quad = \sum_{t=1}^{w} \frac{\gamma_t^2 \lambda_{1t1} \alpha_t}{n_{1t} \left(\lambda_{11}^{c}\right)^2} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \mathbf{K}_{mm} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \\
&\quad = \sum_{t=1}^{w} \frac{\gamma_t^2 \lambda_{1t1} \alpha_t}{n_{1t} \left(\lambda_{11}^{c}\right)^2} \left\{ \mathbf{P}_1^{c} \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \right\} \left\{ \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \otimes \mathbf{P}_1^{c} \right\} \mathbf{K}_{mm} \\
&\quad = \sum_{t=1}^{w} \frac{\gamma_t^2 \lambda_{1t1} \alpha_t}{n_{1t} \left(\lambda_{11}^{c}\right)^2} \left\{ \mathbf{P}_1^{c} \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \otimes \left(\mathbf{I}_m - \mathbf{P}_1^{c}\right) \mathbf{P}_1^{c} \right\} \mathbf{K}_{mm} \\
&\quad = \mathbf{0}.
\end{aligned} \tag{A4}$$

By Equation (4):

$$\begin{aligned}
\mathbf{\Psi}_1 &= \mathbf{H}_1 \operatorname{var}\{\operatorname{vec}(\mathbf{\Lambda}_1)\} \mathbf{H}_1 = \mathbf{H}_1 (\mathbf{\Delta} + \mathbf{\Delta}\mathbf{K}_{mm}) \mathbf{H}_1 \\
&= \mathbf{H}_1 \mathbf{\Delta} \mathbf{H}_1 + \mathbf{H}_1 \mathbf{\Delta} \mathbf{K}_{mm} \mathbf{H}_1 = \mathbf{H}_1 \mathbf{\Delta} \mathbf{H}_1,
\end{aligned}$$

which gives Equation (10).

Let $C = \eta / (\lambda_{11}^{c})^2$. Thus, the optimal weights can be obtained through setting the first derivatives of C to be zero, that is

$$\frac{\partial C}{\partial \gamma_t} = 0, \quad t = 1, 2, \ldots, w. \tag{A5}$$

Let $\lambda_{11}^{c} = f = \sum_{r=1}^{w} \gamma_r \lambda_{1r1}$, then

$$C = \frac{\eta}{f^2} = \frac{1}{f^2} \left\{ \frac{\lambda_{1t1} \alpha_t}{n_{1t}} \gamma_t^2 + \sum_{r \neq t} \frac{\lambda_{1r1} \alpha_r}{n_{1r}} \gamma_r^2 \right\}. \tag{A6}$$

Consequently,

$$\begin{aligned}
\frac{\partial C}{\partial \gamma_t} &= \frac{\partial (f^{-2})}{\partial \gamma_t} \eta \\
&\quad + \frac{\partial \left\{ (\lambda_{1t1} \alpha_t / n_{1t}) \gamma_t^2 + \sum_{r \neq t} (\lambda_{1r1} \alpha_r / n_{1r}) \gamma_r^2 \right\}}{\partial \gamma_t} \frac{1}{f^2} \\
&= -2 f^{-3} \lambda_{1t1} \eta + 2 f^{-2} \frac{\lambda_{1t1} \alpha_t}{n_{1t}} \gamma_t.
\end{aligned}$$

And the solution for Equation (A5) is

$$\gamma_t^* = \frac{\eta^*}{f^*} \frac{n_{1t}}{\alpha_t},$$

where

$$\eta^* = \sum_{r=1}^{w} \frac{\gamma_r^{*2} \lambda_{1r1} \alpha_r}{n_{1r}} \text{ and } f^* = \sum_{r=1}^{w} \gamma_r^* \lambda_{1r1}.$$

It follows that,

$$\gamma_1^* : \gamma_2^* : \ldots : \gamma_w^* = \frac{n_{11}}{\alpha_1} : \frac{n_{12}}{\alpha_2} : \ldots : \frac{n_{1w}}{\alpha_w}. \tag{A7}$$

Based on Equation (A6), it is easy to validate that C is unchanged for any two sets of weights: $\{\gamma_t, t = 1, \ldots, w\}$ and $\{c\gamma_t, t = 1, \ldots, w\}$, where c is an arbitrary constant. Thus, we can normalize the weights in Equation (A7) and then Equation (11) results. ■

Appendix II

Proof of the iterative updating procedure in Section 3.3.

As $k = 1$, $\gamma_1 = \hat{\omega}_1$, $\gamma_{12} = \hat{\omega}_1 + \hat{\omega}_2$, and thus

$$\mathbf{S}_{12}^* = \frac{\hat{\omega}_1}{\hat{\omega}_1 + \hat{\omega}_2}\mathbf{S}_{11} + \frac{\hat{\omega}_2}{\hat{\omega}_1 + \hat{\omega}_2}\mathbf{S}_{12} = \hat{\gamma}_1^*\mathbf{S}_{11} + \hat{\gamma}_2^*\mathbf{S}_{12},$$
$$\hat{\eta}_{12} = \left(\frac{\hat{\omega}_1}{\hat{\omega}_1 + \hat{\omega}_2}\right)^2 \hat{\varsigma}_1 + \left(\frac{\hat{\omega}_2}{\hat{\omega}_1 + \hat{\omega}_2}\right)^2 \hat{\varsigma}_2 = \hat{\gamma}_1^{*2}\hat{\varsigma}_1 + \hat{\gamma}_2^{*2}\hat{\varsigma}_2.$$

Obviously, $\mathbf{S}_{12}^*$ and $\hat{\eta}_{12}$ are the same as $\mathbf{S}_1^*$ and $\hat{\eta}^*$ in Equation (13) and Equation (15) respectively when $w = 2$.

Assume $\mathbf{S}_{1\ldots k}^*$and $\hat{\eta}_{1\ldots k}$ are the same as $\mathbf{S}_1^*$ and $\hat{\eta}^*$ when $w = k$, that is

$$\mathbf{S}_{1\ldots k}^* = \sum_{t=1}^{k} \hat{\gamma}_t^*\mathbf{S}_{1t} \quad \text{where } \hat{\gamma}_t{}^* = \frac{\hat{\omega}_t}{\hat{\gamma}_{1\ldots k}},$$
$$\hat{\eta}_{1\ldots k} = \sum_{t=1}^{k} \hat{\gamma}_t^{*2}\hat{\varsigma}_t.$$

Then by Equation (16):

$$\begin{aligned}\mathbf{S}_{1\ldots k+1}^* &= \frac{\hat{\gamma}_{1\ldots k}}{\hat{\gamma}_{1\ldots k+1}} \sum_{t=1}^{k} \frac{\hat{\omega}_t}{\hat{\gamma}_{1\ldots k}}\mathbf{S}_{1t} + \frac{\hat{\omega}_{k+1}}{\hat{\gamma}_{1\ldots k+1}}\mathbf{S}_{1k+1} \\ &= \sum_{t=1}^{k} \frac{\hat{\omega}_t}{\hat{\gamma}_{1\ldots k+1}}\mathbf{S}_{1t} + \frac{\hat{\omega}_{k+1}}{\hat{\gamma}_{1\ldots k+1}}\mathbf{S}_{1k+1} \\ &= \sum_{t=1}^{k+1} \frac{\hat{\omega}_t}{\hat{\gamma}_{1\ldots k+1}}\mathbf{S}_{1t} \\ &= \sum_{t=1}^{k+1} \hat{\gamma}_t^*\mathbf{S}_{1t}.\end{aligned}$$

By Equation (17):

$$\begin{aligned}\hat{\eta}_{1\ldots k+1} &= \left(\frac{\gamma_{1\ldots k}}{\gamma_{1\ldots k+1}}\right)^2 \sum_{t=1}^{k} \left(\frac{\hat{\omega}_t}{\gamma_{1\ldots k}}\right)^2 \hat{\varsigma}_t + \left(\frac{\hat{\omega}_{k+1}}{\gamma_{1\ldots k+1}}\right)^2 \hat{\varsigma}_{k+1} \\ &= \sum_{t=1}^{k} \left(\frac{\hat{\omega}_t}{\gamma_{1\ldots k+1}}\right)^2 \hat{\varsigma}_t + \left(\frac{\hat{\omega}_{k+1}}{\gamma_{1\ldots k+1}}\right)^2 \hat{\varsigma}_{k+1} \\ &= \sum_{t=1}^{k+1} \hat{\gamma}_t^{*2}\hat{\varsigma}_t.\end{aligned}$$

These are the same as $\mathbf{S}_1^*$ and $\hat{\eta}^*$ when $w = k+1$. Thus, the procedure holds. ■

Biographies

Li Zeng is a Ph.D. student in the Department of Industrial and Systems Engineering at the University of Wisconsin-Madison. Her research interests are statistical modeling and variation source diagnosis in large complex systems.

Nong Jin received his Ph.D. in Industrial Engineering at the University of Wisconsin-Madison in 2006. Currently, he works at Capital One Financial Corporation. His areas of interest are credit risk analysis and valuations in credit card business.

Shiyu Zhou is an Associate Professor in the Department of Industrial and Systems Engineering at the University of Wisconsin-Madison. He was awarded has B.S. and M.S. in Mechanical Engineering by the University of Science and Technology of China in 1993 and 1996 respectively, and his Masters in Industrial Engineering and Ph.D. in Mechanical Engineering by the University of Michigan in 2000. His research interests are the in-process quality and productivity improvement methodologies by integrating statistics, system and control theory, and engineering knowledge. The objective is to achieve automatic process monitoring, diagnosis, compensation, and their implementation in various manufacturing processes. His research is sponsored by the National Science Foundation, Department of Energy, NIST-ATP and industries. He is a recipient of the CAREER Award from the National Science Foundation in 2006. He is a member of IIE, INFORMS, ASME and SME.