IIE Transactions (2009) **41**, 593–604

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170802389308

Detectability study for statistical monitoring of multivariate dynamic processes

NAN CHEN and SHIYU ZHOU*

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

Received March 2007 and accepted April 2008

Fault detection and diagnosis for dynamic processes is an intensively investigated area. However, the problem of determining whether or not system faults can be successfully detected based on the output measurements for a given dynamic process remains an open research topic. An intrinsic definition of fault detectability in multivariate dynamic processes is proposed in this paper. It defines the detectability in an intrinsic manner as a system property, without any reference to any specific fault detection algorithm. Furthermore, the relationship between system structure and the detectability for mean change faults and variability change faults are investigated. Analytical criteria for checking the system detectability are established. The results presented in this paper can provide guidelines on system design improvement for process monitoring and control. A case study is presented that illustrates the effectiveness of the proposed methods.

Keywords: Fault detection, detectability, multivariate dynamic system, statistical monitoring

1. Introduction

Statistical Process Control (SPC) plays a very important role in quality and productivity improvement of a manufacturing or service enterprise (Montgomery, 2005). Following SPC methodology, measurements are taken from the process or the product, they are then treated as random variables and their distributions are compared with the distributions under normal working conditions. If the measurements show that some characteristics are "out of control" (e.g., deviation from the target or variability is too high), an alarm is generated to indicate that changes have occurred in the process. For most of the available SPC techniques, the measurements are often either explicitly or implicitly assumed to be independent and identically distributed (i.i.d.).

Due to the rapid development of information and sensing technologies in recent years, a large amount of data is now readily available in many processes. Multi-dimensional measurements for discrete manufacturing processes with 100% inspection rate and very high sampling rates for continuous processes are no longer rare in practice. For example, in autobody assembly processes, 100% dimensional inspections have been achieved by the use of in-line optical coordinate measurement machines (Ceglarek and Shi, 1995). The extensive datasets provide significant opportunities for more sophisticated analysis for process monitoring, however, on the other hand, it also poses great challenges for SPC. Because of the very high sampling frequency and system inertia, the measurements often exhibit significant autocorrelation (Montgomery, 2005). Many researchers have demonstrated that the performance of SPC methods developed with i.i.d. assumptions will degrade when there exists dependence between successive samples (e.g., Harris and Ross (1991) and Montgomery and Mastrangelo (1991)). These works indicate that, for measurements with significant autocorrelation, dynamic models rather than static models should be used to fully utilize the measurements for process monitoring purposes.

The most commonly used dynamic models in SPC are time series models such as the AR(p) model (Alwan and Roberts, 1989; Montgomery and Mastrangelo, 1991) and the IMA (1, 1) model (MacGregor, 1988; Box and Kramer, 1992; Vander Wiel, 1996). People often fit the autocorrelated measurements using these models and because the model residuals are considered to be i.i.d., the conventional SPC techniques can then be applied to the residuals. Most of the available SPC methods that are based on time series model approaches only handle *univariate* cases. To monitor multivariate dynamic processes a state space model is often used to model the process dynamics. Negiz and Cinar (1997) used subspace identification methods to establish a state space model for the measurement data under

*Corresponding author

normal conditions, and then the prediction of the state vector is monitored through a T^2 control chart. Simoglou *et al.* (2002) used a similar approach and compared different techniques for the fitting of state space models. In addition to model-based methods, various model-free methods have also been developed for multivariate dynamic process monitoring, for example, methods based on asymptotical optimal sequential testing (Basseville and Nikiforov, 1993; Lai, 1995) and various Principal Component Analysis (PCA)-related methods, such as recursive PCA (Li *et al.*, 2000), multiscale PCA (Bakshi, 1998) and dynamic PCA (Ku *et al.*, 1995).

Although various works exist on the statistical monitoring of univariate and multivariate dynamic processes, a fundamental issue regarding process monitoring has not been thoroughly studied in past work. The issue is whether the process measurements contain sufficient information for the detection of process changes. This is referred to as detectability analysis and it is very important, especially in the design phase, because if a process is not detectable due to a problematic system structure, then no matter how much effort we put into the creation of a monitoring algorithm, we will not be able to detect process faults. Thus, if we blindly use statistical monitoring methods without any consideration of the process detectability, we may miss the changes that we want to detect. For example, Harris and Ross (1991) and Wardell *et al.* (1994) noticed that the one-step-ahead residual approach lacks the ability to detect changes in a time series with AR poles close to the unit circle. However, despite its importance, few papers have been published on the topic of exploring general detectability issues.

Since process changes are always caused by the occurrence of process faults, we treat "change detection" and "fault detection" as being interchangeable in this paper. Basseville (2001) provided a good review on various definitions of fault detectability for different types of faults. In that paper, the process faults are classified as being either: (i) additive faults on a linear system, which are manifested as a mean change of the process inputs and process outputs, or (ii) component (or system) faults, which are manifested as process structure changes such as the variance change of the multivariate process inputs. For each type of fault, the detectability definitions can be put into one of two categories: (i) an intrinsic definition that defines the detectability in an intrinsic manner as a system property, without any reference to any specific fault detection algorithms; or (ii) a performance-based definition that defines the detectability with explicit reference to a specific algorithm, taking into account its performance. Using the above classification, Zhou *et al.* (2003) and Apley and Ding (2005) investigated the properties of an intrinsic fault detection definition for both additive (i.e., mean shift of the input) and component faults (i.e., variance change of the input) in a *static* system. In their work, the process measurements are viewed as being i.i.d. and no autocorrelation and system dynamics are considered. On the other hand, Liu and Si (1997) and Gustafsson (2002) studied the properties of a performance-based detection definition only for additive faults in general *dynamic* systems. In their work, a state space model is adopted to model the system dynamics and the system observability is utilized to derive the system fault detectability.

In this paper, we will provide *an intrinsic fault detectability definition for both additive and component faults in multivariate dynamic processes*. A state space model is used to describe the system dynamics. A random vector is added to the system input to represent the faults. In this way, both the mean shift fault (i.e., additive fault) and variance change fault (i.e., a type of component fault) can be modeled. This fault representation can describe a wide range of practical situations and has been adopted in several previous works such as in Negiz and Cinar (1997), Chen *et al.* (1998) and Zhou *et al.* (2003). With the fault representation, an intrinsic detectability definition is proposed and the relationship between the defined detectability and the system structure is further investigated in this paper. The conditions under which the system detectability is guaranteed are derived. The results developed in this paper can be used to analyze the process structure to check if certain process faults are detectable and consequently provide quantitative guidelines on system design to improve the performance of the statistical monitoring scheme.

The paper is organized as follows: In Section 2, we will give a formal definition of the detectability. Then, the conditions of mean detectability and variance detectability will be explored, respectively. We also provide a corollary to easily check a system's detectability. In Section 3, a case study will be given to show the effectiveness of our method, and some possible applications of the result will also be pointed out. The paper is concluded in Section 4.

2. Intrinsic detectability of multivariate dynamic processes

2.1. *Problem formulation*

Linear Time Invariant (LTI) state space models are widely used because this approach can closely approximate many real systems and has a simple structure for analysis. Indeed, many dynamic models such as time series models can be easily transformed to a state space format (Aoki, 1990). In this paper, we adopt the state space model as a description of system dynamics. In the most generalized form, a stochastic LTI model with random noises and disturbances can be expressed in the following way:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}_u\mathbf{u}_k + \mathbf{B}_f\mathbf{f}_k + \mathbf{B}_w\mathbf{w}_k, \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}_u\mathbf{u}_k + \mathbf{D}_f\mathbf{s}_k + \mathbf{v}_k, \end{cases} \tag{1}$$

where $\mathbf{x}_k \in \Re^{n\times 1}$, $\mathbf{u}_k \in \Re^{l\times 1}$, $\mathbf{f}_k \in \Re^{p\times 1}$, $\mathbf{s}_k \in \Re^{s\times 1}$, $\mathbf{w}_k \in \Re^{r\times 1}$, $\mathbf{v}_k \in \Re^{m\times 1}$, $\mathbf{y}_k \in \Re^{m\times 1}$ are state variables, input variables, process faults, sensor faults, process noises, output noises and output measurements, respectively; the matrices $\mathbf{A}$, $\mathbf{B}_u$, $\mathbf{B}_f$, $\mathbf{B}_w$, $\mathbf{C}$, $\mathbf{D}_u$, $\mathbf{D}_f$ are the corresponding coefficient matrices with appropriate dimensions. For a LTI system, these matrices are constant matrices and are determined by the intrinsic structure of the process. In many cases, this model can be further simplified to fit real conditions. For example, in many processes, the input does not have a direct influence on the output. Therefore, the term of $\mathbf{D}_u\mathbf{u}_k$ can be often omitted in the output equation in model (1). Additionally, the sensor faults, represented by $\mathbf{s}_k$, can be often detected through redundancy by placing additional sensors to measure the same variables. Moreover, many sensor faults can be transformed into the form of an actuator fault (e.g., Massoumnia *et al.* (1989)). Other techniques such as the direct redundancy method can be also adopted to handle sensor faults. Properly designed algorithms can be further used to differentiate the sensor and actuator faults after studying the detectability of the system. With these considerations and for the sake of simplicity, we mainly focus on the detectability of the actuator faults $\mathbf{f}_k$. The simplified model adopted in this paper is

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}_u\mathbf{u}_k + \mathbf{B}_f\mathbf{f}_k + \mathbf{B}_w\mathbf{w}_k, \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k. \end{cases} \tag{2}$$

The following assumptions are made regarding this model:

1. The system and measurement noises $\mathbf{w}_k$ and $\mathbf{v}_k$ are a white noise sequence with zero mean and covariance matrix $\mathbf{\Sigma}_w$, $\mathbf{\Sigma}_v$ respectively.
2. Because different faults are often due to different physical causes, it is reasonable to assume the independence among different faults, i.e., $\mathbf{\Sigma}_f$ is assumed to be a diagonal matrix. We also assume that the fault vector $\mathbf{f}_k$ is independent with system noises $\mathbf{w}_k$ and $\mathbf{v}_k$.

These assumptions are not strict in real applications and similar assumptions are adopted in various previous studies such as Ding *et al.* (2002) and Zhou *et al.* (2003).

With the dynamic process model (2) and the above assumptions, the problem we want to address in this paper can be formulated as follows: Under normal working conditions, the mean and variance of process faults $\mathbf{f}_k$ are $\boldsymbol{\mu}_1$ and $\mathbf{\Sigma}_{f1}$, while under an abnormal working condition, the corresponding values become $\boldsymbol{\mu}_2$ and $\mathbf{\Sigma}_{f2}$, respectively. Can these changes be detected based on the monitoring of the measurements of the system outputs $\mathbf{y}_k$? In other words, what are the sufficient and necessary conditions for the detectability of process faults and how to test if the current system structure satisfies these conditions? The answers to these questions can guide us to design a system and ensure system detectability.

To address the above problem, model (2) needs to be transformed into a form that can more conveniently capture the autocorrelation structure of the output measurements as follows:

$$\begin{bmatrix} \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \vdots \\ \mathbf{y}_{k+N} \end{bmatrix} = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^N \end{bmatrix} \mathbf{x}_k + \mathbf{\Gamma}_u \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+N-1} \end{bmatrix} + \mathbf{\Gamma}_f \begin{bmatrix} \mathbf{f}_k \\ \mathbf{f}_{k+1} \\ \vdots \\ \mathbf{f}_{k+N-1} \end{bmatrix} + \mathbf{\Gamma}_w \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+N-1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{k+1} \\ \mathbf{v}_{k+2} \\ \vdots \\ \mathbf{v}_{k+N} \end{bmatrix}, \tag{3}$$

where $N \geq n$. To simplify the notation, we define $\mathbf{\Gamma}_\alpha$ (α can be u, f or w) and $\mathbf{\Xi}$ as follows:

$$\mathbf{\Gamma}_\alpha = \begin{bmatrix} \mathbf{CB}_\alpha & & & \\ \mathbf{CAB}_\alpha & \mathbf{CB}_\alpha & & \\ \vdots & \vdots & \ddots & \\ \mathbf{CA}^{N-1}\mathbf{B}_\alpha & \mathbf{CA}^{N-2}\mathbf{B}_\alpha & \cdots & \mathbf{CB}_\alpha \end{bmatrix}, \quad \mathbf{\Xi} = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^N \end{bmatrix}. \tag{4}$$

We further define $\mathbf{Y}_{k+1}^N = [\mathbf{y}_{k+1}^{\mathrm{T}}\ \mathbf{y}_{k+2}^{\mathrm{T}} \cdots \mathbf{y}_{k+N}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{U}_k^N = [\mathbf{u}_k^{\mathrm{T}}\ \mathbf{u}_{k+1}^{\mathrm{T}} \cdots \mathbf{u}_{k+N-1}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{F}_k^N = [\mathbf{f}_k^{\mathrm{T}}\ \mathbf{f}_{k+1}^{\mathrm{T}} \cdots \mathbf{f}_{k+N-1}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{W}_k^N = [\mathbf{w}_k^{\mathrm{T}}\ \mathbf{w}_{k+1}^{\mathrm{T}} \cdots \mathbf{w}_{k+N-1}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{V}_{k+1}^N = [\mathbf{v}_{k+1}^{\mathrm{T}}\ \mathbf{v}_{k+2}^{\mathrm{T}} \cdots \mathbf{v}_{k+N}^{\mathrm{T}}]^{\mathrm{T}}$, and finally we have:

$$\mathbf{H}_k = \mathbf{\Xi}\mathbf{x}_k + \mathbf{\Gamma}_f\mathbf{F}_k^N + \mathbf{\Gamma}_w\mathbf{W}_k^N + \mathbf{V}_{k+1}^N, \tag{5}$$

where $\mathbf{H}_k$ is defined as $\mathbf{Y}_{k+1}^N - \mathbf{\Gamma}_w\mathbf{U}_k^N$ and it is measurable because both $\mathbf{Y}_{k+1}^N$ and $\mathbf{U}_k^N$ are measurable and the coefficient matrices are assumed to be known, while the terms on the right-hand side involve variables that could not be directly observed. Based on the model expressed in Equation (5), an intrinsic detectability definition can be established as follows. A similar definition for a static process instead of a dynamic process has been adopted in previous works (Rao and Kleffe, 1988; Zhou *et al.*, 2003).

Definition 1. In model (5), the system is said to be mean detectable, if $\forall E({}^1\mathbf{F}_k^N)$, $E({}^2\mathbf{F}_k^N)$ and k:

$$E\left({}^1\mathbf{F}_k^N\right) \neq E\left({}^2\mathbf{F}_k^N\right) \quad \text{implies} \quad E\left({}^1\mathbf{H}_k\right) \neq E\left({}^2\mathbf{H}_k\right). \tag{6}$$

The system is said to be variance detectable, if $\forall$ var$({}^1\mathbf{F}_k^N)$, var$({}^2\mathbf{F}_k^N)$ and k:

$$\mathrm{var}\left({}^1\mathbf{F}_k^N\right) \neq \mathrm{var}\left({}^2\mathbf{F}_k^N\right) \quad \text{implies} \quad \mathrm{var}\left({}^1\mathbf{H}_k\right) \neq \mathrm{var}\left({}^2\mathbf{H}_k\right), \tag{7}$$

where $E(\cdot)$ and var$(\cdot)$ represent the expectation and variance of the random variables. The left superscript 1 and 2 represent two different working conditions of the process and in each working condition, the mean and variance of the faults are kept as constants. The basic idea of this definition is that if the system is detectable, the changes in the mean and covariance of the faults in two different working conditions should result in the mean and covariance changes of observations. This property is an intrinsic property of the system because it is not related to any specific monitoring algorithm. The autocorrelation of the system output is automatically considered in the definition because $\mathbf{Y}_{k+1}^N$ is a stack up of N output values and the covariance of $\mathbf{Y}_{k+1}^N$ actually represents the autocorrelation among the output values. In the definition, it seems that the process detectability is related to the number of the stacked measurements N. However, the analysis in the following sections shows that the process detectability is actually not related to N given $N \geq n$. Note also that in the definition, we are interested in detecting the change of the fault vector, which may be caused by a single fault element or multiple concurrent faults. Under this setting, the detectability analysis would automatically consider multiple concurrent faults. Another point that needs to be mentioned regarding the definition is that we only consider the relationship of mean and covariance between system observations and faults. If the random variables involved in the system follow a normal distribution, then the mean and covariance can completely determine the distribution. However, for other distributions, there could be conditions that even if the system is not detectable based on the mean and variance, the change in the system could still be detected through the analysis of higher order statistics of the measurements. However, to limit the scope of this paper, we will focus on the mean and variance detectability as above defined. We only consider step changes in the mean and variance, although gradual drifting changes can be analyzed through similar methodologies.

2.2. *System detectability analysis*

Based on model (5), we can get the mean and covariance of the observations:

$$E(\mathbf{H}_k) = \boldsymbol{\Xi}\, E(\mathbf{x}_k) + \boldsymbol{\Gamma}_f E\big(\mathbf{F}_k^N\big) + \boldsymbol{\Gamma}_w E\big(\mathbf{W}_k^N\big) + E\big(\mathbf{V}_{k+1}^N\big), \tag{8a}$$

$$\mathrm{var}(\mathbf{H}_k) = \boldsymbol{\Xi} \times \mathrm{var}(\mathbf{x}_k) \times \boldsymbol{\Xi}^{\mathrm{T}} + \boldsymbol{\Gamma}_f \times \mathrm{var}\big(\mathbf{F}_k^N\big) \times \boldsymbol{\Gamma}_f^{\mathrm{T}} + \boldsymbol{\Gamma}_w \times \mathrm{var}\big(\mathbf{W}_k^N\big) \times \boldsymbol{\Gamma}_w^{\mathrm{T}} + \mathrm{var}\big(\mathbf{V}_{k+1}^N\big). \tag{8b}$$

Please note that because $\mathbf{F}_k^N$, $\mathbf{W}_k^N$ and $\mathbf{V}_{k+1}^N$ collect the faults and noises that occur at and after time k, they are independent of the system state $\mathbf{x}_k$ at time k. Since we assume that $\mathbf{w}_k$, and $\mathbf{v}_k$ have zero means, therefore $E(\mathbf{W}_k^N) = 0$ and $E(\mathbf{V}_{k+1}^N) = 0$. In many applications, the variance of noises is much smaller than that of faults, thus $\boldsymbol{\Gamma}_w \times \mathrm{var}(\mathbf{W}_k^N) \times \boldsymbol{\Gamma}_w^{\mathrm{T}}$ and $\mathrm{var}(\mathbf{V}_{k+1}^N)$ can be neglected with the presence of $\boldsymbol{\Gamma}_f \times \mathrm{var}(\mathbf{F}_k^N) \times \boldsymbol{\Gamma}_f^{\mathrm{T}}$. Another justification is that, under normal conditions without process faults, we can estimate the covariance of both $\mathbf{w}_k$ and $\mathbf{v}_k$ accurately with large samples. Therefore, we explore the detectability condition of faults without taking the last two terms of Equation (8b) into consideration. Meanwhile, under normal working conditions, we can estimate the mean and covariance matrix of the state variable $\mathbf{x}_k$ as well. With these considerations, it is easy to see that the detectability analysis boils down to the following problem: does $E({}^1\mathbf{F}_k^N) \neq E({}^2\mathbf{F}_k^N)$ imply $\boldsymbol{\Gamma}_f \times \mathrm{E}({}^1\mathbf{F}_k^N) \neq \boldsymbol{\Gamma}_f \times \mathrm{E}({}^2\mathbf{F}_k^N)$? And does $\mathrm{var}({}^1\mathbf{F}_k^N) \neq \mathrm{var}({}^2\mathbf{F}_k^N)$ imply $\boldsymbol{\Gamma}_f \times \mathrm{var}({}^1\mathbf{F}_k^N) \times \boldsymbol{\Gamma}_f^{\mathrm{T}} \neq \boldsymbol{\Gamma}_f \times \mathrm{var}({}^2\mathbf{F}_k^N) \times \boldsymbol{\Gamma}_f^{\mathrm{T}}$? The following two lemmas provide the sufficient and necessary conditions for mean detectability and variance detectability of a given system, respectively. It is worth mentioning that for a fault with both mean shift and variance change, the following analysis is still applicable as long as our concern is the fault detectability.

Lemma 1. *The system is mean detectable, if and only if* $\mathrm{rank}(\mathbf{OB}_{\mathrm{f}}) = \mathrm{p}$, *where* p *is the dimension of fault vectors, and* $\mathbf{O}$ *is the observability matrix of the system, defined as*

$$\mathbf{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \quad (n \text{ is the system order}). \tag{9}$$

Proof. First, we prove if $rank(\mathbf{OB}_f) = p$, then the system is mean detectable. Since $E(\mathbf{f}_k)$ is assumed to be constant for a given working condition, we have:

$$\boldsymbol{\Gamma}_f\big[E\big({}^1\mathbf{F}_k^N\big) - E\big({}^2\mathbf{F}_k^N\big)\big] = \begin{bmatrix} \mathbf{I} & & & \\ \mathbf{I} & \mathbf{I} & & \\ \vdots & \vdots & \ddots & \\ \mathbf{I} & \mathbf{I} & \cdots & \mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{N-1} \end{bmatrix} \times \mathbf{B}_f \times (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \overset{\mathrm{def}}{=} \mathbf{M} \times (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2), \tag{10}$$

by rearranging the $\boldsymbol{\Gamma}_f$ matrix, where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the mean of the process faults under two different conditions, respectively. Furthermore, through the Cayley Hamilton theorem (Lancaster, 1969), we have:

$$\mathbf{O}^N = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{N-1} \end{bmatrix} = \mathbf{QO}, \text{ where } \mathbf{Q} \in \Re^{Nm \times nm} \text{ and}$$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{I}_{nm} \\ \mathbf{Q}^* \end{bmatrix} \text{ when } N \geq n \text{ } (\mathbf{Q}^* \text{ is the submatrix with dimension } (N-n)m \times nm), \text{ which is full column rank.}$$

Therefore, $rank(\mathbf{O}^N \times \mathbf{B}_f) = rank(\mathbf{QOB}_f) = rank(\mathbf{OB}_f) = p$ when $N \geq n$, since left multiplication by a full column rank matrix does not change the rank. With the same argument, we can find that $\mathbf{M}$ can be obtained by left multiplying a full column rank matrix on $\mathbf{O}^N \times \mathbf{B}_f$, therefore, $rank(\mathbf{M}) = p$, which indicates that $\mathbf{M}$ is full column rank. Thus the null space of $\mathbf{M}$ is empty and for any $\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, we will have $\mathbf{M} \times (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \neq \mathbf{0}$, and hence $\boldsymbol{\Gamma}_f \times E({}^1\mathbf{F}_k^N) \neq \boldsymbol{\Gamma}_f \times E({}^2\mathbf{F}_k^N)$.

Next, we prove if the system is mean detectable, then *rank* $(\mathbf{OB}_f) = p$. Using similar arguments, we can show that if the system is mean detectable, then the null space of the $\mathbf{M}$ matrix defined in Equation (10) should be empty and thus we should have $rank(\mathbf{M}) = p$. Furthermore, since $p = rank(\mathbf{M}) \leq rank(\mathbf{OB}_f) \leq p$, we can get $rank(\mathbf{OB}_f) = p$ accordingly. ■

There is an intuitive understanding of this lemma from the space mapping point of view as shown in Fig. 1. $\mathbf{B}$ is the mapping from fault space $\Re^{p\times 1}$ to the state space $\Re^{n\times 1}$, and the $\mathbf{C}$ matrix is the mapping from state space $\Re^{n\times 1}$ to output space $\Re^{m\times 1}$. In statistical monitoring, we try to detect the process change based on the analysis of process outputs, thus the mapping is actually in the reverse direction: $\mathbf{C}^{-1}$ is the inverse output mapping from $\Re^{m\times 1}$ to $\Re^{n\times 1}$, and $\mathbf{B}^{-1}$ is the inverse input mapping from $\Re^{n\times 1}$ to $\Re^{p\times 1}$ with Ω defined as $\mathrm{Im}(\mathbf{B}^{-1})$. We can further define the unobservable subspace $\boldsymbol{\Theta} = \{\mathbf{x}|\mathbf{Ox} = \mathbf{0}\}$ in the state space, while the inverse mapping from output space to state space cannot be located in the unobservable subspace.

In Fig. 1, the darker area in the state space of $\Re^{n\times 1}$ represents the image of mapping $\mathbf{B}$, denoted by $\mathrm{Im}(\mathbf{B})$; the white area represents the unobservable subspace. While in statistical monitoring, if we want to detect the changes in the fault vectors, we need to have the mapping $\mathbf{B}$: $\Re^{p\times 1} \rightarrow \mathrm{Im}(\mathbf{B})$ be bijective, which means $\mathbf{B}$ should be full column rank, to make sure of the one-to-one correspondence between the fault and its image. Furthermore, if there are any images falling into the unobservable space, i.e., $\exists \mathbf{f} \neq 0,\ s.t.\ \mathbf{OB}_f\mathbf{f} = 0$, then for some changes in $\mathbf{f}$, the observations will not reflect that change. Therefore, to be fully detectable, the intersection between the image of mapping $\mathbf{B}$ and the unobservable subspace should be empty, i.e., $\boldsymbol{\Theta} \cap \mathrm{Im}(\mathbf{B}) = \varnothing$, which requires $\forall \mathbf{f} \neq 0,\ s.t.\ \mathbf{OB}_f\mathbf{f} \neq 0$. This condition is exactly equivalent to $rank(\mathbf{OB}_f) = p$.

Similarly, we can also get the sufficient and necessary condition for variance detectability, as stated by Lemma 2.

Lemma 2. *The system is variance detectable, if and only if* $\mathrm{rank}(\Pi(\mathbf{OB}_\mathrm{f})) = \mathrm{p}$, *where* $\Pi(\cdot)$ *is defined as*

$$\Pi(\boldsymbol{\Gamma}) = \begin{bmatrix} \boldsymbol{\gamma}_1 \otimes \boldsymbol{\gamma}_1 \\ \vdots \\ \boldsymbol{\gamma}_1 \otimes \boldsymbol{\gamma}_{mN} \\ \boldsymbol{\gamma}_2 \otimes \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\gamma}_2 \otimes \boldsymbol{\gamma}_{mN} \\ \vdots \\ \boldsymbol{\gamma}_{mN} \otimes \boldsymbol{\gamma}_{mN} \end{bmatrix},$$

where $\boldsymbol{\gamma}_i$ *is the* i*th row of the matrix* $\boldsymbol{\Gamma}$, mN *is the total number of rows in* $\boldsymbol{\Gamma}$ *and* $\otimes$ *represents the Hadamard product defined as* $(\mathbf{A} \otimes \mathbf{B})_{\mathrm{ij}} = \mathbf{A}_{\mathrm{ij}} \times \mathbf{B}_{\mathrm{ij}}$, *where the subscript* ij *represents the* i*th row*, j*th column element of the corresponding matrix.*

Proof. First, we prove if $rank(\Pi(\mathbf{OB}_f)) = p$, then the system is variance detectable. Denote $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ as the covariance matrix of the process faults for two different conditions. By rearranging the matrix $\boldsymbol{\Gamma}_f$ and noting that the covariance matrix of the process faults is a diagonal matrix, we can get

$$\mathrm{vec}\left(\boldsymbol{\Gamma}_f \times \mathrm{var}\left({}^1\mathbf{F}_k^N\right) \times \boldsymbol{\Gamma}_f^{\mathrm{T}}\right) = \Pi(\boldsymbol{\Gamma}_f) \times diag \times \left(\begin{bmatrix} \boldsymbol{\Sigma}_1 & & & \\ & \boldsymbol{\Sigma}_1 & & \\ & & \ddots & \\ & & & \boldsymbol{\Sigma}_1 \end{bmatrix}\right) = \Pi(\boldsymbol{\Gamma}_f) \times \begin{bmatrix} \mathbf{I}_p \\ \mathbf{I}_p \\ \vdots \\ \mathbf{I}_p \end{bmatrix} \times diag(\boldsymbol{\Sigma}_1), \tag{11}$$

where $vec(\cdot)$ is an operator that stacks up the upper triangle of a symmetric matrix to form a column vector, and $diag(\mathbf{A})$ is the column vector composed of the diagonal elements in $\mathbf{A}$. A similar identity holds for the second condition. We can further rewrite $\boldsymbol{\Gamma}_f$ as $\boldsymbol{\Gamma}_f = [\mathbf{T}_1\ \mathbf{T}_2 \cdots \mathbf{T}_N]$ where $\mathbf{T}_i$ is the matrix consisting from $[(i-1)p+1]$th to $[ip]$th columns of $\boldsymbol{\Gamma}_f$. Due to the property of Π transformation, we have $\Pi(\boldsymbol{\Gamma}_f) = [\,\Pi(\mathbf{T}_1)\ \Pi(\mathbf{T}_2) \cdots \Pi(\mathbf{T}_N)]$.

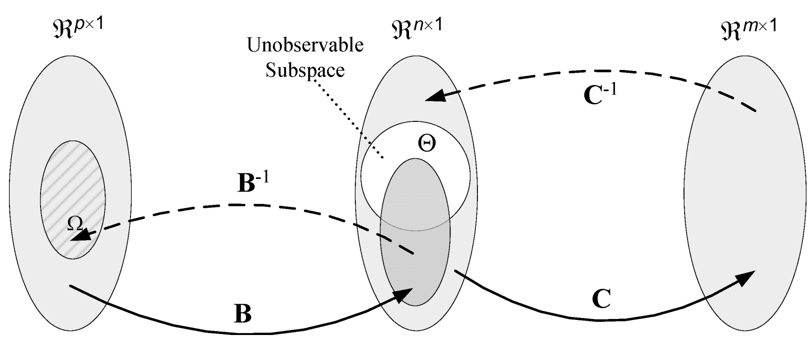

Fig. 1. Space mapping in fault diagnosis.

Furthermore, due to the property of $\boldsymbol{\Gamma}_f$, each non-zero row in $\mathbf{T}_i$ $(i \geq 2)$ is identical to one particular row in $\mathbf{T}_1$. Therefore, each non-zero row in $\Pi(\mathbf{T}_i)$ is identical to one row in $\Pi(\mathbf{T}_1)$. We have:

$$\Pi(\mathbf{T}_i) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{P}_i^* & \mathbf{0} \end{bmatrix} \times \Pi(\mathbf{T}_1) \stackrel{def}{=} \mathbf{P}_i \times \Pi(\mathbf{T}_1),$$

where $\mathbf{P}_i$ is a permutation matrix and $\mathbf{P}_i^*$ is in the lower triangle of $\mathbf{P}_i$ for $i \geq 2$. Therefore

$$\Pi(\boldsymbol{\Gamma}_f) \times \begin{bmatrix} \mathbf{I}_p \\ \mathbf{I}_p \\ \vdots \\ \mathbf{I}_p \end{bmatrix} = \sum_{i=1}^{N} \Pi(\mathbf{T}_i) = \left(\mathbf{I} + \sum_{i=1}^{N} \mathbf{P}_i \right) \times \Pi(\mathbf{T}_1)$$
$$= \begin{bmatrix} \mathbf{I} & 0 \\ \mathbf{P}^* & \mathbf{I} \end{bmatrix} \times \Pi(\mathbf{T}_1) \stackrel{def}{=} \mathbf{P} \times \Pi(\mathbf{T}_1). \quad (12)$$

Obviously, $\mathbf{P}$ is an invertible matrix. Additionally, when $N \geq n$, by the definition of Π transformation, we have:

$$\Pi(\mathbf{O}^N \mathbf{B}_f) = \begin{bmatrix} \mathbf{I} \\ \mathbf{Q}^* \end{bmatrix} \times \Pi(\mathbf{O}\mathbf{B}_f) \stackrel{def}{=} \mathbf{Q} \times \Pi(\mathbf{O}\mathbf{B}_f),$$

where

$$\mathbf{O}^N = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{N-1} \end{bmatrix}.$$

Clearly, $\mathbf{Q}$ is a full column rank matrix, and noticing $\mathbf{T}_1 = \mathbf{O}^N \mathbf{B}_f$, we have $rank(\mathbf{P} \times \Pi(\mathbf{T}_1)) = p$ if $rank(\Pi(\mathbf{O}\mathbf{B}_f)) = p$. Therefore, if $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$ and $rank(\Pi(\mathbf{O}\mathbf{B}_f)) = p$, then $\boldsymbol{\Gamma}_f \times \mathrm{var}(^1\mathbf{F}_k^N) \times \boldsymbol{\Gamma}_f^{\mathrm{T}} \neq \boldsymbol{\Gamma}_f \times \mathrm{var}(^2\mathbf{F}_k^N) \times \boldsymbol{\Gamma}_f^{\mathrm{T}}$ because the null space of $\mathbf{P} \times \Pi(\mathbf{T}_1)$ is empty.

Next, we prove that if the system is detectable, then $rank(\Pi(\mathbf{O}\mathbf{B}_f)) = p$. If the system is variance detectable, $\forall \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$, and thus $\mathrm{var}(^1\mathbf{F}_k^N) \neq \mathrm{var}(^2\mathbf{F}_k^N)$, we can get:

$$vec\left[\boldsymbol{\Gamma}_f \times \left(\mathrm{var}\left(^1\mathbf{F}_k^N\right) - \mathrm{var}\left(^2\mathbf{F}_k^N\right)\right) \times \boldsymbol{\Gamma}_f^T\right] = \mathbf{P} \times \Pi(\mathbf{T}_1)$$
$$\times (diag(\boldsymbol{\Sigma}_1) - diag(\boldsymbol{\Sigma}_2)) \neq 0.$$

This implies that $\mathbf{P} \times \Pi(\mathbf{T}_1)$, as defined above, should be full column rank. We have $p \leq rank(\Pi(\mathbf{T}_1)) \leq rank(\Pi(\mathbf{O}\mathbf{B}_f)) \leq p$. This proves that $rank(\Pi(\mathbf{O}\mathbf{B}_f)) = p$ is the necessary condition of variance detectability. ■

The above two lemmas provide a powerful tool to check whether the system is detectable in terms of mean shift and variance change of the process faults. Although step change assumptions are made during the derivation, the two conditions still work for linear drifting changes through a straightforward extension, and can be used to check the detectability of that kind of change. From the lemma, it is also clear that the detectability has a close relationship with the coefficient matrix $\mathbf{B}_f$ and the observability matrix $\mathbf{O}$ of the system. The above results can be further simplified through a state transformation to change the coordinate of the system.

From control theory, we can always find an invertible matrix $\mathbf{G}$ to transform the state to $\tilde{\mathbf{x}} = \mathbf{G}^{-1}\mathbf{x}$ (Rugh, 1996):

$$\begin{cases} \tilde{\mathbf{x}}_{k+1} = \begin{bmatrix} \tilde{\mathbf{A}}_{11} & \mathbf{0} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{bmatrix} \tilde{\mathbf{x}}_k + \begin{bmatrix} \tilde{\mathbf{B}}_{u,q} \\ \tilde{\mathbf{B}}_{u,n-q} \end{bmatrix} \mathbf{u}_k + \begin{bmatrix} \tilde{\mathbf{B}}_{f,q} \\ \tilde{\mathbf{B}}_{f,n-q} \end{bmatrix} \mathbf{f}_k \\ \qquad + \begin{bmatrix} \tilde{\mathbf{B}}_{w,q} \\ \tilde{\mathbf{B}}_{w,n-q} \end{bmatrix} \mathbf{w}_k, \\ \mathbf{y}_k = \begin{bmatrix} \tilde{\mathbf{C}}_q & \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}}_k + \mathbf{v}_k, \end{cases} \quad (13)$$

where $\tilde{\mathbf{B}}_{u,q} \in \Re^{q \times l}$, $\tilde{\mathbf{B}}_{f,q} \in \Re^{q \times p}$, $\tilde{\mathbf{B}}_{w,q} \in \Re^{q \times r}$ and $\tilde{\mathbf{C}}_q \in \Re^{n \times q}$. Thus, the observability matrix can be written as

$$\tilde{\mathbf{O}} = \begin{bmatrix} \tilde{\mathbf{C}}_q & \mathbf{0} \\ \tilde{\mathbf{C}}_q \tilde{\mathbf{A}}_{11} & \mathbf{0} \\ \vdots & \vdots \\ \tilde{\mathbf{C}}_q \tilde{\mathbf{A}}_{11}^{n-1} & \mathbf{0} \end{bmatrix},$$

and hence

$$\tilde{\mathbf{O}}\tilde{\mathbf{B}}_f = \begin{bmatrix} \tilde{\mathbf{C}}_q \\ \tilde{\mathbf{C}}_q \tilde{\mathbf{A}}_{11} \\ \vdots \\ \tilde{\mathbf{C}}_q \tilde{\mathbf{A}}_{11}^{n-1} \end{bmatrix} \times \tilde{\mathbf{B}}_{f,q} \stackrel{def}{=} \tilde{\mathbf{O}}_q \times \tilde{\mathbf{B}}_{f,q} = \mathbf{O}\mathbf{B}_f. \quad (14)$$

State transformation does not change the observability, therefore $rank(\tilde{\mathbf{O}}_q) = rank(\mathbf{O}) = q$. However, the $\tilde{\mathbf{O}}_q$ has full column rank, thus $rank(\mathbf{O}\mathbf{B}_f) = rank(\tilde{\mathbf{O}}\tilde{\mathbf{B}}_f) = rank(\tilde{\mathbf{B}}_{f,q})$; similar results hold for variance condition $rank(\Pi(\mathbf{O}\mathbf{B}_f)) = rank(\Pi(\tilde{\mathbf{O}}\tilde{\mathbf{B}}_f)) = rank(\Pi(\tilde{\mathbf{B}}_{f,q}))$. Based on these results, we have the following corollary.

Corollary 1. *Let* $\mathbf{G}^{-1}$ *be the state transformation to the form of Equation* (*13*) *and* $\tilde{\mathbf{B}}_{\mathrm{f,q}}$ *be the first* q *rows of the transformed fault matrix* $\mathbf{G}^{-1}\mathbf{B}_{\mathrm{f}}$, *then*:

1. The system is mean detectable, if $\tilde{\mathbf{B}}_{f,q}$ has full column rank.
2. The system is variance detectable, if the columns of $\tilde{\mathbf{B}}_{f,q}$ do not meet all the following three conditions (define $\tilde{\mathbf{B}}_{f,q} = [\mathbf{t}_1\ \mathbf{t}_2 \cdots \mathbf{t}_p]$):
   C1. There exist two disjoint subsets of column vectors of $\tilde{\mathbf{B}}_{f,q}$, $\{\mathbf{t}_{i1}, \mathbf{t}_{i2}, ..., \mathbf{t}_{ih}\} \cap \{\mathbf{t}_{j1}, \mathbf{t}_{j2}, ..., \mathbf{t}_{jg}\} = \phi$, and one subset can be represented as the linear combination of the other: $[\mathbf{t}_{i1}\ \mathbf{t}_{i2} ... \mathbf{t}_{ih}] = [\mathbf{t}_{j1}\ \mathbf{t}_{j2} \cdots \mathbf{t}_{jg}]\boldsymbol{\beta}$.
   C2. $g \leq h$.
   C3. There exists a non-zero diagonal matrix $\boldsymbol{\Lambda}$ such that $\boldsymbol{\beta}\boldsymbol{\Lambda}\boldsymbol{\beta}^T$ is diagonal.

Proof.

1. Since $rank(\mathbf{OB}_f) = rank(\tilde{\mathbf{O}}\tilde{\mathbf{B}}_f) = rank\,(\tilde{\mathbf{B}}_{f,q})$, therefore, the mean detectable condition $rank(\mathbf{OB}_f) = p$ is equivalent to $\tilde{\mathbf{B}}_{f,q}$ has full column rank.
2. With $rank(\Pi(\mathbf{OB}_f)) = rank(\Pi(\tilde{\mathbf{O}}\tilde{\mathbf{B}}_f)) = rank\,(\Pi(\tilde{\mathbf{B}}_{f,q}))$, we can also obtain the sufficient and necessary variance detectability condition to be $rank(\Pi(\tilde{\mathbf{B}}_{f,q})) = p$. Zhou *et al.* (2003) pointed out that:

$$rank(\Pi(\tilde{\mathbf{B}}_{f,q})) = rank\left(\begin{bmatrix} (\mathbf{t}_1^{\mathrm{T}}\mathbf{t}_1)^2 & (\mathbf{t}_1^{\mathrm{T}}\mathbf{t}_2)^2 & \cdots & (\mathbf{t}_1^{\mathrm{T}}\mathbf{t}_p)^2 \\ (\mathbf{t}_2^{\mathrm{T}}\mathbf{t}_1)^2 & (\mathbf{t}_2^{\mathrm{T}}\mathbf{t}_2)^2 & \cdots & (\mathbf{t}_2^{\mathrm{T}}\mathbf{t}_p)^2 \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{t}_p^{\mathrm{T}}\mathbf{t}_1)^2 & (\mathbf{t}_p^{\mathrm{T}}\mathbf{t}_2)^2 & \cdots & (\mathbf{t}_p^{\mathrm{T}}\mathbf{t}_p)^2 \end{bmatrix}\right). \tag{15}$$

Based on the study reported in Apley and Ding (2005), the right-hand side of Equation (15) being p is equivalent to the conditions C1 to C3 not being simultaneously satisfied. This completes our proof. ■

Although conditions C1 to C3 seem more complicated than the condition of $rank(\Pi(\mathbf{OB}_f)) = p$, Corollary 1 provides us with an easy way to check whether the system is variance detectable based on the system matrices directly without Π transformation. As we will see in the next section, this is especially helpful when we want to design a system matrix to improve system detectability.

2.3. *Application of detectability analysis in system design*

Based on the lemmas and corollary introduced in the previous section, we can easily check whether the system is detectable given the state space model. A point that needs mentioning is that in Section 2.2 it is assumed that different faults are not cross-correlated. However, this restriction can be relaxed once we know their cross-correlation structure. If the covariance matrix of $\mathbf{f}$ is not diagonal, there always exists a linear transformation of $\mathbf{f}$ to make the covariance matrix of the transformed variables a diagonal matrix. Then, the above results on detectability analysis can still be applied. It is also clear that mean detectability always implies variance detectability. Therefore, if the system is mean detectable, it is always variance detectable; but not *vice versa*. For Autoregressive Moving Average (ARMA) or Vector ARMA (VARMA) models, we first transform them into a state space model as summarized in Aoki (1990), and then use the lemmas above to study their detectability conditions. For the ARMA model, since there is only one univariate input variable, we could always use minimal realization to make the state space model observable. Therefore, the ARMA model should always be both mean and variance detectable. However, for the VARMA model, because there are multiple inputs and outputs, the situation becomes more complicated. There is no simple conclusion about the detectability of VARMA models. The above lemmas need to be applied to get the results for individual cases.

If the system is not detectable due to the design, we can find possible solutions to revise the system design based on the guidelines provided by the detectability analysis. We would like to articulate this point as follows.

For mean detectability, the guideline is intuitive: just modify $\tilde{\mathbf{B}}_{f,q}$ to make $\tilde{\mathbf{B}}_{f,q}$ be full column rank. If we can make the system mean detectable, the variance detectability would be automatically ensured. There are also cases in which the mean detectability is not important but the variance detectability is required. It is still possible to make the variance detectable even if the mean detectability cannot be reached. The variance detectability requires $\Pi(\tilde{\mathbf{B}}_{f,q})$ be full rank. However, the relationship between the changes in the matrix $\tilde{\mathbf{B}}_{f,q}$ and the rank of $\Pi(\tilde{\mathbf{B}}_{f,q})$ is not clear. Corollary 1 provides us with some guidelines to look into this relationship. The corollary indicates that the rank deficiency of $\Pi(\tilde{\mathbf{B}}_{f,q})$ is caused by the existence of two column subsets which meet the conditions C1 to C3 simultaneously. Therefore, when the variance change is not detectable, we can identify these subsets and redesign them accordingly to make them violate those conditions. An example is demonstrated in the numerical case study section. The above mentioned principles are summarized into the following procedures.

If the system is not mean detectable, the following steps can be followed.

1. Transform the system into the form of Equation (13), and get $\tilde{\mathbf{B}}_{f,q}$, which is $q \times p$ matrix.
2. If $q < p$, only revising $\mathbf{B}$ is not enough. The sensor placement needs to be revised, i.e., to modify the $\mathbf{C}$ matrix to increase the observability index to at least p. If this does not happen then the system cannot be mean detectable given the current dimension of state, input and output.
3. If $q \geq p$ or procedure 2 succeeds, but the new $\tilde{\mathbf{B}}_{f,q}$ is still rank deficient, revise the corresponding columns of $\tilde{\mathbf{B}}_{f,q}$ to make it full rank.
4. Transform the $\tilde{\mathbf{B}}_{f,q}$ back to the original coordination to get $\mathbf{B}$, and double check whether this revised design can be implemented. If not, return to procedure 3 for another optional design.

If the system is not variance detectable, and $\tilde{\mathbf{B}}_{f,q}$ is a $q \times p$ matrix, the following steps can be followed.

1. If $p > \max(2q, q(q+1)/2)$, then revise the $\mathbf{C}$ matrix to increase the observability index q to violate the inequality. If failed to do so, the system could not be variance detectable.
2. Otherwise, find the column subsets that can meet conditions C1 to C3 in the corollary simultaneously, and then revise the corresponding vectors to make them violate at least one condition.

3. Transform $\tilde{\mathbf{B}}_{f,q}$ back to $\mathbf{B}$ to see if this design can be implemented. If not, return to step 2 and get another optional design.

We would like to mention that we can change the entries of $\tilde{\mathbf{B}}_{f,q}$ freely by only modifying the **B** matrix, however, we could not change its dimension in this way. If the dimension of $\tilde{\mathbf{B}}_{f,q}$ is not capable of being detectable, we need to increase the observability index by changing **A** and **C**. We can also change the entries of $\tilde{\mathbf{B}}_{f,q}$ by changing **A** and **C**, but this method is not so straightforward compared with changing the **B** matrix. It can also be noted that although the design of mean detectability seems straightforward, the one for variance detectability may need several trials. The following section presents a numerical case study to demonstrate the effectiveness of the above methods.

3. Numerical case study

3.1. *Case studies for detectability analysis*

In this section, we will illustrate the effectiveness of our detectability analysis and its applications in dynamic system monitoring. The monitoring of multivariate dynamic processes such as chemical processes is quite difficult because of the autocorrelated structures of the measurement data. As widely adopted, the dynamic processes are usually modeled as state space models in statistical monitoring literature:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{f}_k, \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \boldsymbol{\varepsilon}_k. \end{cases} \tag{16}$$

Clearly, model (16) is a special case of model (2). This model has been used to monitor chemical processes such as the HTST Pasteurization process (Negiz and Cinar,1997; Kosebalaban and Cinar, 2001; Lee *et al.*, 2004). In HTST processes, measurements such as product temperature, inlet temperature, residence time and steam temperature are collected. The process dynamic comes from the heat transfer between different flows. Typical faults in this process include a high variability of the inlet temperature, variable constant flow rate, and so forth. Negiz (1995) used conservation laws to build up the mathematical model and the canonical variates method to get a data-driven model in the form of Equation (16). One of the monitoring techniques for this model is to monitor the residuals of this model using a T^2 statistic. In our case study, we will adopt the same model structure as given in Equation (16). However, we will make necessary modifications on the dimensions and entries of the system matrices to better illustrate the different diagnosability conditions and their impacts on the effectiveness of monitoring schemes. The example we will use to demonstrate our analysis is shown in Equation (17) and the residual-based T^2 chart will serve as the monitoring scheme for detectability comparison.

$$\begin{cases} \mathbf{x}_{k+1} = \begin{bmatrix} -0.5 & -1 & 1 & -1 \\ 0 & -1.5 & 1 & -1 \\ 1 & -3 & 2.5 & -3 \\ 1 & -2 & 2 & -2.5 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & -1 \end{bmatrix} \mathbf{f}_k, \\ \mathbf{y}_k = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & -1 \\ 2 & -3 & 1 & -1 \end{bmatrix} \mathbf{x}_k + \boldsymbol{\varepsilon}_k. \end{cases} \tag{17}$$

Without loss of generality, we assume that $\mathbf{f}_k$ and $\boldsymbol{\varepsilon}_k$ follow a multivariate normal distribution with zero mean and covariance matrix as identity under normal working conditions. According to Corollary 1, the model is first transformed into the form of Equation (13), and $\tilde{\mathbf{B}}_{f,q}$ can be obtained as

$$\tilde{\mathbf{B}}_{f,q} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 3 & 2 \end{bmatrix}. \tag{18}$$

It is easy to verify that $rank(\tilde{\mathbf{B}}_{f,q}) = 2$, and $rank(\Pi(\tilde{\mathbf{B}}_{f,q})) = 3$. Based on Lemmas 1 and 2, it can be concluded that the system is neither mean detectable nor variance detectable. Alternatively, we can also check the variance detectability by looking into the dependence structure of $\tilde{\mathbf{B}}_{f,q}$ directly by using Corollary 1. It is easy to find that:

$$\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \times \begin{bmatrix} -0.5 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & -0.5 \end{bmatrix}. \tag{19}$$

In other words, there exist two column subsets which satisfy conditions C1 to C3, therefore, the system is not variance detectable.

In order to illustrate the consequences of the non-detectability, we established a residual-based T^2 control chart developed by Kosebalaban and Cinar (2001) to monitor the process. In the monitoring scheme, first, the Kalman filter method is used to estimate the state variable $\hat{\mathbf{x}}_k$, then the output residual $\mathbf{e}_k = \mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k$ is computed. Finally a T^2 statistic of $\mathbf{e}_k$ is established for monitoring, i.e., we monitor the variable $T_k = (\mathbf{e}_k - \bar{\mathbf{e}})^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}(\mathbf{e}_k - \bar{\mathbf{e}})$, where $\bar{\mathbf{e}}$ and $\boldsymbol{\Sigma}_k$ are the mean and covariance of the residuals respectively under normal conditions. If the population values of $\bar{\mathbf{e}}$ and $\boldsymbol{\Sigma}_k$ are known, the T^2 statistic of the residuals follows a χ^2 distribution. In this numerical study, we use a stationary Kalman filter to estimate state vectors, and then we use the steady-state covariance of the residuals, denoted as $\boldsymbol{\Sigma}_e$, to construct a T^2 control chart. These approximations would make the T^2 statistic deviate from the χ^2 distribution slightly. However, because the setup of the T^2 control chart is not the focus of this article and for the sake of simplicity, we will ignore the influence of these approximations and compute the upper control limits still based on the χ^2

Table 1. ARL and EMM of mean change with different direction and strength

	Magnitude = 3.87		*Magnitude = 7.75*		*Magnitude = 15.49*	
Direction	*ARL*	*EMM*	*ARL*	*EMM*	*ARL*	*EMM*
[−3 −2 1 1]	310.260	0	315.740	0	315.690	0
[3 2 −1 1]	84.946	0.367 39	14.397	1.4696	2.472	5.8783
[1 −2 −3 1]	13.609	1.4696	2.527	5.8783	1.999	23.513
[1 1 3 −2]	13.334	1.513	2.437	6.0521	2.000	24.208
[3 −2 1 1]	9.181	2.1581	2.132	8.6323	1.998	34.529
[−3 1 −2 1]	8.677	2.1581	2.116	8.6323	2.000	34.529
[3 2 1 1]	5.858	2.5175	2.050	10.07	2.000	40.28
[1 1 −3 −2]	4.367	3.2213	2.016	12.885	1.999	51.541
[1 −1 3 2]	2.865	4.9886	2.000	19.954	1.999	79.818
[1 1 3 2]	2.560	5.6644	2.000	22.658	2.000	90.63

distribution. In this study, we choose $\alpha = 0.0033$, and the corresponding upper control limit is $UCL = 13.70$. The in-control average run length is around 300.

The system given in model (17) is neither mean nor variance detectable, and consequently according to the definition of detectability, it can be concluded that some directions of changes cannot be detected by any monitoring schemes. To demonstrate this point, we arbitrarily change the directions of the fault vectors but keep their magnitudes constant, and use the average run length of the control chart to evaluate the detection efficiency. Table 1 lists the average run length after 1000 simulations to detect a fault with different magnitude and directions. In Table 1, an index denoted as EMM is also used to characterize the effective mean change magnitude. If we denote $\mathbf{q} = \mathbf{C} \times (\sum_{i=0}^{n-1} \mathbf{A}^i) \times \mathbf{B} \times (\mathbf{f} - \mathbf{f}_0)$, then EMM can be expressed as: $\text{EMM}(\mathbf{f}) = \mathbf{q}^{\text{T}} \boldsymbol{\Sigma}_e^{-1} \mathbf{q}$, where $\mathbf{f} - \mathbf{f}_0$ is the mean shift vector, and $\boldsymbol{\Sigma}_e$ is the estimated covariance matrix of the residual in normal conditions. In fact, for a given system and mean shift vector, this index is equivalent to the statistical distance between $\mathbf{f}_1$ and $\mathbf{f}_0$. This quantity takes into account both the magnitude and direction of the fault vectors, and thus can fully represent the influence of the changes. The numerical study indicates that the EMM index is closely related to the average run length: when EMM is small, the ARL is large; when the EMM is large, the ARL is small. Thus, the EMM value is a good indicator of the sensitivity of the monitoring scheme. Clearly, if the system is not detectable, then for a certain mean shift vector the EMM index will be zero.

For illustration purposes, control charts for the detection of two different mean shifts are illustrated in Fig. 2. In the figure, the control chart can clearly detect the mean shift to $E(\mathbf{f}_2)$, whereas it is unable to detect the mean shift to $E(\mathbf{f}_1)$. This situation is exactly as we predict through mean detectability analysis, and is consistent with the results in the table.

The Average Run Length (ARL) of the T^2 control chart under different conditions of covariance changes is also identified through simulation. The results are listed in Table 2. The results are similar to those of the mean shift cases, i.e., for a non-detectable system, the sensitivity of the monitoring scheme is different for different change directions and for some particular directions, the monitoring scheme fails to detect the change. In principle, we could also develop an index to indicate the monitoring sensitivity for different change directions. However, because the T^2 chart was mainly developed for mean shift detection, such an index of variance change detection is not needed and is somewhat difficult to derive. Again, for illustration purposes, control charts for the detection of two different variance change conditions are illustrated in Fig. 3.

3.2. *System design improvement based on detectability analysis*

When a system is not detectable, the detectability condition of the system can be modified by redesigning the system matrices.

We still use the system of Equation (17) as an example. Since the row number of $\tilde{\mathbf{B}}_{f,q}$ is smaller than the column number, the $\mathbf{C}$ matrix needs to be revised to increase the observability index. Actually, if the $\mathbf{C}$ matrix in Equation (17)

Table 2. ARL of variance change with different direction and magnitude

Direction of *diag*$(\boldsymbol{\Sigma}_f - \boldsymbol{\Sigma}_{f0})$	*Magnitude = 0.625*	*Magnitude = 1.25*	*Magnitude = 1.875*
[0.5 −0.25 −0.25 0.5]	311.07	308.76	302.16
[0.5 0.25 −0.25 0.5]	267.01	214.07	203.53
[0.5 0.25 0.25 0.5]	126.51	72.286	42.366
[0.25 0.5 0.5 0.25]	117.06	54.044	31.993
[−0.25 0.5 0.5 −0.25]	166.02	96.948	62.198
[0.25 0.25 0.5 0.5]	112.57	46.848	29.466
[0.25 −0.25 0.5 0.5]	116.62	57.747	33.769
[−0.25 −0.25 0.5 0.5]	138.63	70.808	39.737

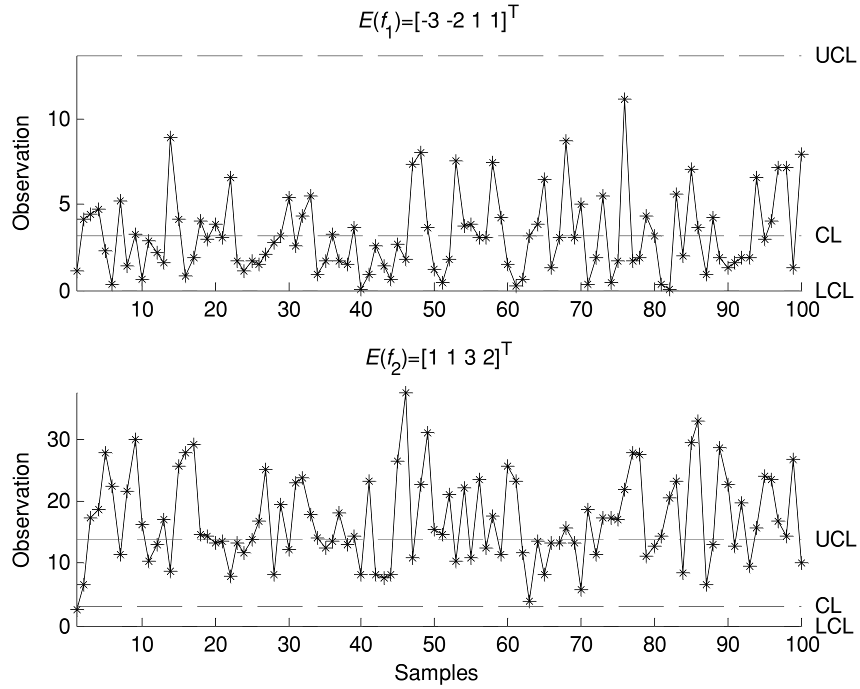

Fig. 2. Comparisons of the detection of two different mean shifts.

is revised to the following matrix, which only involves very minor changes, the system can be made fully observable:

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -2 & 2 & -2 \\ 2 & -3 & 1 & -2 \end{bmatrix}. \tag{20}$$

In this case, $\tilde{\mathbf{B}}_{f,q} = \mathbf{B}_f$, and it can also be checked that $rank(\mathbf{B}_f) = 4$, therefore the system is both mean detectable and variance detectable. In this example, we do not need to further revise the $\mathbf{B}_f$ matrix. We would also like to mention that in most cases, the detectability condition can be

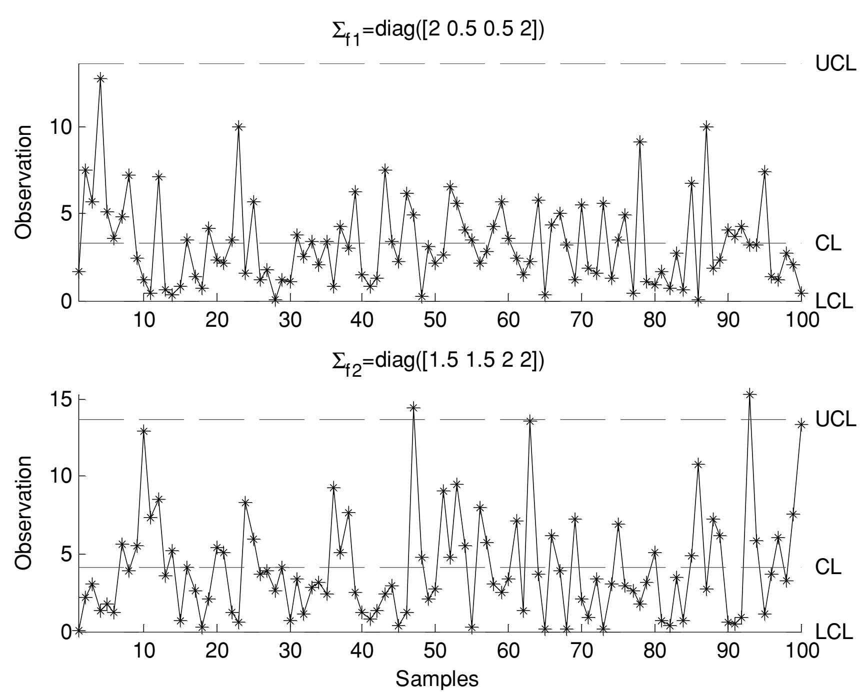

Fig. 3. Comparison of the results of the detection of two different variance shifts.

improved by slightly revising the system matrix because one or two entry changes in the matrix are usually enough to make the system detectable. To illustrate this point, we can evaluate and compare the following three $\tilde{\mathbf{B}}_{f,q}$ matrices:

$$\tilde{\mathbf{B}}_{f1} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{bmatrix}, \tilde{\mathbf{B}}_{f2} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 2 & 1 \end{bmatrix},$$
$$\tilde{\mathbf{B}}_{f3} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 3 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 1 \end{bmatrix}. \tag{21}$$

In the first case, $\tilde{\mathbf{B}}_{f1}$ is of full rank, so the system is both mean and variance detectable; in the second case, $\tilde{\mathbf{B}}_{f2}$ is not full rank, so it is not mean detectable. However, we could not find subsets of columns satisfying conditions C1 to C3. Therefore, the system is variance detectable. In the third case, $\tilde{\mathbf{B}}_{f3}$ is still rank-deficient. Furthermore, it can be shown that the third and fourth columns of $\tilde{\mathbf{B}}_{f3}$ fall into the space spanned by the first and second columns, and these two subsets satisfy conditions C1 to C3. Therefore, the system with $\tilde{\mathbf{B}}_{f3}$ is neither mean nor variance detectable. We can then follow the design guidelines in Section 2.3 to revise the $\tilde{\mathbf{B}}_{f,q}$ matrix to make the system detectable.

These examples demonstrate that although these $\tilde{\mathbf{B}}_{f,q}$ matrices have very similar structures with only one or two entries being different, the detectability results are quite different. This provides us with the possibility to change the detectability condition by only slightly modifying the system matrices.

4. Concluding remarks and future work

This paper investigates detectability issues in multivariate dynamic systems. In this paper, process faults are represented by a random vector input to the dynamic system with possible mean and variance changes. The relationship between the intrinsic detectability and system structures is investigated and criteria for checking detectability are established. The major results of this paper are summarized into two lemmas and one corollary. These results can be used for system detectability analysis and provide guidelines for system design.

There are still some open issues in the proposed methods. In our detectability definition, we have no restriction on the change types of the faults, which can be step change or drifting change or others. However, for easy checking of the detectability criteria we made the assumption that all the changes are step changes. Although the conditions are still applicable to linear drifting changes, they cannot handle more sophisticated changes. It would therefore be of interest to consider more general cases. In our analysis, the system parameters are also assumed to be precisely known. If the system model is driven from physical analysis, this might be true. However, in a real application where it is unlikely that we can obtain a physical process model, the robustness of the detectability analysis needs further attention. Furthermore, although the detectability analysis can identify which vectors can be detected, it does not provide a continuous indicator of how sensitive the detection scheme is for a specific process change. Therefore, it is desirable to study sensitivity issues in change detection, including both mean shift and variance shift. However, unlike the detectability study, the sensitivity will have a close relationship with the particular change detection algorithms. We will concentrate on these open issues, from both theoretical and practical perspectives, in our future studies.

Acknowledgement

Financial support of this work was provided by NSF grants CMMI-0545600 and CMMI-0529327. The authors thank the editors and reviewers for their valuable comments and suggestions.

References

Alwan, L.C. and Roberts, H.V. (1989) Time series modeling for statistical process control. *Journal of Business & Economic Statistics*, **6**, 87–95.

Aoki, M. (1990) *State Space Modeling of Time Series*, second edition, Springer-Verlag, New York, NY.

Apley, D.W. and Ding, Y. (2005) A characterization of diagnosability conditions for variance components analysis in assembly operations. *IEEE Transactions on Automation Science and Engineering*, **2**(2), 101–110.

Bakshi, B.R. (1998) Multiscale PCA with application to multivariate statistical process monitoring. *American Institute of Chemical Engineering Journal*, **44**(7), 1596–1610.

Basseville, M. (2001) On fault detectability and isolability. *European Journal of Control*, **7**, 625–637.

Basseville, M. and Nikiforov, I.V. (1993) *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, Upper Saddle River, NJ.

Box, G. and Kramer, T. (1992) Statistical process monitoring and feedback adjustment—a discussion. *Technometrics*, **34**(3), 251–267.

Ceglarek, D. and Shi, J. (1995) Dimensional variation reduction for automotive body assembly. *Manufacturing Review*, **8**, 139–154.

Chen, S.G., Ulsoy, A.G. and Koren, Y. (1998) Error source diagnostics using a turning process simulator. *Journal of Manufacturing Science and Engineering*, **120**(2), 409–416.

Ding, T., Shi, J. and Ceglarek, D. (2002) Diagnosability analysis of multistage manufacturing processes. *ASME Journal of Dynamic Systems, Measurement and Control*, **124**, 1–13.

Gustafsson, F. (2002) Stochastic fault diagnosability in parity spaces. Presented at the 15th IFAC World Congress, Barcelona, July 2002.

Harris, T.J. and Ross, W.H. (1991) Statistical process control procedures for correlated observations. *Canadian Journal of Chemical Engineering*, **69**(1), 48–57.

Kosebalaban, F. and Cinar, A. (2001) Integration of multivariate SPM and FDD by parity space technique for a food pasteurization process. *Computers and Chemical Engineering*, **25**, 473–491.

Ku, W., Storer, R.H. and Georgakis, C. (1995) Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **30**(1), 179–196.

Lai, T.L. (1995) Sequential changepoint detection in quality control and dynamic systems. *Journal of the Royal Statististical Society, Series B* (*Methodological*), **57**(4), 613–658.

Lancaster, P. (1969) *Theory of Matrices*, first edition, Academic Press, New York, NY.

Lee, J.M., Yoo, C.K. and Lee, I.B. (2004) Statistical process monitoring with independent component analysis. *Journal of Process Control*, **14**(5), 467–485.

Li, W., Yue, H., Valle-Cervantes, S. and Qin, S. (2000) Recursive PCA for adaptive process monitoring. *Journal of Process Control*, **10**(5), 471–486.

Liu, B.F. and Si, J. (1997) Fault isolation filter design for linear time-invariant systems. *IEEE Transaction on Automatic Control*, **42**(5), 704–707.

MacGregor, J.F. (1988) On-line statistical process control. *Chemical Engineering Progress*, **84**, 21–31.

Massoumnia, M., Verghese, G.C. and Willsky, A.L. (1989) Failure detection and identification. *IEEE Transaction on Automatic Control*, **34**(3), 316–321.

Montgomery, D.C. (2005) *Introduction to Statistical Quality Control*, fifth edition, Wiley, New York, NY.

Montgomery, D.C. and Mastrangelo, C.M. (1991) Some statistical process control methods for autocorrelated data. *Journal of Quality Technology*, **23**, 179–193.

Negiz, A. (1995) Statistical dynamic modeling and monitoring methods for multivariable continuous processes. PhD Dissertation, Illinois Institute of Technology, Chicago, IL.

Negiz, A. and Cinar, A. (1997) Statistical monitoring of multivariable dynamic processes with state-space models. *AIChE Journal*, **43**(8), 2002–2020.

Rao, C.R. and Kleffe, J. (1988) *Estimation of Variance Components and Applications*, North-Holland, Amsterdam, The Netherlands.

Rugh, W.J. (1996) *Linear System Theory*, second edition, Prentice Hall, Upper Saddle River, NJ.

Simoglou, A., Martin, E.B., and Morris, A.J. (2002) Statistical performance monitoring of dynamic multivariate processes using state space modeling. *Computers and Chemical Engineering*, **26**(6), 909–920.

Vander Wiel, S.A. (1996) Monitoring processes that wander using integrated moving average models. *Technometrics*, **38**(2), 139–151.

Wardell, D.G., Moskowitz, H. and Plante, R.D. (1994) Run-length distributions of special cause control charts for correlated processes. *Technometrics*, **36**(1), 3–17.

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.

Biographies

Nan Chen is a Ph.D. student in the Department of Industrial and Systems Engineering at University of Wisconsin-Madison. He received his B.S. degree in Automation at Tsinghua University, China in 2006. Currently, his research interests include modeling, monitoring, and diagnosis in complex engineering systems, especially discrete event dynamic systems.

Shiyu Zhou is an associate professor in the Department of Industrial and Systems Engineering at the University of Wisconsin-Madison. He received his B.S. and M.S. in Mechanical Engineering at University of Science and Technology of China in 1993 and 1996 respectively, and received his Master's in Industrial Engineering and Ph.D. in Mechanical Engineering at the University of Michigan in 2000. Dr. Zhou's research interests include the in-process quality and productivity improvement methodologies by integrating statistics, system and control theory, and engineering knowledge. His research is sponsored by National Science Foundation, Department of Energy, NIST-ATP, and industries. He is a recipient of the CAREER Award from the National Science Foundation and the Best Application Paper award from *IIE Transactions* in 2006. Dr. Zhou is a member of IIE, INFORMS, ASME, and SME.