A Comparison of Process Variation Estimators for In-Process Dimensional Measurements and Control

Yu Ding
Department of Industrial Engineering, Texas A&M University, College Station, TX 77843-3131

Shiyu Zhou
Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI 53706-1572

Yong Chen
Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City, IA 52242

Dimensional variation reduction is critical to assure high product quality in discrete-part manufacturing. Recent innovations in sensor technology enable in-process implementation of laser-optical coordinate sensors and continuous monitoring of product dimensional quality. The abundance of measurement data provides an opportunity to develop next generation process control technologies that not only detect process change, but also provide guidelines respective of root cause identification. Given continuous product dimensional measurements, a critical step leading to root cause identification is the variance estimation of process variation sources. A few on-line variance estimators are available. The focus of this paper is to study the interrelationships and properties of the available variance estimators and compare their performance. An operating characteristics curve is developed as a convenient tool to guide the appropriate use of on-line variance estimators under specific circumstances. The method is illustrated using examples of dimensional control for a panel assembly process. [DOI: 10.1115/1.1870041]

1 Introduction

Dimensional integrity is a major quality concern in many discrete-part manufacturing processes, such as assembly and machining. The dimensional variation of a product is affected by many sources of process variability, e.g., positioning of fixture locators, alignment of machine tools, and random deformation of compliant parts. In order to meet high product quality standards and reduce defect-induced downtime, manufacturers tend to deploy in-process laser-optical coordinate sensors, a recent innovation, to obtain 100% inspection of product dimensional quality characteristics [1]. The in-process measurements collected by these automatic-sensing devices yield exceptional opportunities for identification of process variation sources responsible for product quality problems.

To illustrate, consider the in-process coordinate measurement system used in the automotive assembly process in Fig. 1. In this process, laser-optical coordinate sensors are installed in a space frame and measure each automobile body assembly that passes through. The entire system (coordinate sensors and sensor frame housing) is referred to as an optical coordinate measuring machine (OCMM). The in-process OCMM can perform parallel measurements of multiple product characteristics and is therefore capable of measuring as many as 150 product features a minute on the same automobile body. By comparison, a mechanical coordinate measuring machine (CMM), which takes measurements sequentially by means of a touch stylus, is substantially slower than an in-process OCMM. A CMM can measure only 6–8 automobile body assemblies per day in a production line having a daily throughput of 1000 units.

Given continuous dimensional measurements, statistical control chart methods [2] can be employed to monitor part quality and production processes. When a process change is detected by control chart analysis, it becomes necessary to determine the appropriate corrective actions needed to restore the manufacturing system to its normal condition. This determination of corrective action is referred to as root cause analysis [3,4]. Unfortunately, control chart methods are of limited use with respect to root cause analysis; the identification of process variation sources is left primarily to human operators. Effective statistical methods that help determine process variation sources are consequently highly desirable.

A more specific example of root cause analysis can be demonstrated using the panel assembly process illustrated in Fig. 2. In panel assembly, of which automobile body assembly is characteristic, the workpiece (panel) is held by a set of fixtures. Figure 2 shows a two-dimensional (2-D) workpiece held by a set of fixtures consisting of a four-way locator (P_1), which controls motion in both x and z directions ($\delta P_1(x,z)$), and a two-way locator (P_2) which controls motion only in the z direction ($\delta P_2(z)$). In order to control the dimensional quality of the panel, coordinate sensors are used to monitor dimensional variation of key product features. In this example, three sensors are used to measure three corners of the panel, as shown in Fig. 2.

Small positional perturbation of the workpiece exists even when all locators function properly within designed tolerances. However, when there is damage to pinholes or locator wear, the workpiece will experience large random positional deviation, resulting in excessive dimensional variation of the final assembly. Suppose locator P_2 exhibits a large deviation $\delta P_2(z)$ due to pinhole damage, resulting in a large product variation detectable by the sensor at M_1. In this instance, the process variation source is fixturing variation, i.e., unacceptably low fixturing repeatability. Application of root cause analysis identifies the malfunctioning locator pair to be the cause of the excessive dimensional variability.

Since product dimensional measurement is readily available, a critical step in performing root cause analysis is to estimate the magnitude of process variation based on product measurements. If a source exhibits variability larger than a specified tolerance, it is reasonable to designate it to be the root cause of product dimensional defects.

If the relationship between product measurement and process variation sources can be represented using linear models, several variation estimators are available. The principal statistical meth-

Contributed by the Dynamic Systems, Measurement, and Control Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL. Manuscript received by the ASME Dynamic Systems and Control Division May 16, 2003; final revision March 27, 2004. Review conducted by: K. Stelson.

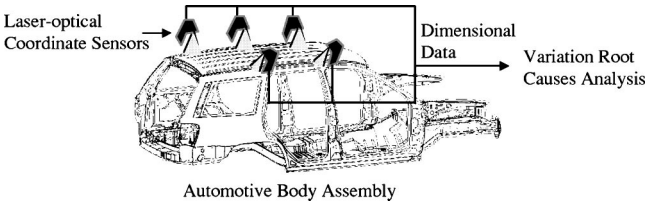

Fig. 1 Optical coordinate measuring machine for automotive body assembly

odology relevant to linear models is variance components analysis (VCA) [5,6]. Although applicable to general linear models, many VCA techniques were developed for application domains other than process quality control, primarily specifically designed off-line experiments involving random effects in one-way or two-way layout models. In addition, variance estimators developed for VCA often use iterative numerical procedures to solve the maximum likelihood estimator (MLE), restricted maximum likelihood estimator, and Rao's iterative MINQUE estimators. These methods are as well primarily applicable to off-line experiments where sample size is small and computation time is not a concern. For on-line variance estimation and quality control in large-scale systems having potentially large data samples, direct application of the VCA-theory estimators may not be appropriate since these procedures are computationally costly and time consuming.

In contrast, variance estimators have also been developed for signal processing [7–9] and quality engineering [10,11] applications. In general, these estimators entail closed-form expressions and are more cost-effective calculations than the MLE method, particularly for large sample sizes, and are thus more suitable for on-line quality control. The development of on-line variance estimators was often based on a least-squares criterion or followed empirical intuition. The properties of the resulting estimators have not been thoroughly evaluated. Issues with respect to differences in the performance of on-line estimators and selection of a proper estimator have not been clearly resolved. Consequently, variation estimation for a specific application can be misconstrued if an inappropriate estimator is used. As a result, the variation sources that result in poor product quality are still existent and the root cause analysis effort is wasted.

To resolve these issues, a comparison of process variation estimation methods is necessary. Several comparative studies have been conducted as found in the literature [12–14]. These studies, however, are not appropriate for on-line variance estimation. Rather they are based on one-way or two-way layout models, which are more applicable to off-line experimental designs than root cause identification.

This paper will address the interrelationships among on-line variance estimators, investigate the statistical properties of these estimators, compare performance, and develop convenient tools to aid practitioners in selecting the appropriate estimator for specific needs. The MLE, as developed in VCA theory, will be used as a benchmark reference for comparing the on-line variance estimators presented in this study.

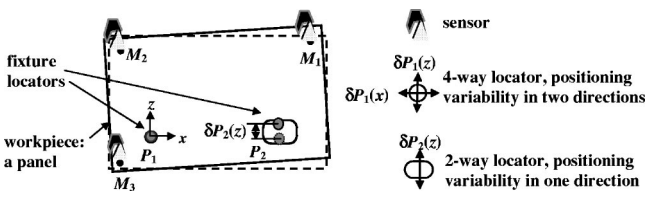

Fig. 2 A simplified fixture setup for a 2-D workpiece

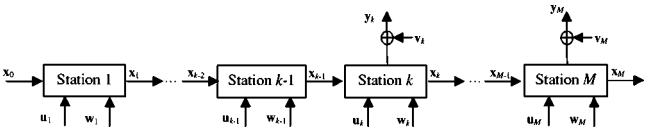

Fig. 3 Diagram of a multistation manufacturing process

Section 2 presents the formulation for root cause identification problems. Section 3 briefly reviews the on-line variance estimators and presents their interrelationships and properties. The results of a comparative analysis of variance estimator performance are presented in Sec. 4. An operating characteristics (OC) curve is introduced to determine appropriate conditions for the use of these estimators. Section 5 revisits the panel assembly process example and illustrates estimator properties and the use of the OC curve. Section 6 provides a brief summary and conclusion.

2 Formulation of the Root Cause Identification Problem

The first step in identifying the root cause of product quality problems is to establish a model that links product measurements to process variance sources. In quality control applications, deviation in product/process features is usually much smaller than nominal values. Hence, a linear model, or a linearization of nonlinear systems, is often acceptable for representation of discrete-part manufacturing processes. The panel assembly process, illustrated in Fig. 2, is used as an example to introduce symbolic notation and to demonstrate the validity of the linear model. The applicability of the resulting model is not limited to panel assembly processes.

For the panel assembly process in Fig. 2, let $\mathbf{u}$ denote deviations at fixturing points such as P_1 and P_2 and $\mathbf{y}$ denote deviations measured by coordinate sensors. In this example, $\mathbf{u}=[\delta P_1(x)\ \delta P_1(z)\ \delta P_2(z)]^T$ and $\mathbf{y}=[\delta M_1(x)\ \delta M_1(z)\ \delta M_2(x)\ \delta M_2(z)\ \delta M_3(x)\ \delta M_3(z)]^T$. Employing standard kinematics analysis, the relationship between $\mathbf{y}$ and $\mathbf{u}$ can be approximated by a linear model as

$$\mathbf{y}(i) = \mathbf{A}\mathbf{u}(i) + \mathbf{v}(i), \quad i = 1,2, \ldots, N, \tag{1}$$

where $\mathbf{y}(i) \in \mathrm{Re}^{n\times 1}$ is the measurement vector, $\mathbf{u}(i) \in \mathrm{Re}^{p\times 1}$ is the random vector of p process variation sources, and $\mathbf{v}(i) \in \mathrm{Re}^{n\times 1}$ is the additive sensor noise. Matrix $\mathbf{A}$ can be determined using kinematics analysis as found in Apley and Shi [10], Chang and Gossard [15], and Carlson et al. [16]. The index i is the sample index and N is the sample size. For the sake of convenience, n, i.e., the dimension of $\mathbf{y}$, designates the number of sensors, although it actually equals the product of the number of physical sensors and the degrees of freedom measured by each sensor.

N independent, identically distributed observations, $\{\mathbf{y}(i)\}_{i=1}^{N}$, are typically collected for variance estimation of a stationary process. Matrix $\mathbf{A}$ does not vary during the data-collecting process. Thus, model (1) is referred to as a linear model with replicated observations or a linear replicated model [5].

Actual manufacturing systems, such as automobile body assembly processes and transfer-line machining processes, typically consist of multiple stations. In a multistation process, variation sources can originate in every station and propagate along the production line. Consider the M-station process illustrated in Fig. 3, where the subscript k denotes the station index. The product dimensional deviation is represented by the state vector $\mathbf{x}_k$ and the process variation sources at station k are included as inputs $\mathbf{u}_k$. The unmodeled errors are represented by a random vector $\mathbf{w}_k$, while vector $\mathbf{v}_k$ still corresponds to sensor noise. The variation propagation in a multistation process can be modeled as

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_k + \mathbf{w}_k \quad \text{and} \quad \mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k, \tag{2}$$

where $\mathbf{A}_{k-1}\mathbf{x}_{k-1}$ represents the transformation of product dimensional deviation from station $k-1$ to station k, $\mathbf{B}_k\mathbf{u}_k$ represents product deviations resulting from process variations at station k, and $\mathbf{C}_k$ maps product dimensional states to quality measurements. Matrices $\mathbf{A}_k$, $\mathbf{B}_k$, and $\mathbf{C}_k$ are again determined through kinematics analysis. This station-indexed state space model has been used to model variation propagation in various multistation manufacturing processes, e.g., rigid-part assembly processes [17–19], compliant-part assembly processes [20], machining processes [21–23], and sheet stretch forming processes [24]. A detailed modeling procedure as well as the determination of model parameters in $\mathbf{A}_k$, $\mathbf{B}_k$, and $\mathbf{C}_k$ can be obtained from these references.

An algebraic transformation can be used to convert model (2) into a linear replicated model. First, $\mathbf{y}_k$ can be expressed as

$$\mathbf{y}_k = \Sigma_{l=1}^{k}\mathbf{C}_k\mathbf{\Phi}_{k,l}\mathbf{B}_l\mathbf{u}_l + \mathbf{C}_k\mathbf{\Phi}_{k,0}\mathbf{x}_0 + \Sigma_{l=1}^{k}\mathbf{C}_k\mathbf{\Phi}_{k,l}\mathbf{w}_l + \mathbf{v}_k, \tag{3}$$

where $\mathbf{\Phi}_{k,l} \equiv \mathbf{A}_{k-1}\mathbf{A}_{k-2}\cdots\mathbf{A}_l$ for $k>l$ and $\mathbf{\Phi}_{k,k} \equiv \mathbf{I}$. Vector $\mathbf{x}_0$ corresponds to the initial condition of a product prior to entering the manufacturing line. If $\mathbf{x}_0$ is available, $\mathbf{C}_k\mathbf{\Phi}_{k,0}\mathbf{x}_0$ can be moved to the left side of Eq. (3) and $\mathbf{y}_k - \mathbf{C}_k\mathbf{\Phi}_{k,0}\mathbf{x}_0$ can then be treated as a new measurement. If $\mathbf{x}_0$ is not available, it can be treated as an additional variation source. Without loss of generality, $\mathbf{x}_0$ can be set to $\mathbf{0}$. Combining all available measurements from station 1 through station M, yields

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_M\end{bmatrix} = \mathbf{\Gamma}\cdot\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\\ \vdots\\ \mathbf{u}_M\end{bmatrix} + \mathbf{\Psi}\cdot\begin{bmatrix}\mathbf{w}_1\\ \mathbf{w}_2\\ \vdots\\ \mathbf{w}_M\end{bmatrix} + \begin{bmatrix}\mathbf{v}_1\\ \mathbf{v}_2\\ \vdots\\ \mathbf{v}_M\end{bmatrix}, \tag{4}$$

where

$$\mathbf{\Gamma} = \begin{bmatrix}\mathbf{C}_1\mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{C}_2\mathbf{\Phi}_{2,1}\mathbf{B}_1 & \mathbf{C}_2\mathbf{B}_2 & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{C}_M\mathbf{\Phi}_{M,1}\mathbf{B}_1 & \mathbf{C}_M\mathbf{\Phi}_{M,2}\mathbf{B}_2 & \cdots & \mathbf{C}_M\mathbf{B}_M\end{bmatrix},$$

$$\mathbf{\Psi} = \begin{bmatrix}\mathbf{C}_1 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{C}_2\mathbf{\Phi}_{2,1} & \mathbf{C}_2 & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{C}_M\mathbf{\Phi}_{M,1} & \mathbf{C}_M\mathbf{\Phi}_{M,2} & \cdots & \mathbf{C}_M\end{bmatrix},$$

and $\mathbf{C}_k = \mathbf{0}$ if no measurement is available at station k.

The model in Eq. (4) becomes equivalent to the linear replicated model in Eq. (1) if $\mathbf{y}^T = [\mathbf{y}_1^T \cdots \mathbf{y}_M^T]$, $\mathbf{u}^T = [\mathbf{u}_1^T \cdots \mathbf{u}_M^T | \mathbf{w}_1^T \cdots \mathbf{w}_M^T]$, $\mathbf{v}^T = [\mathbf{v}_1^T \cdots \mathbf{v}_M^T]$, and $\mathbf{A} = [\mathbf{\Gamma}\ \mathbf{\Psi}]$. It should be noted that the sample index is not explicitly defined in the above equations.

The linear replicated model in Eq. (1) represents the relationship between process variation and product measurement in a manufacturing process and serves as the basis for variance estimation. For the purposes of this study, it is assumed that:

(1) The underlying distributions of $\mathbf{u}$ and $\mathbf{v}$ are normal,
(2) Noise vector $\mathbf{v}$ has a zero mean, is independent of $\mathbf{u}$, and has the variance-covariance matrix $\sigma_v^2\mathbf{I}_n$ ($\mathbf{I}_n$ is an $n\times n$ identity matrix), where σ_v^2 is sensor noise variance,
(3) The p variation sources are independent, such that $\mathbf{u}$ has a diagonal covariance matrix $\mathbf{\Sigma}_u = \text{diag}\{\sigma_1^2\ \sigma_2^2 \cdots \sigma_p^2\}$, where σ_i^2 is the variance of the ith input in $\mathbf{u}$. It is further assumed that $\mathbf{u}$ has a zero mean since it represents the deviation from the designed nominal position. A brief discussion that considers a nonzero mean $\mathbf{u}$ is presented in Sec. 4.

Given these assumptions, the covariance relationship between quality measurements (having zero mean) and process variation sources can be obtained from model (1) as

$$\mathbf{\Sigma}_y = \mathbf{A}\mathbf{\Sigma}_u\mathbf{A}^T + \sigma_v^2\mathbf{I}_n = \Sigma_{j=1}^{p+1}\sigma_j^2\mathbf{V}_j, \tag{5}$$

where $\mathbf{V}_j = \mathbf{a}_j\mathbf{a}_j^T$ for $j=1,\ldots,p$, $\mathbf{a}_j$ is the jth column vector of $\mathbf{A}$, $\mathbf{V}_{p+1} = \mathbf{I}_n$, and $\sigma_{p+1}^2 = \sigma_v^2$. $\{\sigma_j^2\}_{j=1}^{p+1}$ constitute the *variance components* of process variation sources (including sensor noise variance). The identification of variation sources is then achieved by estimating variance components from observations $\{\mathbf{y}(i)\}_{i=1}^{N}$. Section 3 presents several available on-line variance estimators and their properties.

3 Variance Estimators and Their Properties

3.1 Brief Review of Variance Estimators. The following discussion briefly reviews four types of on-line variance estimators and the MLE method for variance estimation. All discussion is based on model (1).

3.1.1 Least-Squares (LS) fit estimator. The basic premise of the LS fit estimator is to minimize the sum of the squared difference between the estimate $\hat{\mathbf{\Sigma}}_y \equiv \Sigma_{j=1}^{p+1}\hat{\sigma}_j^2\mathbf{V}_j$ and the observation covariance matrix $\mathbf{S}_y$, where $\mathbf{S}_y$ is defined as

$$\mathbf{S}_y \equiv \frac{1}{N}\Sigma_{i=1}^{N}(\mathbf{y}(i)\mathbf{y}(i)^T). \tag{6}$$

Namely,

$$\min\|\hat{\mathbf{\Sigma}}_y - \mathbf{S}_y\|^2, \tag{7}$$

where $\|\cdot\|$ is the Euclidean norm of a matrix (i.e., the Frobenius norm) and minimization is performed with respect to the estimates of variance components, $\{\hat{\sigma}_j^2\}_{j=1}^{p+1}$. Denoting $\boldsymbol{\sigma}^2 \equiv [\sigma_1^2 \cdots \sigma_{p+1}^2]^T$ and $\hat{\boldsymbol{\sigma}}^2$ as its estimated value, D'Assumpcao [7] and Bohme [8] derived the LS fit estimator as

$$\{\text{tr}(\mathbf{V}_j\mathbf{V}_i)\}_{i,j=1}^{p+1}\cdot\hat{\boldsymbol{\sigma}}^2 = \{\text{tr}(\mathbf{V}_i\mathbf{S}_y)\}_{i=1}^{p+1}, \tag{8}$$

where $\text{tr}(\cdot)$ is the matrix trace, $\{\cdot\}_{i,j=1}^{p+1}$ is a $(p+1)\times(p+1)$ matrix, and $\{\cdot\}_{i=1}^{p+1}$ is a $(p+1)\times 1$ column vector.

3.1.2 Estimator in Ding, Shi, and Ceglarek [11]. Let $\text{vec}(\cdot)$ denote an operator that stacks the columns of a matrix on top of one another, e.g., $\text{vec}(\mathbf{S}) = [s_{11}\ s_{21}\ s_{12}\ s_{22}]^T$ for a 2×2 symmetric $\mathbf{S}$. Using this operator, Eq. (5) can be written as

$$\text{vec}(\mathbf{\Sigma}_y) = [\pi(\mathbf{A})\ \text{vec}(\mathbf{I}_n)]\cdot\boldsymbol{\sigma}^2, \tag{9}$$

where $\pi(\cdot)$ is a matrix transform defined as

$$\pi(\mathbf{A}) = [(\mathbf{a}^1 * \mathbf{a}^1)^T\ \cdots\ (\mathbf{a}^1 * \mathbf{a}^n)^T\ |\cdots|\ (\mathbf{a}^n * \mathbf{a}^1)^T\ \cdots\ (\mathbf{a}^n * \mathbf{a}^n)^T]^T, \tag{10}$$

where $\mathbf{a}^j$ is the jth row vector of $\mathbf{A}$ and $*$ represents the Hadamard product [25]. Replacing $\mathbf{\Sigma}_y$ with $\mathbf{S}_y$ in Eq. (9), $\hat{\boldsymbol{\sigma}}^2$ is obtained as

$$\begin{aligned}\hat{\boldsymbol{\sigma}}^2 &= ([\pi(\mathbf{A})\ \text{vec}(\mathbf{I}_n)]^T[\pi(\mathbf{A})\ \text{vec}(\mathbf{I}_n)])^{-1}\\ &\quad\times[\pi(\mathbf{A})\ \text{vec}(\mathbf{I}_n)]^T\cdot\text{vec}(\mathbf{S}_y)\\ &= (\mathbf{\Pi}^T\mathbf{\Pi})^{-1}\mathbf{\Pi}^T\cdot\text{vec}(\mathbf{S}_y)\\ &= \mathbf{\Pi}^+\text{vec}(\mathbf{S}_y),\end{aligned} \tag{11}$$

where $\mathbf{\Pi} \equiv [\pi(\mathbf{A})\ \text{vec}(\mathbf{I}_n)]$ and $\mathbf{\Pi}^+ \equiv (\mathbf{\Pi}^T\mathbf{\Pi})^{-1}\mathbf{\Pi}^T$.

3.1.3 Estimator in Stoica and Nehorai [9]. For matrix $\mathbf{A}$ of full column rank, i.e., $\mathbf{A}^T\mathbf{A}$ is full rank, Stoica and Nehorai [9] defined the estimator

$$\hat{\mathbf{\Sigma}}_u = \mathbf{A}^+\mathbf{S}_y\mathbf{A}^{+T} - \hat{\sigma}_v^2(\mathbf{A}^T\mathbf{A})^{-1}$$

$$\hat{\sigma}_v^2 = \mathrm{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\mathbf{S}_y)/(n-p), \tag{12}$$

where $\mathbf{A}^+ \equiv (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$. When random variables in $\mathbf{u}$ are known to be independent, as assumed in assumption (3), the heuristic estimator of $\{\sigma_i^2\}_{i=1}^p$ uses the diagonal elements of $\hat{\mathbf{\Sigma}}_u$.

3.1.4 Estimator in Apley and Shi [10]. The variance estimator in Apley and Shi [10] was also developed for a system in which matrix $\mathbf{A}$ has full column rank. The estimator in Apley and Shi [10] is obtained as follows:

(i) Estimate $\hat{\mathbf{u}}(i) = \mathbf{A}^+\mathbf{y}(i)$ for $i = 1, \ldots, N$,
(ii) Estimate the variance of $\mathbf{v}$: $\hat{\sigma}_v^2 = (1/N(n-p))\Sigma_{i=1}^N \hat{\mathbf{v}}^T(i)\hat{\mathbf{v}}(i)$, where $\hat{\mathbf{v}}(i) = \mathbf{y}(i) - \mathbf{A}\hat{\mathbf{u}}(i)$;
(iii) Estimate the variance of $\mathbf{u}$: $\hat{\sigma}_j^2 = (1/N)\Sigma_{i=1}^N \hat{u}_j^2(i) - \hat{\sigma}_v^2(\mathbf{A}^T\mathbf{A})_{j,j}^{-1}$, for $j = 1, \ldots, p$, where $\hat{u}_j(i)$ is the jth element of $\hat{\mathbf{u}}(i)$ and $(\mathbf{A}^T\mathbf{A})_{j,j}^{-1}$ is the $(j,j)^{\text{th}}$ element of $(\mathbf{A}^T\mathbf{A})^{-1}$.

3.1.5 MLE. Given the assumption of normality, the probability density function (p.d.f.) of $\mathbf{y}(i)$ is defined as $f(\mathbf{y}(i)) = (2\pi)^{-(1/2)n}|\mathbf{\Sigma}_y|^{-1/2}e^{-(1/2)\mathbf{y}(i)^T\mathbf{\Sigma}_y^{-1}\mathbf{y}(i)}$. Denoting the complete set of observations as $\mathbf{y}^T = [\mathbf{y}(1)^T\ \mathbf{y}(2)^T \cdots \mathbf{y}(N)^T]^{Nn\times 1}$, the p.d.f. of $\mathbf{y}$ is $f(\mathbf{y}) = \Pi_{i=1}^N f(\mathbf{y}(i))$. The log-likelihood function of $\mathbf{\Sigma}_y$ is then

$$L(\mathbf{\Sigma}_y|\mathbf{y}) = \ln f(\mathbf{y}) = -\frac{Nn}{2}\ln(2\pi) - \frac{N}{2}\ln|\mathbf{\Sigma}_y| - \frac{N}{2}\mathrm{tr}(\mathbf{\Sigma}_y{}^{-1}\mathbf{S}_y). \tag{13}$$

Based on this log-likelihood function, Anderson [26,27] derived an MLE as a solution to the nonlinear equation

$$\{\mathrm{tr}(\hat{\mathbf{\Sigma}}_y^{-1}\mathbf{V}_i\hat{\mathbf{\Sigma}}_y^{-1}\mathbf{V}_j)\}_{i,j=1}^{p+1}\hat{\boldsymbol{\sigma}}^2 = \{\mathrm{tr}(\hat{\mathbf{\Sigma}}_y^{-1}\mathbf{V}_i\hat{\mathbf{\Sigma}}_y^{-1}\mathbf{S}_y)\}_{i=1}^{p+1}. \tag{14}$$

Anderson [26,27] also presented iterative numerical algorithms to solve Eq. (14).

3.2 Properties of On-Line Variance Estimators. Although the four on-line variance estimators reviewed in the previous section appear to significantly differ, some are intrinsically equivalent, as will be shown in the following. As a result, these four variance estimators can be grouped into two distinct categories.

First, it will be shown that the LS fit estimator and the estimator in Ding et al. [11] are identical, i.e., that $\mathbf{\Pi}^T\mathbf{\Pi} = \{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}$ and $\mathbf{\Pi}^T \cdot \mathrm{vec}(\mathbf{S}_y) = \{\mathrm{tr}(\mathbf{V}_i\mathbf{S}_y)\}_{i=1}^{p+1}$, as declared by the following Lemma (the proof is included in Appendix A).

Lemma 1. $\mathbf{\Pi}^T\mathbf{\Pi} = \{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}$ and $\mathbf{\Pi}^T \cdot \mathrm{vec}(\mathbf{S}_y) = \{\mathrm{tr}(\mathbf{V}_i\mathbf{S}_y)\}_{i=1}^{p+1}$.

This result is not surprising since both variance estimators are obtained from application of the least-squares criterion. The term "LS fit estimator" is used hereinafter to refer to both estimators.

Second, it will be shown that the estimator in Apley and Shi [10] comprises the diagonal elements of the estimator proposed by Stoica and Nehorai [9] in Eq. (12). The proof is as follows:

Substituting $\hat{\mathbf{v}}(i) = \mathbf{y}(i) - \mathbf{A}\hat{\mathbf{u}}(i)$ and $\hat{\mathbf{u}}(i) = \mathbf{A}^+\mathbf{y}(i)$ into $\hat{\sigma}_v^2$ in Apley and Shi [10] yields

$$\begin{aligned}\hat{\sigma}_v^2 &= \frac{1}{N(n-p)}\Sigma_{i=1}^N \mathbf{y}^T(i)(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)^T(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\mathbf{y}(i)\\ &= \frac{1}{N(n-p)}\mathrm{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\cdot\Sigma_{i=1}^N \mathbf{y}(i)\mathbf{y}^T(i))\\ &= \frac{1}{(n-p)}\cdot\mathrm{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\cdot\mathbf{S}_y),\end{aligned} \tag{15}$$

which is the same as $\hat{\sigma}_v^2$ in Eq. (12).

Further, $\hat{u}_j(i) = \mathbf{A}_j^+\mathbf{y}(i)$, where $\mathbf{A}_j^+$ is the jth *row* vector of $\mathbf{A}^+$. $\hat{\sigma}_j^3$ can then be expressed as

$$\begin{aligned}\hat{\sigma}_j^2 &= \frac{1}{N}\Sigma_{i=1}^N \hat{u}_j^2(i) - \hat{\sigma}_v^2(\mathbf{A}^T\mathbf{A})_{j,j}^{-1} = \frac{1}{N}\Sigma_{i=1}^N(\mathbf{A}_j^+\mathbf{y}(i))(\mathbf{A}_j^+\mathbf{y}(i))^T\\ &\quad - \hat{\sigma}_v^2(\mathbf{A}^T\mathbf{A})_{j,j}^{-1} = \mathbf{A}_j^+\cdot\left(\frac{1}{N}\Sigma_{i=1}^N\mathbf{y}(i)(\mathbf{y}(i))^T\right)\cdot(\mathbf{A}_j^+)^T - \hat{\sigma}_v^2(\mathbf{A}^T\mathbf{A})_{j,j}^{-1}\\ &= \mathbf{A}_j^+\cdot\mathbf{S}_y\cdot(\mathbf{A}_j^+)^T - \hat{\sigma}_v^2(\mathbf{A}^T\mathbf{A})_{j,j}^{-1}.\end{aligned} \tag{16}$$

It is clear that the above result is the (j,j)th diagonal element of $\hat{\mathbf{\Sigma}}_u$ in Eq. (12). Thus, the $\{\hat{\sigma}_j^2\}_{j=1}^p$ terms in Apley and Shi [10] constitute the diagonal elements of $\hat{\mathbf{\Sigma}}_u$ in Stoica and Nehorai [9]. This estimator is referred to hereinafter as a "diagonal-elements (DE) estimator."

A matrix expression of the DE estimator is developed as follows, and will be used subsequently to analyze estimation variance.

Let $\mathbf{e}_j$ be the jth column vector of $\mathbf{I}_p$. The first p elements in the DE estimator can then be expressed as

$$\hat{\boldsymbol{\sigma}}_{1,p}^2 \equiv [\hat{\sigma}_1^2 \ \cdots \ \hat{\sigma}_p^2]^T = \mathbf{Q}\cdot[(\mathbf{A}^+\otimes\mathbf{A}^+)\mathrm{vec}(\mathbf{S}_y) - \mathrm{vec}((\mathbf{A}^T\mathbf{A})^{-1})\hat{\sigma}_v^2], \tag{17}$$

where $\otimes$ is the matrix Kronecker product [25], $\hat{\sigma}_v^2$ is the same as in Eq. (12), and $\mathbf{Q}$ is defined as

$$\mathbf{Q} \equiv \begin{bmatrix}(\mathbf{e}_1\otimes\mathbf{e}_1)^T\\ \vdots\\ (\mathbf{e}_p\otimes\mathbf{e}_p)^T\end{bmatrix}. \tag{18}$$

From the above, it is evident that the four on-line variance estimators presented in Sec. 3.1 can be characterized into two distinct types: the LS fit estimator and the DE estimator. (The MLE is considered an offline estimator). Some key properties of the two types of on-line estimators, including existence condition, unbiasedness, and estimation uncertainty (i.e., dispersion level) are evaluated in the following.

3.2.1 Existence condition of variance estimators. The condition for DE estimator existence is that $\mathbf{A}^T\mathbf{A}$ is full rank and $n \geqslant p+1$, where $p+1$ is the number of independent variance components, including sensor noise variance. The existence condition for an LS fit estimator is that $\mathbf{\Pi}^T\mathbf{\Pi}$ or $\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}$ is full rank.

The existence condition of a variance estimator is related to the so-called "diagnosability condition" of variance sources, as presented in Zhou et al. [28]. In general, the diagnosability condition characterizes whether or not the observations of $\mathbf{y}$ contain enough information to assure that the variance components can be estimated. This condition is independent of specific estimation algorithms but should be required as a necessary condition for all variance estimators. Individual variance estimators may, however, require stronger diagnosability conditions. Zhou et al. [28] defined the diagnosability condition of variance components for a linear replicated model (1) as $\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}$ being full rank, the same as that of an LS fit estimator. The condition that $\mathbf{A}^T\mathbf{A}$ be full rank and $n \geqslant p+1$ is a stronger condition than $\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}$ or $\mathbf{\Pi}^T\mathbf{\Pi}$ be full rank. It can be proven that the existence of a DE estimator guarantees the existence of an LS fit estimator. The result is stated in Lemma 2 and the proof is included in Appendix B. However, the converse is not true; i.e., $\mathbf{\Pi}^T\mathbf{\Pi}$ could be full rank even if $\mathbf{A}^T\mathbf{A}$ is singular. An example of this occurrence in a multistation assembly process is presented in Sec. 5.

Lemma 2. If $\mathbf{A}^T\mathbf{A}$ is full rank and $n \geqslant p+1$, $\mathbf{\Pi}^T\mathbf{\Pi}$ is full rank.

An MLE exists if $\mathbf{S}_y$ is positive definite and $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_{p+1}$ are linearly independent [5,26]. The positive definiteness of sample variance $\mathbf{S}_y$ is usually satisfied in practice since independent sensor noise exists having nonzero variances. It is easy to verify that $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_{p+1}$ are linearly independent if and only if $\mathbf{\Pi}^T\mathbf{\Pi}$ is full rank. Therefore, the existence condition for an MLE is the same as the diagnosability condition for a linear replicated

model (1) and likewise an LS fit estimator.

3.2.2 Unbiasedness of variance estimators. When $\mathbf{u}$ and $\mathbf{v}$ are assumed to be zero-mean vectors, $E(\mathbf{S}_y)=\boldsymbol{\Sigma}_y$, i.e., $\mathbf{S}_y$ is the unbiased estimate of $\boldsymbol{\Sigma}_y$, where $E(\cdot)$ is the expectation operator. From Eq. (11), it can be readily seen that the LS fit estimator is unbiased.

The unbiasedness of a DE estimator can be more easily determined from Eq. (12) than from Eq. (17). Note that $E(\hat{\sigma}_v^2) = \mathrm{tr}((\mathbf{I}_n-\mathbf{A}\mathbf{A}^+)\boldsymbol{\Sigma}_y)/(n-p))$. Applying Eq. (5) yields

$$\begin{aligned} E(\hat{\sigma}_v^2) &= \mathrm{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)(\mathbf{A}\boldsymbol{\Sigma}_u\mathbf{A}^T + \sigma_v^2\mathbf{I}_n))/(n-p) \\ &= \sigma_v^2 \cdot \{\mathrm{tr}(\mathbf{I}_n) - \mathrm{tr}(\mathbf{A}\mathbf{A}^+)\}/(n-p) \\ &= \sigma_v^2 \cdot \{\mathrm{tr}(\mathbf{I}_n) - \mathrm{tr}((\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A})\}/(n-p) \\ &= \sigma_v^2 \cdot \{\mathrm{tr}(\mathbf{I}_n) - \mathrm{tr}(\mathbf{I}_p)\}/(n-p) \\ &= \sigma_v^2. \end{aligned} \tag{19}$$

The expectation of $\hat{\boldsymbol{\Sigma}}_u$ is then taken as

$$E(\hat{\boldsymbol{\Sigma}}_u) = \mathbf{A}^+\boldsymbol{\Sigma}_y\mathbf{A}^{+^T} - E(\hat{\sigma}_v^2)\cdot(\mathbf{A}^T\mathbf{A})^{-1}. \tag{20}$$

Utilizing the results from Eqs. (5) and (19), it is easy to show that $E(\hat{\boldsymbol{\Sigma}}_u)=\boldsymbol{\Sigma}_u$. Thus, the DE estimator is unbiased.

Although an MLE is generally biased, it is asymptotically unbiased. On-line estimators have a distinct advantage over an MLE with respect to unbiasedness when the sample size is not large. Nevertheless, the MLE bias is not pertinent here since only on-line estimators are being compared. MLE is utilized simply to provide a performance reference for comparing estimation uncertainty, i.e., the dispersion level, as discussed in the following.

3.2.3 Dispersion of variance estimators. Variance estimator dispersion is characterized herein by the trace of its variance-covariance matrix, i.e., $\mathrm{var}_{ML,DE,\mathrm{or}LS}\equiv\mathrm{tr}(Cov(\hat{\boldsymbol{\sigma}}^2_{ML,DE,\mathrm{or}LS}))$. The same criterion has been used for a composite comparison of variance estimation in Corbeil and Searle [13]. Since the MLE will be used as the reference for comparison of DE and LS fit estimators, MLE dispersion is also presented.

First, the variance of $\hat{\sigma}_v^2$ for a DE estimator is derived as

$$Var(\hat{\sigma}_v^2) = \frac{2\sigma_v^4}{N(n-p)}, \tag{21}$$

where $Var(\cdot)$ is the variance of a random variable. The complete derivation is given in Appendix C. Using Eq. (17), var_{DE} can be calculated as

$$\begin{aligned} \mathrm{var}_{DE} = \mathrm{tr}(Cov(\hat{\boldsymbol{\sigma}}^2_{1,p})) + Var(\hat{\sigma}_v^2) &= \frac{1}{N}\mathrm{tr}(\mathbf{P}_1(\mathbf{I}_{n^2}+\mathbf{K}_n)(\boldsymbol{\Sigma}_y\otimes\boldsymbol{\Sigma}_y)\mathbf{P}_1^T) \\ &- \frac{1}{N(n-p)}\mathrm{tr}(\mathbf{P}_1(\mathbf{I}_{n^2}+\mathbf{K}_n)(\boldsymbol{\Sigma}_y\otimes\boldsymbol{\Sigma}_y)\mathbf{P}_3^T) \\ &+ \mathrm{tr}(\mathbf{P}_2\mathbf{P}_2^T)\frac{2\sigma_v^4}{N(n-p)} + \frac{2\sigma_v^4}{N(n-p)}, \end{aligned} \tag{22}$$

where $\mathbf{P}_1=\mathbf{Q}\cdot(\mathbf{A}^+\otimes\mathbf{A}^+)$, $\mathbf{P}_2=\mathbf{Q}\cdot\mathrm{vec}((\mathbf{A}^T\mathbf{A})^{-1})$, $\mathbf{P}_3=\mathbf{P}_2\cdot(\mathrm{vec}(\mathbf{I}_n-\mathbf{A}\mathbf{A}^+))^T$, and $\mathbf{K}_n$ is as defined in Magnus and Neudecker [29]. The complete derivation of Eq. (22) is given in Appendix D.

Similarly, the var_{LS} for the LS fit estimator in Eq. (11) is

$$\mathrm{var}_{LS} = \mathrm{tr}(Cov(\hat{\boldsymbol{\sigma}}^2_{LS})) = \frac{1}{N}tr(\boldsymbol{\Pi}^+(\mathbf{I}_{n^2}+\mathbf{K}_n)(\boldsymbol{\Sigma}_y\otimes\boldsymbol{\Sigma}_y)(\boldsymbol{\Pi}^+)^T). \tag{23}$$

The covariance matrix of an MLE is approximated by the inverse of its Fisher information matrix [6], given as

$$\boldsymbol{\Psi}(\hat{\boldsymbol{\sigma}}^2_{ML}) \equiv -E\left(\left\{\frac{\partial^2 L(\boldsymbol{\Sigma}_y|\mathbf{y})}{\partial\sigma_i^2\partial\sigma_j^2}\right\}_{i,j=1}^{p+1}\right). \tag{24}$$

Thus, var_{ML} is approximated as

$$\mathrm{var}_{ML} \approx \mathrm{tr}(\boldsymbol{\Psi}^{-1}) = \frac{2}{N}\mathrm{tr}\{(\{\mathrm{tr}(\boldsymbol{\Sigma}_y^{-1}\mathbf{V}_i\boldsymbol{\Sigma}_y^{-1}\mathbf{V}_j)\}_{i,j=1}^{p+1})^{-1}\}. \tag{25}$$

It should be noted that the sample size N has the same effect on the dispersion of all three estimators for Eqs. (22), (23), and (25).

4 Performance Comparison and Selection Guidelines for On-Line Estimators

The performance of the two types of on-line variance estimators is compared in this section. Since DE and LS fit estimators are both unbiased, estimator dispersion is employed as the criterion for performance comparison. One objective of this comparison is to determine the condition under which the LS fit and DE estimators may be effective alternatives to MLE for on-line variance estimation. LS fit and DE estimators can be computed using their closed-form expressions and, consequently, should require less computation time than the MLE. The primary disadvantage of the LS fit or DE estimator is that either may demonstrate unacceptably higher variances than MLE. var_{ML} is used as the reference for this performance comparison. The relative difference between a DE (or LS) estimator and an MLE is characterized by the percentage difference (*Diff*), defined as

$$Diff_{DE(\mathrm{or}LS)\mathrm{vs}ML} \equiv \frac{\mathrm{var}_{DE(LS)} - \mathrm{var}_{ML}}{\mathrm{var}_{ML}} \times 100\,\%\,. \tag{26}$$

A direct analytical comparison of variance estimators is difficult, if not impossible. To address this issue, a general understanding of the performance of DE and LS fit estimators is provided, followed by a numerical evaluation to illustrate the conclusions of this study.

An LS fit or DE estimator becomes an MLE estimator under special conditions. A DE estimator is an MLE in a noise free environment; i.e., when $\sigma_v^2=0$. In this case, when a DE estimator exists, observation of $\mathbf{y}(i)$ is equivalent to direct observation of $\mathbf{u}(i)$. The sample variances computed from direct observation of $\mathbf{u}(i)$, $i=1,\ldots,N$, are the maximum likelihood estimators of $\{\sigma_j^2\}_{j=1}^p$. Under noise-free conditions, the diagonal elements of $\hat{\boldsymbol{\Sigma}}_u$ (defined in Eq. (12)) are the same as the sample variances of $\mathbf{u}$. Therefore, a DE estimator is the MLE of $\{\sigma_j^2\}_{j=1}^p$. In contrast, an LS fit estimator becomes an MLE when the signal $\mathbf{u}$ is not random. Randomness in $\mathbf{y}$ is due solely to sensor noise; i.e., $\boldsymbol{\Sigma}_u=\mathbf{0}$ but $\sigma_v^2\neq 0$. This equivalence can be easily demonstrated by substituting $\hat{\boldsymbol{\Sigma}}_y=\hat{\sigma}_v^2\mathbf{I}_n$ into Eq. (14) (ML equation). The results for $\hat{\sigma}_v^2$ will be the same as those obtained from Eq. (8) (LS fit equation).

Hence, the LS fit and DE estimators are two extreme cases of the MLE. It is expected that a DE estimator will perform as well as an MLE when sensor noise is relatively small, while an LS fit estimator will perform as well as an MLE when sensor noise is dominant in the process. In order to characterize sensor noise dominance, an average signal-to-noise ratio (SNR) is defined as follows:

$$\mathrm{SNR} = \frac{\mathrm{tr}(\boldsymbol{\Sigma}_u)}{p\cdot\sigma_v^2}, \tag{27}$$

where $\mathrm{tr}(\boldsymbol{\Sigma}_u)/p=\Sigma_{i=1}^p\sigma_i^2/p$ is the average signal power and σ_v^2 is the sensor noise power.

In addition to the SNR, the structure of matrix $\mathbf{A}$ will also affect the estimator variance. However, since the existence of matrix $\mathbf{A}$ in dispersion indicators (22), (23), and (25) is complicated, it is recommended that the percentage difference (*Diff*) versus SNR be plotted for a given matrix $\mathbf{A}$ using Eqs. (22), (23), and (25). As an example, let $n=6$, $p=3$, and matrix $\mathbf{A}$ be

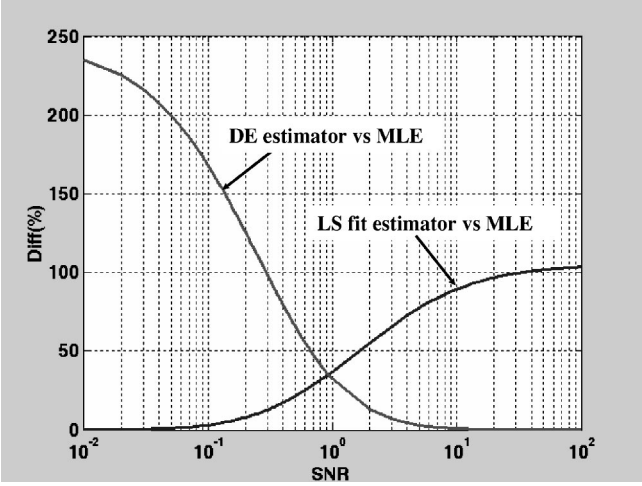

Fig. 4 OC curves for DE and LS fit estimators

$$\mathbf{A}^T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{28}$$

for which the SNR range is selected to be 0.01, ~100. This SNR range is achieved by fixing σ_v^2 at unit variance, while varying the value of $\text{tr}(\boldsymbol{\Sigma}_u)$. While the components σ_i^2 in $\boldsymbol{\Sigma}_u$ can take different values for a given $\text{tr}(\boldsymbol{\Sigma}_u)$, for this case all σ_i^2 are assigned equal values; i.e., $\sigma_i^2 = \text{tr}(\boldsymbol{\Sigma}_u)/p$. For instance, if SNR=1, given $\sigma_v^2 = 1$, then $\text{tr}(\boldsymbol{\Sigma}_u) = 1$. Subsequently, $\sigma_i^2 = 0.333$ for $i = 1, \ldots, 3$.

Diff versus SNR for both DE versus ML and LS versus ML are plotted in Fig. 4. Figure 4 provides a quantitative characterization of the relationship between the two estimators, and is consistent with the general concept described in the preceding. The graphs in Fig. 4 are referred to as the OC curves for DE and LS fit estimators. In practice, the curves are plotted for a given matrix **A** to determine the SNR range over which a DE or an LS fit estimator demonstrates acceptable performance compared with an MLE. With respect to matrix **A** in Eq. (28), the point at which the DE and LS fit estimators exhibit the same performance is approximately SNR=1 (SNR=10^0). If 10% is selected as the maximum allowed difference from an MLE, then a DE estimator is a good alternative for an MLE when SNR>2. Similarly, an LS fit estimator is an effective alternative when SNR<0.2.

The characteristic relationship between the two types of estimators illustrated in Fig. 4 is by and large true for most system designs exhibiting the general structure of matrix **A**. For some special structures of matrix **A**, however, there are cases where a DE and/or LS fit estimator could be exactly the same as an MLE, regardless of the SNR value. The general conditions concerning the effects of matrix **A** merit further study since the results can lead to more effective design of the system structure of matrix **A**, and subsequently lead to more efficient and accurate estimation of variance components.

To this point, discussion and methodology development has progressed under the assumption that **u** is of a zero mean. When **u** has a nonzero unknown mean, model (1) can be expressed as

$$\mathbf{y}(i) = \mathbf{A}\boldsymbol{\mu}_u + \mathbf{A}\tilde{\mathbf{u}}(i) + \mathbf{v}(i), \quad i = 1,2, \ldots, N, \tag{29}$$

where $\boldsymbol{\mu}_u$ is an unknown mean and $\tilde{\mathbf{u}}$ is the zero-mean random vector. The variance components to be estimated are associated with $\tilde{\mathbf{u}}$. With respect to DE and LS fit estimators, Eqs. (8), (11), and (17) can still be employed by replacing $\mathbf{S}_y$ with the sample covariance matrix for samples having a nonzero mean; e.g.,

$$\mathbf{S}_y = \frac{1}{N-1}\sum_{i=1}^{N} (\mathbf{y}(i) - \bar{\mathbf{y}})(\mathbf{y}(i) - \bar{\mathbf{y}})^T, \tag{30}$$

where $\bar{\mathbf{y}}$ is the sample mean of **y**. The resulting estimators are unbiased, and their existence conditions do not change. Again, the general behavior presented in the preceding performance comparison is still valid—the same procedures can be used to develop similar OC curves. A summary comparison of the two types of estimators is given in Table 1.

5 Case Study

The automotive body assembly process described in Sec. 1 is used in the following for case study analysis. Two cases are presented: (1) an assembly process involving only one station in which both assembly and inspection operations are performed; and (2) a two-station assembly process in which the assembly operation is performed at the first station and the inspection performed at the second station.

5.1 Single-Station Assembly Process. A single-station automotive body assembly process is described and modeled in Apley and Shi [10] Sec. 4. For this particular problem, there are nine measurements ($n=9$) and three independent variation sources ($p=3$). The associated matrix **A** given by Apley and Shi [10] is

$$\mathbf{A}^T = \begin{bmatrix} .093 & 0 & -.093 & .093 & 0 & .647 & -.370 & 0 & .647 \\ .577 & 0 & 0 & .577 & 0 & 0 & .577 & 0 & 0 \\ -.120 & 0 & .843 & -.120 & 0 & -.120 & .482 & 0 & -.120 \end{bmatrix}. \tag{31}$$

Matrix **A** in this example is of full column rank and $n > p+1$, suggesting that both DE and LS fit estimators are applicable. The OC curve for this system design is presented in Fig. 5.

In order to evaluate the performance of variance estimators, the SNR is first estimated from engineering design specifications. The sensor used in this example is a type of noncontact coordinate sensor having a regular precision. The specified sensor accuracy is $(6\sigma)_{sensor} = 0.1$ mm. In contrast, the tolerance of the pinhole contact is roughly 0.2 mm. If the tolerance is approximated by the six-sigma value, then $(6\sigma)_{locator} = 0.2$ mm, implying that the repeatability of a malfunctioning locator will have a six-sigma value

Table 1 Summarized comparison of on-line variance estimators

	DE estimator	LS fit estimator
Definition	Equations (12) and (17)	Equations (8) and (11)
Pros	• Closed-form solution, computationally efficient. • Unbiased. • Close to MLE when SNR large	• Closed-form solution, computationally efficient. • Unbiased. • Same as MLE if SNR=0.
Cons	• Performance deteriorates in noisy environment (small SNR).	• Performs much worse than a DE estimator and an MLE for a large SNR.
Existence condition	• $\mathbf{A}^T\mathbf{A}$ full rank and $n \geq p+1$. Stronger condition than that $\boldsymbol{\Pi}^T\boldsymbol{\Pi}$ full rank.	• $\boldsymbol{\Pi}^T\boldsymbol{\Pi}$ full rank. Estimator may exist when $\mathbf{A}^T\mathbf{A}$ is singular.

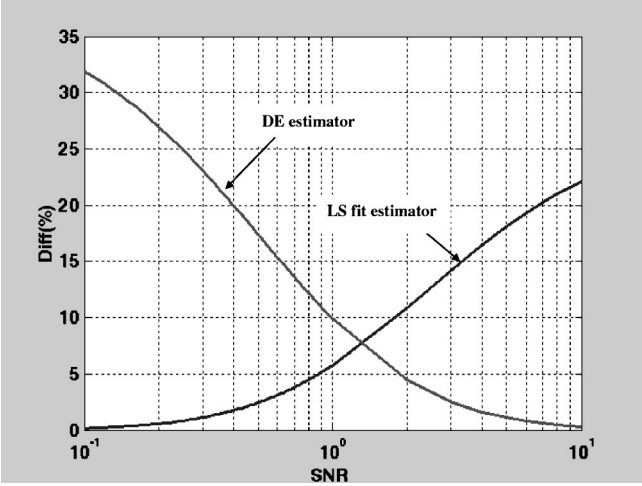

Fig. 5 OC curves for the linear system with A as in Eq. (31)

larger than 0.2 mm. Based on this approximation, SNR $\geqslant((6\sigma)_{locator}/(6\sigma)_{sensor})^2=4$. Given an SNR of around 4, a DE estimator can exhibit a 1% difference from an MLE in terms of estimation variance, whereas the difference for an LS fit estimator would be around 18%.

To determine if meaningful guidance can be obtained from the OC curve, a simulation (500 trials) was performed with variance components $\boldsymbol{\sigma}^2=[0.0011\ 0.0025\ 0.0044\ 0.0006]$ (including sensor noise) and a sample size of $N=25$. The results from the two on-line variance estimators are compared in Table 2. MLE results are also included in Table 2 as a reference. The first row in Table 2 ($\overline{\hat{\boldsymbol{\sigma}}^2}$) presents the sample average of variance estimation from all three estimators and indicates that both the DE and LS fit estimators are unbiased. The bias of the MLE in this example is not noticeable. Table 2, Row 2 details the sample variance of each estimator. The computed *Diff* values (Row 3) are near those predicted by the OC curve. The simulation result confirms our conclusion from the OC curve. The small difference between simulation results and the OC curve is due to the fact that the MLE variance in Eq. (25) is actually a large sample approximation. In real circumstances, the sample size may not satisfy the large sample requirement.

5.2 Multistation Assembly. Figure 6 depicts a two-station process, derived as a segment of the simplified automotive body assembly process offered in Ding et al. [11]. In this example, three workpieces are welded together at Station I, with the first workpiece, a subassembly from prior assembly operations, consisting of two components. Once welding operations are completed, the entire assembly is transferred to a dedicated in-process OCMM station (Station II) for inspection. The nine points at which measurements are taken are indicated in Fig. 6(*b*). High-precision laser-optic coordinate sensors are used to measure two directional coordinates at each measurement point; therefore, $n=18$. Since the dedicated OCMM station is assumed to be always well maintained, only those variation sources associated with locators on

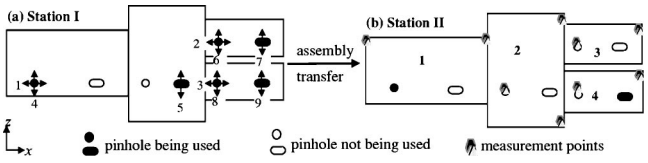

Fig. 6 A two-station assembly process

Station I are considered. Fixturing variation sources are marked from 1 to 9 in Fig. 6(*a*); 1 through 3 represent fixturing variations in the *x* direction and 4 through 9 are variations in the *z* direction. Hence, $p_x=3$, $p_z=6$, and $p=p_x+p_z=9$.

Since this process involves more than one station, the state-space modeling approach outlined in Sec. 2 is employed to generate the variation propagation model. A detailed modeling procedure for this example can be found in Jin and Shi [18] and specific model parameters are given in Ding et al. [11]. The state-space model was converted to a linear replicated model as

$$\mathbf{y}(i)=\mathbf{A}\mathbf{u}(i)+\mathbf{v}(i)=[\mathbf{A}_x\quad\mathbf{A}_z]\mathbf{u}(i)+\mathbf{v}(i),\quad i=1,2,\ldots N, \tag{32}$$

where $\mathbf{A}_x$ and $\mathbf{A}_z$ are two blocks in matrix $\mathbf{A}$, corresponding to horizontal and vertical fixturing variations, respectively. The numerical expression of matrix $\mathbf{A}$ is included in Appendix E. For the purposes of this case study, it was assumed that only horizontal variance sources or vertical sources exist, so that only $\mathbf{A}_x$ or $\mathbf{A}_z$ is needed for variance estimation. This simplified treatment is employed only to demonstrate different variance estimator performances. Similarly, both horizontal and vertical variance sources can be simultaneously estimated using the entire $\mathbf{A}$ matrix. First $\mathbf{A}_x\in \mathrm{Re}^{18\times 3}$ is evaluated. $\mathbf{A}_x^T\mathbf{A}_x$ is

$$\mathbf{A}_x^T\mathbf{A}_x=\begin{bmatrix}4&-2&-2\\-2&2&0\\-2&0&2\end{bmatrix}, \tag{33}$$

which has a rank of 2. In this case, $\mathbf{A}_x^T\mathbf{A}_x$ is singular, illustrating a critical aspect of many multistation processes, i.e., matrix $\mathbf{A}$ often does not have full column rank.

In this singular system, the DE estimator cannot be used. The existence of an LS fit estimator is dependent on the rank of $\mathbf{\Pi}_x^T\mathbf{\Pi}_x$, where $\mathbf{\Pi}_x=[\pi(\mathbf{A}_x)\ \mathrm{vec}(\mathbf{I}_n)]$. In this case,

$$\mathbf{\Pi}_x^T\mathbf{\Pi}_x=\begin{bmatrix}16&4&4&4\\4&4&0&2\\4&0&4&2\\4&2&2&18\end{bmatrix}, \tag{34}$$

Table 2 Comparison of three estimators for linear system with A as in Equation (31)

	DE	LS fit	MLE
$\overline{\hat{\boldsymbol{\sigma}}^2}$	[0.0011 0.0025 0.0044 0.0006]	[0.0011 0.0025 0.0044 0.0006]	[0.0011 0.0025 0.0044 0.0006]
$\mathrm{tr}(\mathrm{Cov}(\hat{\boldsymbol{\sigma}}^2))$	3.20×10^{-6}	3.58×10^{-6}	3.17×10^{-6}
Diff (%)	1.1%	13.1%	—

Note: The true values of $\boldsymbol{\sigma}^2=[0.0011\ 0.0025\ 0.0044\ 0.0006]$.

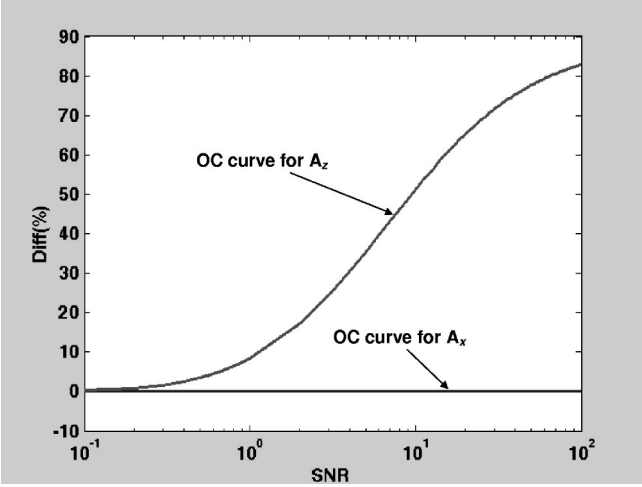

Fig. 7 OC curves for A_x and A_z

which is of full rank. Hence, an LS fit estimator exists. The OC curve for the LS fit estimator is given in Fig. 7, a roughly flat line near the $Diff=0$ value. As a result, to estimate variance location in the x direction, an LS fit estimator can be used in place of an MLE.

The same evaluation is performed for $\mathbf{A}_z \in \mathrm{Re}^{18\times 6}$. Likewise, $\mathbf{A}_z^T\mathbf{A}_z$ is also singular (rank=5), however, $\mathbf{\Pi}_z^T\mathbf{\Pi}_z$ (defined similarly as $\mathbf{\Pi}_x^T\mathbf{\Pi}_x$) is full rank. For convenience $\mathbf{A}_z^T\mathbf{A}_z$ and $\mathbf{\Pi}_z^T\mathbf{\Pi}_z$ are not computed here, although the OC curve for $\mathbf{A}_z$ is also given in Fig. 7. Unfortunately, when the SNR is large, the LS fit estimator for a system having $\mathbf{A}_z$ can exhibit an unacceptable deterioration from the preferred MLE performance. Hence, the LS fit estimator may not be appropriate for use in estimating variance location in the z direction, depending on the amplitude of the process SNR.

In this multistation example, a dedicated OCMM station houses highly accurate coordinate sensors with $(6\sigma)_{sensor}=0.02$ mm. The tolerance for fixture locators is the same as before, resulting in an SNR=100. Thus, the variance of an LS fit estimator can be as much as 80% greater than that of an MLE. Again, 500 simulation trials were run with $N=25$. The variance estimation sample average (Table 3, Row 1) indicates that the LS fit estimators are unbiased and the bias of the MLE is not noticeable. The *Diff* values are also near those obtained from the OC curve. Hence, the behavior of the OC curve in Fig. 7 is verified by the simulation results given in Table 3.

5.3 Computation Time. As mentioned earlier, computation of an MLE typically involves an iterative procedure and can be expensive for a large-scale system and a moderately large sample size. The required computation time was recorded for the three estimators in each of the above case examples for sample sizes of $N=10$, 25, 50, 75, and 100. Estimators were computed using MATLAB, consequently a relatively longer time was required than similar computation using C or FORTRAN programming. However, the objective here is to determine the relative performance between estimators. The time-recording accuracy in MATLAB is 0.01 s; therefore, computation time is reported as "<0.01 s" if the returned algorithm execution time is zero. All algorithms are executed using the same computer platform (Pentium™ III CPU operating at 1.0 GHz).

The single-station case example entailed the simplest computational burden. In this case, the dimension of matrix $\mathbf{A}$ is 9×3, representing a single rigid part with three coordinate sensors. The required computation time for all LS fit and DE estimators was

Table 3 Simulation comparison of LS fit and ML estimators

	LS fit ($\mathbf{A}_x$)	MLE ($\mathbf{A}_x$)	LS fit ($\mathbf{A}_z$)	MLE ($\mathbf{A}_z$)
$\overline{\hat{\boldsymbol{\sigma}}^2}$	[0.0045 0.0003 0.0003 1.11×10^{-5}]	[0.0045 0.0003 0.0003 1.11×10^{-5}]	[0.0005 0.0005 0.0014 0.0025 0.0005 0.0005 1.15×10^{-5}]	[0.0005 0.0005 0.0014 0.0025 0.0005 0.0005 1.11×10^{-5}]
tr(Cov($\hat{\boldsymbol{\sigma}}^2$))	1.741×10^{-6}	1.675×10^{-6}	1.678×10^{-6}	0.954×10^{-6}
Diff (%)	3.97%	—	75.93%	—
True values	[.0045 .0003 .0003, 1.11×10^{-5}]		[.0005 .0005 .0015 .0025 .0005 .0005, 1.11×10^{-5}]	

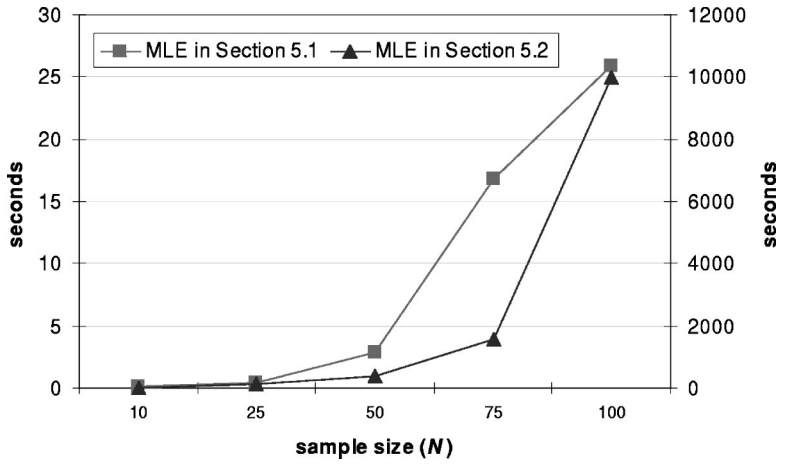

Fig. 8 Computation time for MLE (s)

<0.01 s. The computation time required for the MLE is presented in Fig. 8, reading from the left vertical axis. For this simple single part case, the MLE computation time is minimal: roughly 26 s for a sample size of $N=100$. The MLE computation requirement is, however, much larger than that of either type of on-line estimator, and exhibits a nonlinear (exponential) increase as the sample size increases.

The second case study represents a moderately complex process with multiple stations. The dimension of matrix $\mathbf{A}$ is 18×9. The values in matrix $\mathbf{A}$ were modified slightly to obtain a matrix of full rank to test DE estimator computation for this example. The computation times for both DE and LS fit estimators remained <0.01 s for all sample sizes. The required MLE computation time for this example is also given in Fig. 8, reading from the right vertical axis. It is observed that for this two-station process, MLE computation time for $N=75$ is approximately 26 min (1570 s) and for $N=100$ MLE computation will require 2 h, 50 min (9967 s), far too long a time for effective use in on-line variance estimation. As before, MLE computation time exponentially increases as sample size increases.

6 Concluding Remarks

A variance estimator for linear replicated models is of central importance in identifying sources of variation. Variance estimators with closed-form expressions are demonstrably more appropriate for on-line quality control than an MLE, which require multiple numerical iterations to achieve convergence. This paper has exposed the intrinsic relationship among several variance estimators developed for signal processing and quality engineering applications.

The results of study will significantly facilitate efforts supporting root cause analysis by enabling rapid and accurate estimation of underlying variation sources and providing information for corrective action. Shalon et al. [30] and Ceglarek and Shi [31] reported that dimensional problems contribute to roughly two-thirds of quality-related problems during new product launch in automotive and aerospace assembly applications. Timely identification and elimination of variation sources will greatly enhance product quality and significantly reduce launch time. Meanwhile, evaluation of variance estimator properties and comparison of performances will guide the proper use of estimators for different conditions. For example, for the multistation example presented in this paper, fault estimation will fail if a DE estimator is used arbitrarily. In addition, as illustrated in Fig. 4, for a case where $\mathrm{SNR}<0.1$, an LS fit estimator exhibits a variance more than three times that of a DE estimator. In other words, a sample size at least three times larger is needed to achieve the same estimation accuracy if an LS fit estimator is used. If an MLE is used, an even longer time may be required to reach a conclusion and the cost of the delay could be prohibitive.

The two-station example used in this study is still relatively simple compared to most actual manufacturing systems. For a full-scale manufacturing system, the necessity of using DE and LS fit estimators for on-line quality control is obvious. For instance, in the automotive industry in-process OCMMs measure each car body produced. Given a body shop with a daily throughput of 1000 units (two shifts), a car body assembly passes through an OCMM station every 2 min. Allocating one minute for actual measurement, roughly one minute is available to the OCMM to make a decision with respect to body assembly acceptance/process fault identification. In addition, as many as 150 measurements are taken for a full-sized car body at one OCMM station and as many as ten OCMM stations are distributed along an autobody shop assembly line [1]. The aggregated measurement vector $\mathbf{y}^T=[\mathbf{y}_1^T\cdots\mathbf{y}_M^T]$ (corresponding to Eq. (4)) for a multistation assembly system could have several hundred variables. Given the computation time analysis in Sec. 5.3 it is clear that the computationally expensive MLE is not well suited for high-dimension problems subject to high productivity requirements.

Instances of high data dimension and high throughputs are prevalent in other industries. For example, in a printed circuit board assembly process, throughput can be as high as 4000 units per day, which translates to a window of 20 s or less for both measurement and decision-making action. Polyester film processing [32] involves a total of 308 original variables, although this number can be reduced to 103 variables depending on the empirical knowledge of plant engineers. Given in-process sensors capable of measuring these variables at near real-time, the time pressure for rapid decision-making increases substantially in order to fully utilize new sensor capability. Since a linear model exhibits the same structure as Eq. (1), and is often used to describe the input-output relationship for both printed circuit board manufacturing and polyester film processing [32,33], the results from this study are likely to be applicable for quality control in systems other than automotive assembly.

Acknowledgment

The authors gratefully acknowledge financial support from the NSF Engineering Research Center for Reconfigurable Machining Systems under grant EEC95-92125, from the NSF under grants DMI-0217481, DMI-0322147, DMI-0348150, and from the State of Texas Advanced Technology Program under grant 000512-0237-2003.

Appendix A: Proof of Lemma 1 in Sec. 3.2

Proof. Using the definition of $\mathbf{\Pi}$ in Eq. (11), then

$$\mathbf{\Pi}^T\mathbf{\Pi}=\begin{bmatrix}\pi(\mathbf{A})^T\pi(\mathbf{A}) & \pi(\mathbf{A})^T\mathrm{vec}(\mathbf{I}_n)\\ \mathrm{vec}(\mathbf{I}_n)^T\pi(\mathbf{A}) & \mathrm{vec}(\mathbf{I}_n)^T\mathrm{vec}(\mathbf{I}_n)\end{bmatrix}=\begin{bmatrix}\pi(\mathbf{A})^T\pi(\mathbf{A}) & \pi(\mathbf{A})^T\mathrm{vec}(\mathbf{I}_n)\\ \mathrm{vec}(\mathbf{I}_n)^T\pi(\mathbf{A}) & n\end{bmatrix}. \tag{A1}$$

However, since $\mathbf{V}_{p+1}=\mathbf{I}_n$, $\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}$ can be expressed as

$$\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p+1}=\begin{bmatrix}\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p} & \{\mathrm{tr}(\mathbf{V}_i)\}_{i=1}^{p}\\ (\{\mathrm{tr}(\mathbf{V}_i)\}_{i=1}^{p})^T & n\end{bmatrix}. \tag{A2}$$

Further, it can be shown that $\pi(\mathbf{A})^T\pi(\mathbf{A})=\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p}$ and $\pi(\mathbf{A})^T\mathrm{vec}(\mathbf{I}_n)=\{\mathrm{tr}(\mathbf{V}_i)\}_{i=1}^{p}$. Recalling that $\mathrm{tr}(\mathbf{AB})=\mathrm{vec}(\mathbf{A})^T\mathrm{vec}(\mathbf{B})$ for any symmetric matrices $\mathbf{A}$ and $\mathbf{B}$, the (i,j)th element in $\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p}$ is $(\mathrm{vec}(\mathbf{a}_i\mathbf{a}_i^T))^T\mathrm{vec}(\mathbf{a}_j\mathbf{a}_j^T)$. In fact, $\mathrm{vec}(\mathbf{a}_i\mathbf{a}_i^T)$ is the ith column vector in $\pi(\mathbf{A})$, leading to the conclusion that $\pi(\mathbf{A})^T\pi(\mathbf{A})=\{\mathrm{tr}(\mathbf{V}_i\mathbf{V}_j)\}_{i,j=1}^{p}$.

Following the same procedure, $\{\mathrm{tr}(\mathbf{V}_i\mathbf{S}_y)\}_{i=1}^{p}=\{(\mathrm{vec}(\mathbf{a}_i\mathbf{a}_i^T))^T\cdot\mathrm{vec}(\mathbf{S}_y)\}_{i=1}^{p}=\pi(\mathbf{A})^T\mathrm{vec}(\mathbf{S}_y)$. Then, $\pi(\mathbf{A})^T\mathrm{vec}(\mathbf{I}_n)=\{\mathrm{tr}(\mathbf{V}_i)\}_{i=1}^{p}$ constitutes a special case when $\mathbf{S}_y=\mathbf{I}_n$. Moreover, $\mathbf{\Pi}^T\cdot\mathrm{vec}(\mathbf{S}_y)=\{\mathrm{tr}(\mathbf{V}_i\mathbf{S}_y)\}_{i=1}^{p+1}$ immediately results following a similar matrix partition as in Eqs. (A1) and (A2). □

Appendix B: Proof of Lemma 2 in Sec. 3.2

Proof. According to the definition of $\mathbf{\Pi}$ in Eq. (11), $\mathbf{\Pi}$ can be partitioned as

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{\Pi}_1 & \mathbf{e}_1 \\ \vdots & \vdots \\ \mathbf{\Pi}_i & \mathbf{e}_i \\ \vdots & \vdots \\ \mathbf{\Pi}_n & \mathbf{e}_n \end{bmatrix}, \tag{A3}$$

where $\mathbf{e}_i$ is the ith column vector of an identity matrix of size n and $\mathbf{\Pi}_i$ is the ith row block defined as

$$\mathbf{\Pi}_i = \begin{bmatrix} \mathbf{a}^i * \mathbf{a}^1 \\ \vdots \\ \mathbf{a}^i * \mathbf{a}^n \end{bmatrix}. \tag{A4}$$

Furthermore, Eq. (A4) can be expressed as

$$\mathbf{\Pi}_i = [a_{i1} \cdot \mathbf{a}_1 \; a_{i2} \cdot \mathbf{a}_2 \; \cdots \; a_{ip} \cdot \mathbf{a}_p] \tag{A5}$$

where a_{ij} is the (i,j)th element in $\mathbf{A}$ and $\mathbf{a}_i$ is the ith column vector of $\mathbf{A}$. Given that $\mathbf{A}$ is of full column rank, the $\{\mathbf{a}_i\}_{i=1}^p$'s are linearly independent. For this reason, the columns in $\mathbf{\Pi}_i$ are also linearly independent, leading to the conclusion that the first p columns in $\mathbf{\Pi}$ are also linearly independent. Assume the first p columns in $\mathbf{\Pi}$ are a linear combination of the $(p+1)$th column. Then, $\exists$ a nontrivial vector $\boldsymbol{\alpha}=[\alpha_1 \cdots \alpha_{p+1}]^T$ such that

$$\Sigma_{j=1}^p \alpha_j a_{ij} \cdot \mathbf{a}_j + \alpha_{p+1} \cdot \mathbf{e}_i = \mathbf{0} \text{ for } i = 1, \ldots, n, \tag{A6}$$

Denoting $\mathbf{\Lambda} = \text{diag}\{\alpha_1 \cdots \alpha_p\}$, Eq. (A6) can be written as

$$\mathbf{A} \cdot \mathbf{\Lambda} \cdot (\mathbf{a}^i)^T + \alpha_{p+1} \cdot \mathbf{e}_i = \mathbf{0} \text{ for } i = 1, \ldots, n \tag{A7}$$

All n (for $i=1, \ldots, n$) equations can then be arranged to yield

$$\mathbf{A} \cdot \mathbf{\Lambda} \cdot \mathbf{A}^T + \alpha_{p+1} \cdot \mathbf{I}_n = \mathbf{0} \tag{A8}$$

in which the results of $\mathbf{A}^T=[(\mathbf{a}^1)^T \cdots (\mathbf{a}^i)^T \cdots (\mathbf{a}^n)^T]$ and $\mathbf{I}_n = [\mathbf{e}_1 \cdots \mathbf{e}_n]$ are used. Given that $\mathbf{A}$ is of rank p and $\mathbf{\Lambda}$ is also of rank p, the rank of $\mathbf{A} \cdot \mathbf{\Lambda} \cdot \mathbf{A}^T$ is at most p. However, $\alpha_{p+1} \cdot \mathbf{I}_n$ is of rank n if $\alpha_{p+1} \neq 0$. Thus, Eq. (A8) cannot hold since $n > p$. If $\alpha_{p+1}=0$, $\mathbf{A} \cdot \mathbf{\Lambda} \cdot \mathbf{A}^T$ cannot equal $\mathbf{0}$ unless all $\{\alpha_i\}_{i=1}^p$ elements are also zero. This contradiction indicates that the first p columns in $\mathbf{\Pi}$ are not a linear combination of the $(p+1)$th column. Thus $\mathbf{\Pi}^T\mathbf{\Pi}$ is of full rank. □

Appendix C: Derivation of Eq. (21) in Sec. 3.2

From Eq. (12), it is known that $\hat{\sigma}_v^2 = \text{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\mathbf{S}_y)/(n-p)$. Since $\text{tr}(\mathbf{AB}) = \text{vec}(\mathbf{A})^T\text{vec}(\mathbf{B})$ for any symmetric matrices $\mathbf{A}$ and $\mathbf{B}$, then $\hat{\sigma}_v^2 = (\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T\text{vec}(\mathbf{S}_y)/(n-p)$. Consequently,

$$\begin{aligned} Var(\hat{\sigma}_v^2) &= \frac{Var((\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T\text{vec}(\mathbf{S}_y))}{(n-p)^2} \\ &= \frac{(\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T \cdot Cov(\text{vec}(\mathbf{S}_y)) \cdot \text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)}{(n-p)^2}. \end{aligned} \tag{A9}$$

Given $Cov(\text{vec}(\mathbf{S}_y)) = (\mathbf{I}_{n^2} + \mathbf{K}_n)(\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y)/N$ from Magnus and Neudecker [29], where $\mathbf{K}_n$ is defined as the "commutation matrix," then

$$\begin{aligned} &Var(\hat{\sigma}_v^2) \\ &= \frac{(\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T \cdot (\mathbf{I}_{n^2} + \mathbf{K}_n) \cdot (\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y) \cdot \text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)}{N(n-p)^2}. \end{aligned} \tag{A10}$$

From Eq. (2.1) in Magnus and Neudecker [29],

$$\begin{aligned} (\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y) \cdot \text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+) &= \text{vec}(\mathbf{\Sigma}_y(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\mathbf{\Sigma}_y) \\ &= \text{vec}((\mathbf{A}\mathbf{\Sigma}_u\mathbf{A}^T + \sigma_v^2\mathbf{I}_n) \cdot (\mathbf{I}_n \\ &\quad - \mathbf{A}\mathbf{A}^+) \cdot (\mathbf{A}\mathbf{\Sigma}_u\mathbf{A}^T + \sigma_v^2\mathbf{I}_n)) \\ &= \text{vec}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)) \cdot \sigma_v^4 \end{aligned} \tag{A11}$$

Further, the following two results can be obtained:

$$\begin{aligned} &(\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T \cdot (\mathbf{I}_{n^2}) \cdot \text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+) \cdot \sigma_v^4 \\ &= \text{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+) \cdot (\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)) \cdot \sigma_v^4 \\ &= (\text{tr}(\mathbf{I}_n) - \text{tr}(\mathbf{A}\mathbf{A}^+)) \cdot \sigma_v^4 \\ &= (\text{tr}(\mathbf{I}_n) - \text{tr}(\mathbf{A}^+\mathbf{A})) \cdot \sigma_v^4 = (n-p) \cdot \sigma_v^4, \end{aligned} \tag{A12}$$

and

$$\begin{aligned} &(\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T \cdot (\mathbf{K}_n) \cdot \text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+) \cdot \sigma_v^4 \\ &= (\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T \cdot \text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+) \cdot \sigma_v^4 \\ &= (n-p) \cdot \sigma_v^4, \end{aligned} \tag{A13}$$

in which the result $\mathbf{K}_n\text{vec}(\mathbf{D}) = \text{vec}(\mathbf{D}^T)$ was utilized, where $\mathbf{D}$ is an $n \times n$ matrix, also from Magnus and Neudecker ([29], Theorem 3.1-(vii)). Finally, the results from (A11)–(A13) are substituted into (A10) and to yield $Var(\hat{\sigma}_v^2) = 2\sigma_v^4/N(n-p)$ (Eq. (21)). This result is different from that in Bohme ([8], Eq. (16)), as a constant "2" is generated. It is believed that the difference is due to the fact that $\mathbf{S}_y$ was implicitly assumed to be Toeplitz as in Bohme [8]. For a Toeplitz $\mathbf{S}_y$, $Cov(\text{vec}(\mathbf{S}_y)) = \mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y/N$. (See Li et al. [34] for the covariance matrix of a Toeplitz $\mathbf{S}_y$.) □

Appendix D: Derivation of Eq. (22) in Sec. 3.2

From Eq. (17),

$$\begin{aligned} Cov(\hat{\boldsymbol{\sigma}}_{1,p}^2) &= \mathbf{P}_1 \cdot Cov(\text{vec}(\mathbf{S}_y)) \cdot \mathbf{P}_1^T - Cov(\mathbf{P}_1 \cdot \text{vec}(\mathbf{S}_y), \mathbf{P}_2\hat{\sigma}_v^2) \\ &\quad + \mathbf{P}_2\mathbf{P}_2^T Var(\hat{\sigma}_v^2), \end{aligned} \tag{A14}$$

given $Cov(\text{vec}(\mathbf{S}_y)) = (\mathbf{I}_{n^2} + \mathbf{K}_n)(\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y)/N$ [29] and $Var(\hat{\sigma}_v^2)$ in Eq. (21), Eq. (A14) can be expressed as

$$\begin{aligned} Cov(\hat{\boldsymbol{\sigma}}_{1,p}^2) &= \frac{1}{N}\mathbf{P}_1 \cdot (\mathbf{I}_{n^2} + \mathbf{K}_n)(\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y) \cdot \mathbf{P}_1^T \\ &\quad - Cov(\mathbf{P}_1 \cdot \text{vec}(\mathbf{S}_y), \mathbf{P}_2\hat{\sigma}_v^2) + \mathbf{P}_2\mathbf{P}_2^T \cdot \frac{2\sigma_v^4}{N(n-p)}. \end{aligned} \tag{A15}$$

It is then necessary to simplify the second term in (A15). Recalling that $\hat{\sigma}_v^2 = \text{tr}((\mathbf{I}_n - \mathbf{A}\mathbf{A}^+)\mathbf{S}_y)/(n-p) = (\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T\text{vec}(\mathbf{S}_y)/(n-p)$ (see Appendix C), this second term becomes

$$\begin{aligned} Cov(\mathbf{P}_1 \cdot \text{vec}(\mathbf{S}_y), \mathbf{P}_2\hat{\sigma}_v^2) &= Cov\left(\mathbf{P}_1 \cdot \text{vec}(\mathbf{S}_y), \frac{1}{n-p}\mathbf{P}_2(\text{vec}(\mathbf{I}_n \right. \\ &\quad \left. - \mathbf{A}\mathbf{A}^+))^T\text{vec}(\mathbf{S}_y)\right) \\ &= \frac{1}{n-p}\mathbf{P}_1 \cdot Cov(\text{vec}(\mathbf{S}_y), \text{vec}(\mathbf{S}_y)) \cdot \mathbf{P}_3^T \\ &= \frac{1}{N(n-p)}\mathbf{P}_1(\mathbf{I}_{n^2} + \mathbf{K}_n)(\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y)\mathbf{P}_3^T \end{aligned} \tag{A16}$$

where $\mathbf{P}_3 = \mathbf{P}_2 \cdot (\text{vec}(\mathbf{I}_n - \mathbf{A}\mathbf{A}^+))^T$ and the result of $Cov(\text{vec}(\mathbf{S}_y)) = (\mathbf{I}_{n^2} + \mathbf{K}_n)(\mathbf{\Sigma}_y \otimes \mathbf{\Sigma}_y)/N$ is again used. Given Eqs. (A15) and (A16), it can be readily verified that the variance of a DE estimator, $\text{var}_{DE} = \text{tr}(\text{Cov}(\hat{\boldsymbol{\sigma}}_{1,p}^2)) + \text{Var}(\hat{\sigma}_v^2)$, has the same expression as in Eq. (22). □

Appendix E: Matrix A in Eq. (32)

$$\mathbf{A} = [\mathbf{A}_x \ \mathbf{A}_z] = \begin{bmatrix}
0 & 0 & 0 & 0.1215 & -0.3846 & 0 & 0 & 0 & 0.2632 \\
0 & 0 & 0 & 0.0221 & -0.0699 & 0 & 0 & 0 & 0.0478 \\
0 & 0 & 0 & 0.1215 & -0.3846 & 0 & 0 & 0 & 0.2632 \\
0 & 0 & 0 & -0.1817 & 0.5944 & 0 & 0 & 0 & -0.4067 \\
0 & 0 & 0 & -0.0773 & 0.2448 & 0 & 0 & 0 & -0.1675 \\
0 & 0 & 0 & -0.3379 & 1.0699 & 0 & 0 & 0 & -0.7321 \\
0 & 0 & 0 & 0.1656 & -0.5245 & 0 & 0 & 0 & 0.3589 \\
0 & 0 & 0 & -0.3379 & 1.0699 & 0 & 0 & 0 & -0.7321 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -0.2054 & 0.6503 & 0 & 0 & 0 & -0.445 \\
-1 & 1 & 0 & -0.3110 & 0 & 0.4 & -0.4 & 0 & 0.311 \\
0 & 0 & 0 & 0.0574 & 0 & -0.24 & 1.24 & 0 & -1.0574 \\
-1 & 1 & 0 & -0.2153 & 0 & 0 & 0 & 0 & 0.2153 \\
0 & 0 & 0 & -0.2392 & 0 & 1 & 0 & 0 & -0.7608 \\
-1 & 0 & 1 & -0.0957 & 0 & 0 & 0 & 0.4 & -0.3043 \\
0 & 0 & 0 & 0.0574 & 0 & 0 & 0 & -0.24 & 0.1826 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -0.2392 & 0 & 0 & 0 & 1 & -0.7608
\end{bmatrix}_{18\times 9}$$

References

[1] Hu, S. J., 1990, *Impact of 100% Measurement Data on Statistical Process Control (SPC) in Automobile Body Assembly*, Ph.D. thesis, The University of Michigan, Ann Arbor, MI.

[2] Montgomery, D. C., 1996, *Introduction to Statistical Quality Control*, 3rd ed., Wiley & Sons, Inc., New York.

[3] Hu, S. J., and Wu, S., (1992), "Identifying Sources of Variation in Automobile Body Assembly Using Principal Component Analysis," Transactions of NAMRI/SME, **XX**, pp. 311–316.

[4] Ceglarek, D., and Shi, J., 1996, "Fixture Failure Diagnosis for Auto Body Assembly Using Pattern Recognition," ASME J. Eng. Ind., **118**, pp. 55–66.

[5] Rao, C. R., and Kleffe, J., 1988, *Estimation of Variance Components and Applications*, Amsterdam: North-Holland, Amsterdam.

[6] Searle, S. R., Casella, G., and McCulloch, C. E., 1992, *Variance Components*, Wiley & Sons, New York.

[7] D'Assumpcao, H. A., 1980, "Some New Signal Processors for Arrays of Sensors," IEEE Transactions on Information Technology, **IT-26**, pp. 441–453.

[8] Bohme, J. F., 1986, "Estimation of Special Parameters of Correlated Signals in Wavefields," Signal Process., **11**, pp. 329–337.

[9] Stoica, P., and Nehorai, A., 1995, "On the Concentrated Stochastic Likelihood Function in Array Signal Processing," Circuits Syst. Signal Process., **14**, pp. 669–674.

[10] Apley, D. W., and Shi, J., 1998, "Diagnosis of Multiple Fixture Faults in Panel Assembly," ASME ASME J. Manuf. Sci. Eng., **120**, pp. 793–801.

[11] Ding, Y., Shi, J., and Ceglarek, D., 2002, "Diagnosability Analysis of Multi-station Manufacturing Processes," ASME J. Dyn. Syst., Meas., Control, **124**, pp. 1–13.

[12] Bush, N., and Anderson, R. L., 1963, "A Comparison of Three Different Procedures for Estimating Variance Components," Technometrics, **5**, pp. 421–440.

[13] Corbeil, R. R., and Searle, S. R., 1976, "A Comparison of Variance Component Estimators," Biometrics, **32**, pp. 779–791.

[14] Swallow, W. H., and Monahan, J. F., 1984, "Monte-Carlo Comparison of ANOVA, MINQUE, REML, and ML Estimators of Variance Components," Technometrics, **26**, pp. 47–57.

[15] Chang, M., and Gossard, D. C., 1998, "Computational Method for Diagnosis of Variation-Related Assembly Problem," Int. J. Prod. Res., **36**, pp. 2985–2995.

[16] Carlson, J. S., Lindkvist, L., and Soderberg, R., 2000, "Multi-Fixture Assembly System Diagnosis Based on Part and Subassembly Measurement Data," Proceedings of the 2000 ASME Design Engineering Technical Conference, Baltimore, MD, 10–13 September.

[17] Mantripragada, R., and Whitney, D. E., 1999, "Modeling and Controlling Variation Propagation in Mechanical Assemblies Using State Transition Models," IEEE Trans. Rob. Autom., **15**, pp. 124–140.

[18] Jin, J., and Shi, J., 1999, 'State Space Modeling of Sheet Metal Assembly for Dimensional Control," ASME J. Manuf. Sci. Eng., **121**, pp. 756–762.

[19] Ding, Y., Ceglarek, D., and Shi, J., 2000, "Modeling and Diagnosis of Multi-stage Manufacturing Processes: Part I State Space Model," Proceedings of the 2000 Japan/USA Symposium on Flexible Automation, Ann Arbor, MI, 23–26 July 2000, JUSFA-13146.

[20] Camelio, A. J., Hu, S. J., and Ceglarek, D. J., (2001) "Modeling Variation Propagation of Multi-station Assembly Systems with Compliant Parts," Proceedings of the 2001 ASME Design Engineering Technical Conference, Pittsburgh, PA, 9–12 September.

[21] Huang, Q., Zhou, N., and Shi, J., 2000, "Stream of Variation Modeling and Diagnosis of Multi-Station Machining Processes," Proceedings of IMECE 2000, MED-Vol. 11, pp. 81–88, Orlando, FL, 5–10 November.

[22] Djurdjanovic, D., and Ni, J., 2001, "Linear State Space Modeling of Dimensional Machining Errors," Transactions of NAMRI/SME, **XXIX**, pp. 541–548.

[23] Zhou, S., Huang, Q., and Shi, J., 2003b, 'State Space Modeling for Dimensional Monitoring of Multistage Machining Process Using Differential Motion Vector," IEEE Trans. Rob. Autom., **19**, pp. 296–308.

[24] Suri, R., and Otto, K., 1999, "Variation Modeling for a Sheet Stretch Forming Manufacturing System," Annals of CIRP, **48**, pp. 397–400.

[25] Schott, J. R., 1997, *Matrix Analysis for Statistics*, Wiley & Sons, New York.

[26] Anderson, T. W., 1969, "Statistical Inference for Covariance Matrices with Linear Structure," *Proceedings of the Second International Symposium of Multivariate Analysis*, P. R. Krishnaiah, ed., Academic Press, New York, pp. 55–66.

[27] Anderson, T. W., 1970, "Estimation of Covariance Matrices Which Are Linear Combinations or Whose Inverses Are Linear Combinations of Given Matrices," In *Essays in Probability and Statistics*, R. C. Rose et al., eds., University of North Carolina Press, Chapel Hill, pp. 1–24.

[28] Zhou, S., Ding, Y., Chen, Y., and Shi, J., 2003a, "Diagnosability Study of Multistage Manufacturing Process Based on Linear Mixed-Effects Models," Technometrics, **45**, pp. 312–325.

[29] Magnus, J. R., and Neudecker, H., 1979, "The Commutation Matrix: Some Properties And Applications," Ann. Stat., **7**, pp. 381–394.

[30] Shalon, D., Gossard, D., Ulrich, K., and Fitzpatrick, D., 1992, "Representing Geometric Variations in Complex Structural Assemblies on CAD Systems," Proceedings of the 19th Annual ASME Advances in Design Automation Conference, DE-44-2, pp. 21–132.

[31] Ceglarek, D., and Shi, J., 1995, "Dimensional Variation Reduction for Automotive Body Assembly," Manuf. Rev., **8**, pp. 39–154.

[32] Qin, S. J., 2003, "Statistical Process Monitoring: Basics and Beyond," Journal of Chemometrics, **17**, pp. 480–502.

[33] Barton, R. R., and Gonzalez-Barreto, D. R., 1996, "Process-oriented Basis Representations for Multivariate Process Diagnosis," Quality Engineering, **9**, pp. 107–118.

[34] Li, H., Stoica, P., and Li, J., 1999, "Computationally Efficient Maximum Likelihood Estimation of Structured Covariance Matrices," IEEE Trans. Signal Process., **47**, pp. 1314–1323.