

Guidelines for placing additional sensors to improve variation diagnosis in assembly processes

Y. DING*† and D. W. APLEY‡

†Department of Industrial and Systems Engineering, Texas A&M University,
TAMU 3131, College Station, TX 77843-3131, USA

‡Department of Industrial Engineering and Management Sciences,
Northwestern University, Evanston, IL 60208-3119, USA

(Revision received March 2007)

Dimensional variation reduction is critical to assuring high quality in assembly and manufacturing processes. The extent to which data from a multiple-sensor system aids the diagnosis of variation sources depends on the effective placement of the sensors. The diagnostic objective that we consider is to estimate the variance components for potential variation sources. Using a linear structured model to represent the effects of the variation sources on the measurement data, this paper studies the problem of how to add additional sensor(s) to ensure diagnosability and/or improve estimation accuracy. Most prior work on sensor placement focused on automated numerical search algorithms that optimize rather unintuitive mathematical measures of diagnosability and accuracy. Our objective is to translate the measures into expressions that provide better conceptual guidance into how to most appropriately locate additional sensors to improve accuracy and diagnosability. The expressions may be used in conjunction with qualitative judgment and expert knowledge as the basis for locating additional sensors. Alternatively, they can be used in conjunction with existing numerical search routines by providing initial guesses for the sensor location and/or substantially narrowing the space of feasible sensor locations that must be searched during the numerical optimization. The proposed method is illustrated with examples from automotive panel assembly.

Keywords: Assembly systems; Dimensional variation reduction; Fault diagnosis; Sensor placement; Variance component estimation

1. Introduction

Dimensional integrity is a major quality concern in many discrete-part manufacturing processes such as assembly (Shalon *et al.* 1992, Ceglarek and Shi 1995). To ensure high-quality, automated dimensional sensing devices are commonly used, providing measurements of dimensional features on every product and thus enabling a level of manufacturing fault diagnosis that would otherwise be impossible. In recent years there has been considerable work on reducing dimensional variation in assembly processes, in particular in automobile body assembly (Hu and Wu 1992, Ceglarek *et al.* 1994, Ceglarek and Shi 1996, Apley and Shi 1998, Chang and Gossard 1998,

*Corresponding author. Email: yuding@iemail.tamu.edu

Rong *et al.* 2000, Carlson *et al.* 2000, Apley and Shi 2001, Ding *et al.* 2002a). The broad objective is to effectively utilize the dimensional measurement data for the purpose of identifying (and subsequently eliminating) major root causes of part-to-part dimensional variation.

The above referenced papers have developed measurement data analysis algorithms for diagnosing root causes of variation, with particular emphasis on fixture- and other tooling-related variation sources. The typical first step is to establish a model that links product measurements to process variance sources. In quality control applications, a linear or linearized model is often considered acceptable for representing the effects of the variation sources on the measurement data. Almost all the aforementioned approaches employ the following linear structured model:

$$\mathbf{y}(t) = \mathbf{\Gamma}\mathbf{u}(t) + \mathbf{v}(t) = \sum_{i=1}^p \mathbf{\Gamma}_i u_i(t) + \mathbf{v}(t): \quad t = 1, 2, \dots, N, \quad (1)$$

where $\mathbf{y}(t)$ is a vector of n measured product features on part number t , $\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_p(t)]^T$ is a random vector whose elements represent p independent variation sources, $\mathbf{v}(t)$ is an additive random noise vector (e.g. sensor noise), $\mathbf{\Gamma} = [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \dots, \mathbf{\Gamma}_p]$ is an $n \times p$ diagnostic matrix of known coefficients relating the variation sources to the measurement vector, t is an observation or part index, and N is the sample size. The quantity $\mathbf{\Gamma}_i u_i(t)$ therefore represents the effects of the i th variation source on the measurements for part number t of the sample. Because the elements of $\mathbf{u}(t)$ are assumed independent, its covariance matrix is a diagonal matrix $\mathbf{\Sigma}_u = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2\}$. The sensor system is normally assumed to be homogenous, so that the elements of $\mathbf{v}(t)$ are independent with equal variance σ^2 . In other words, the covariance matrix of $\mathbf{v}(t)$ is $\mathbf{\Sigma}_v = \sigma^2 \mathbf{I}$.

In addition to the aforementioned work on panel assembly, models of the form (1) have been used to represent the effects of variation sources in crankshaft manufacturing (Apley and Lee 2003), electronics assembly (Barton and Gonzalez-Barreto 1996), polyester film processing (Qin 2003), machining (Djurdjanovic and Ni 2001, Zhou *et al.* 2003b), and many other applications. It is a special case of the standard linear mixed effects model used in variance components analysis (refer to Searle *et al.* 1992). We can view (1) as more informative than the general linear model, because the columns of $\mathbf{\Gamma}$ incorporate a great deal of information on the geometry of parts, the nature and location of the tooling elements and other variation sources, and the locations of the sensors. There have been a number of recent analytical modelling developments to aid in obtaining $\mathbf{\Gamma}$ based on engineering design knowledge, process physics, and a specified set of potential variation sources (Ceglarek *et al.* 1994, Jin and Shi 1999, Mantripragada and Whitney 1999, Ding *et al.* 2000, Djurdjanovic and Ni 2001, Camelio *et al.* 2003, Zhou *et al.* 2003b).

The diagnostic objective in this paper and in most of the aforementioned references is to estimate the variance components $\{\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2\}$ for each of the p potential variation sources, based on a sample of data and a $\mathbf{\Gamma}$ matrix obtained from prior off-line modelling. The reason that most of the aforementioned work has focused on variance components, as opposed to mean components, is that variance components are more challenging to estimate and to correct in dimensional quality control. A sustained, consistent deviation from nominal (i.e. a mean shift) can often

be compensated relatively easily by process engineers via shimming and other adjustments. In contrast, variance is much more difficult to compensate and requires either some form of on-line feedback control or the removal of the root cause. Estimation of the variance components allows us to assess whether each variation source is present in the current sample and, if so, the severity of the source.

All of the existing diagnostic algorithms (details of which will be given in the subsequent section) require a set of diagnosability conditions to be satisfied in order to produce valid, unique estimates of the variance components. This is analogous to the singularity issue in standard least squares that results in non-unique parameter estimates. The diagnosability conditions for the various algorithms (see Apley and Shi 1998, Chang and Gossard 1998, Ding *et al.* 2002b, Zhou *et al.* 2003a) typically involve checking whether a certain matrix is singular. Further studies have decomposed the diagnosability conditions to reveal coupling relations among the variation sources in partially diagnosable systems (Zhou *et al.* 2003a) or to provide a set of more conceptually meaningful conditions that offers better insight into the reasons behind the diagnosability problems (Apley and Ding 2005). Various measures of estimation accuracy, which is closely related to diagnosability, have also been developed by Wang and Nagarkar (1999), Chang and Gossard (1998), Carlson *et al.* (2000), and Ding *et al.* (2005), often using one of the alphabetic optimality criteria (such as D-, A-, or E-optimality) that originated in optimal experimental design research (Atkinson and Donev 1992).

Diagnosability and accuracy are important issues to consider when laying out a system of sensors. Existing sensor layout strategies typically involve searching over a pre-determined set of candidate sensor layouts in order to optimize appropriate diagnosability and/or accuracy measures (Khan *et al.* 1998, 1999, Wang and Nagarkar 1999, Khan and Ceglarek 2000, Djurdjanovic and Ni 2003, Ding *et al.* 2003, 2004, Liu *et al.* 2005). Because the optimization algorithms are automated and the diagnosability and accuracy measures are mathematical in nature, it is difficult to incorporate other more qualitative sensor layout criteria that result from the system designer's expert knowledge and engineering judgment.

The main purpose of this paper is to study strategies of how to place additional sensors in order to improve an existing system of sensors that currently lacks diagnosability and/or sufficient estimation accuracy. In contrast to previous automated numerical search methods for optimizing a sensor layout, we decompose the diagnosability and accuracy measures into expressions that are more convenient for providing conceptual guidance in where to place an additional sensor. The guidelines can be used by themselves to provide an immediate suboptimal solution of where to place the additional sensor. Alternatively, they can be used in conjunction with numerical optimization algorithms to provide better initial guesses for the additional sensor location(s) and/or narrow down the range of feasible locations over which the optimization algorithm must search.

The format of the remainder of the paper is as follows. In the next section, we review the variance components estimation method and its mathematical diagnosability conditions. Next, we discuss the strategy for adding an additional sensor to an existing layout in order to ensure diagnosability and illustrate with an example. Following this, we present the strategy for adding an additional sensor to an existing layout in order to improve estimation accuracy, as opposed to ensure diagnosability, and illustrate with an example. We then present a more involved

example and show how the proposed method can help place an additional sensor for diagnosability and accuracy requirements in a panel assembly process. The last section concludes the paper.

2. Variance components estimation and diagnosability conditions

Throughout, the ‘ $\hat{\cdot}$ ’ overscore symbol denotes an estimate of a parameter. Unless otherwise noted, we assume the sample mean $\bar{\mathbf{y}} = N^{-1} \sum_{t=1}^N \mathbf{y}(t)$ has been subtracted from the data so that the resulting sample $\{\mathbf{y}(t): t=1, 2, \dots, N\}$ is taken to be zero-mean.

The following approach for estimating the variance components was investigated in D’Assumpcao (1980), Bohme (1986), and Ding *et al.* (2004). Express the covariance matrix of \mathbf{y} as $\Sigma_{\mathbf{y}} = \sum_{i=1}^p \Gamma_i \Gamma_i^T \sigma_i^2 + \sigma^2 \mathbf{I}$, and consider the sample covariance matrix $\mathbf{S}_{\mathbf{y}} = (N-1)^{-1} \sum_{t=1}^N (\mathbf{y}(t) - \bar{\mathbf{y}})(\mathbf{y}(t) - \bar{\mathbf{y}})^T$ as an estimate of $\Sigma_{\mathbf{y}}$. The variance component estimates are taken to be the values that minimize the sum of squares of the elements of the error matrix $\mathbf{S}_{\mathbf{y}} - \sum_{i=1}^p \Gamma_i \Gamma_i^T \hat{\sigma}_i^2 + \hat{\sigma}^2 \mathbf{I}$. The estimates for this approach, which we refer to as matrix least squares (MLS), are given by

$$\hat{\boldsymbol{\sigma}} = \mathbf{G}^{-1} \mathbf{b}, \quad (2)$$

where $\hat{\boldsymbol{\sigma}} = [\hat{\sigma}_1^2 \quad \hat{\sigma}_1^2 \quad \dots \quad \hat{\sigma}_p^2 \quad \hat{\sigma}^2]^T$ is the vector of variance component estimates,

$$\mathbf{G} = \begin{bmatrix} (\Gamma_1^T \Gamma_1)^2 & \dots & (\Gamma_1^T \Gamma_p)^2 & \Gamma_1^T \Gamma_1 \\ \vdots & & \vdots & \\ (\Gamma_1^T \Gamma_p)^2 & \dots & (\Gamma_p^T \Gamma_p)^2 & \Gamma_p^T \Gamma_p \\ \Gamma_1^T \Gamma_1 & \dots & \Gamma_p^T \Gamma_p & n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \Gamma_1^T \mathbf{S}_{\mathbf{y}} \Gamma_1 \\ \vdots \\ \Gamma_p^T \mathbf{S}_{\mathbf{y}} \Gamma_p \\ \text{tr}(\mathbf{S}_{\mathbf{y}}) \end{bmatrix}. \quad (3)$$

The diagnosability condition for the MLS algorithm is that \mathbf{G} has full rank $p+1$, so that its inverse exists in equation (2). Because \mathbf{G} is a Gram matrix of (appropriately defined) inner products between pairs of matrices in the set $\{\Gamma_1 \Gamma_1^T, \Gamma_2 \Gamma_2^T, \dots, \Gamma_p \Gamma_p^T, \mathbf{I}\}$, an equivalent condition for diagnosability is that these matrices are linearly independent. Otherwise, the variance component estimates that minimize the MLS error criterion are not unique.

Ding *et al.* (2004) also presented a weighted least squares version of the MLS algorithm designed to improve estimation accuracy. We refer to this as the weighted MLS (WMLS) algorithm. Because the elements of the estimation error matrix $\mathbf{S}_{\mathbf{y}} - \Sigma_{\mathbf{y}}$ in the MLS algorithm are not uncorrelated with equal variance, the MLS algorithm does not inherit certain desirable properties of least squares estimation. This is corrected if we pre- and post-multiply the error matrix $\mathbf{S}_{\mathbf{y}} - \sum_{i=1}^p \Gamma_i \Gamma_i^T \hat{\sigma}_i^2 + \hat{\sigma}^2 \mathbf{I}$ in the MLS criterion by the inverse square root matrix $\Sigma_{\mathbf{y}}^{-1/2}$. This results in an iterative algorithm (refer to Ding *et al.* 2004 for details), where an iteration over the following two steps is required because we do not know the true $\Sigma_{\mathbf{y}}$.

(1) Based on the estimate $\hat{\boldsymbol{\sigma}}$ at the previous iteration, calculate

$$\hat{\Sigma}_{\mathbf{y}} = \sum_{i=1}^p \Gamma_i \Gamma_i^T \hat{\sigma}_i^2 + \hat{\sigma}^2 \mathbf{I}.$$

(2) Calculate the new estimate $\hat{\sigma} = (\mathbf{G}^*)^{-1}\mathbf{b}^*$ at the current iteration, where

$$\mathbf{G}^* = \begin{bmatrix} (\Gamma_1^T \hat{\Sigma}_y^{-1} \Gamma_1)^2 & \cdots & (\Gamma_1^T \hat{\Sigma}_y^{-1} \Gamma_p)^2 & \Gamma_1^T \hat{\Sigma}_y^{-2} \Gamma_1 \\ \vdots & & \vdots & \\ (\Gamma_p^T \hat{\Sigma}_y^{-1} \Gamma_p)^2 & \cdots & (\Gamma_p^T \hat{\Sigma}_y^{-1} \Gamma_p)^2 & \Gamma_p^T \hat{\Sigma}_y^{-2} \Gamma_p \\ \Gamma_1^T \hat{\Sigma}_y^{-2} \Gamma_1 & \cdots & \Gamma_p^T \hat{\Sigma}_y^{-2} \Gamma_p & tr(\hat{\Sigma}_y^{-2}) \end{bmatrix}, \text{ and } \mathbf{b}^* = \begin{bmatrix} \Gamma_1^T \hat{\Sigma}_y^{-1} \mathbf{S}_y \hat{\Sigma}_y^{-1} \Gamma_1 \\ \vdots \\ \Gamma_p^T \hat{\Sigma}_y^{-1} \mathbf{S}_y \hat{\Sigma}_y^{-1} \Gamma_p \\ tr(\hat{\Sigma}_y^{-2} \mathbf{S}_y) \end{bmatrix}. \tag{4}$$

At the initial iteration, we can use the estimate $\hat{\sigma}$ from the MLS algorithm. Provided that $\hat{\Sigma}_y$ remains positive definite at each iteration, which can be guaranteed by resetting any negative variance component estimates to zero (and a negative $\hat{\sigma}^2$ should be reset to some small positive value), the diagnosability conditions for the WMLS and MLS algorithms are equivalent (Ding *et al.* 2004).

Given that model (1) is a special case of a mixed linear model, one might consider any of the estimation methods developed in the broad body of variance component analysis (VCA) literature (see Searle *et al.* 1992, Rao and Kleffe 1988), such as a maximum likelihood estimator (MLE), a restricted maximum likelihood (REML) estimator, or Rao’s MINQUE. In fact, the MLS estimator is precisely the MINQUEO or MINQUE(0) estimator, a special form of Rao’s MINQUE with the initial Σ_y chosen to be the identity matrix (Searle *et al.* 1992). Anderson (1973) proved that the WMLS estimator is an asymptotical MLE. Moreover, if we start with an MLS estimator and proceed with the WMLS iterations, the entire procedure is equivalent to Rao’s iterative MINQUE. Therefore, the MLS and WMLS variance estimators are equivalent to those developed in VCA theory, retaining their desirable statistical properties. We believe that the above presentation of the MLS and WMLS algorithms are intuitive to practitioners and also make the connection to their diagnosability conditions more transparent.

It is worth noting that the diagnosability condition for variance components estimation is different from the diagnosability condition for the ordinary least-squares (OLS) estimation approach taken in, for example, Apley and Shi (1998). The OLS approach involves calculating $\hat{\mathbf{u}}(t) = [\Gamma^T \Gamma]^{-1} \Gamma^T \mathbf{y}(t)$ for $t = 1, 2, \dots, N$, for which the diagnosability conditions are a non-singular $\Gamma^T \Gamma$ and $n > p$. The (W)MLS algorithm is diagnosable whenever the OLS algorithm is, but the converse is not necessarily true (Ding *et al.* 2004).

To illustrate the model and diagnosability concepts, consider the following example of dimensional variability in the liftgate opening of a minivan (see Apley and Ding 2005 for further details). Figure 1(a) shows the liftgate opening, which is illustrated schematically as a box in figure 1(b). Suppose that sensors 1–4 are positioned around the liftgate opening as in figure 1(c) (temporarily ignore sensors 5–7, which will be used later). The sensors positioned on the bodyside (sensors 1 and 2) measure the left/right dimensional deviation from nominal at their respective locations. The sensors positioned on the roof cross-member (sensors 3 and 4) measure the up/down deviation from nominal. Deviations in the up and right directions are taken to be positive. Deviations in the left and down directions are taken to be negative.

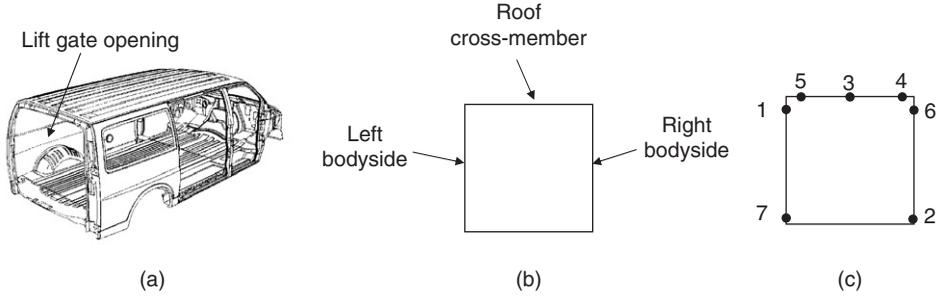


Figure 1. Illustration of the actual liftgate opening (a), a schematic box representation (b), and a potential sensor layout (c) with four sensors numbered 1 to 4.

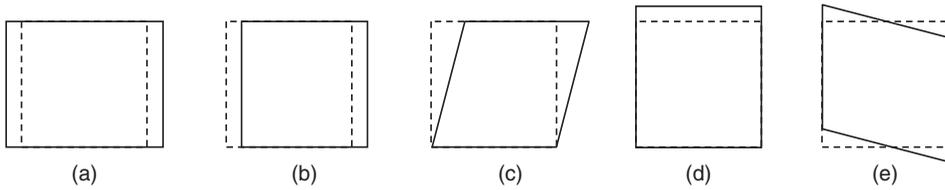


Figure 2. Illustration of five variation patterns affecting the liftgate opening: (a) pattern 1, a horizontal enlargement; (b) pattern 2, a horizontal translation; (c) pattern 3, a horizontal matchboxing; (d) pattern 4, a vertical enlargement; and (e) pattern 5, a vertical matchboxing.

Suppose we are interested in diagnosing the five potential variation patterns illustrated in figure 2, each of which is a relatively common occurrence as the tooling becomes worn, loose, broken, etc. Throughout, we will refer to the effects of a variation source as a variation pattern and the corresponding column of Γ as a pattern vector. Note that each pattern represents part-to-part variation, as opposed to a mean shift. For example, although pattern 1 is shown as a positive enlargement in figure 2(a), on some autobodies in the sample the enlargement may be negative (a contraction) depending on whether the value of $u_1(t)$ was positive or negative for that autobody.

Determination of Γ in this example is fairly straightforward. Recall that we only have sensors 1–4 for the time being. When pattern 1 in figure 2 happens with a magnitude of u_1 , the sensor reading vector is $[-u_1 \ u_1 \ 0 \ 0]^T$, which makes the first column of Γ to be $[-1 \ 1 \ 0 \ 0]^T$ after the effect of input u_1 is taken out from the sensor outputs. The similar procedure can be employed to determine other columns in Γ based on the geometry of the patterns and the locations of the sensors. Eventually, the diagnostic matrix for this example is

$$\Gamma = \begin{bmatrix} -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}. \tag{5}$$

It can be verified that $\mathbf{\Gamma}^T\mathbf{\Gamma}$ is singular, but the Gram matrix

$$\mathbf{G} = \begin{bmatrix} 4 & 0 & 1 & 0 & 0 & 2 \\ 0 & 4 & 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 2 & 2 & 1 & 2 & 1 & 4 \end{bmatrix} \tag{6}$$

is full rank. Thus, the (W)MLS algorithm is diagnosable with the four-sensor layout, whereas the OLS algorithm would not be. In the remainder of this paper, the sensor placement strategies assume that the (W)MLS algorithm is used.

3. Adding individual sensors for diagnosability

Suppose that we have a sensor layout with n sensors currently installed, and we decide that an additional sensor must be added. For example, we may decide that we need to add a sensor to improve accuracy. Alternatively, we may discover an additional potential variation source that must be added to the model. After increasing the number of sources in the model, the original layout with n sensors may no longer be diagnosable, in which case we must add an additional sensor for diagnosability. This section discusses strategies for adding an additional sensor for diagnosability. The following section will discuss adding an additional sensor to improve the accuracy of a system that is already diagnosable.

Let $\mathbf{\Gamma} = [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \dots, \mathbf{\Gamma}_p]$ denote the $(n + 1) \times p$ diagnostic matrix with the additional sensor added, $\tilde{\mathbf{\Gamma}} = [\tilde{\mathbf{\Gamma}}_1, \tilde{\mathbf{\Gamma}}_2, \dots, \tilde{\mathbf{\Gamma}}_p]$ denote the $n \times p$ diagnostic matrix for the original n sensors, and $\boldsymbol{\gamma}^T = [\gamma_1, \gamma_2, \dots, \gamma_p]$ denote the last row of $\mathbf{\Gamma}$. Let \mathbf{G}_n and \mathbf{G}_{n+1} denote the Gram matrices in equation (3) for the original system of n sensors and the system of $n + 1$ sensors, respectively. If the original system of n sensors is non-diagnosable, then \mathbf{G}_n is singular. Note that the upper left $p \times p$ block of the Gram matrix can be written as $(\mathbf{\Gamma}^T\mathbf{\Gamma}) \circ (\mathbf{\Gamma}^T\mathbf{\Gamma})$, where the symbol ‘ \circ ’ denotes the Hadamard (element-by-element) product of two matrices of equal dimension (Schott 1997). From the preceding definitions, it follows that $\mathbf{\Gamma}^T = [\tilde{\mathbf{\Gamma}}^T, \boldsymbol{\gamma}]$, $\mathbf{\Gamma}^T\mathbf{\Gamma} = \tilde{\mathbf{\Gamma}}^T\tilde{\mathbf{\Gamma}} + \boldsymbol{\gamma}\boldsymbol{\gamma}^T$, and it can be shown (the derivation is included in the appendix) that

$$\mathbf{G}_{n+1} = \mathbf{G}_n + \mathbf{A}, \tag{7}$$

where

$$\mathbf{A} = \begin{bmatrix} 2(\tilde{\mathbf{\Gamma}}^T\tilde{\mathbf{\Gamma}}) \circ (\boldsymbol{\gamma}\boldsymbol{\gamma}^T) + (\boldsymbol{\gamma}\boldsymbol{\gamma}^T) \circ (\boldsymbol{\gamma}\boldsymbol{\gamma}^T) & \boldsymbol{\gamma} \circ \boldsymbol{\gamma} \\ (\boldsymbol{\gamma} \circ \boldsymbol{\gamma})^T & 1 \end{bmatrix}.$$

The Gram matrix \mathbf{G}_{n+1} depends on $\boldsymbol{\gamma}$ only via \mathbf{A} , because \mathbf{G}_n does not depend on $\boldsymbol{\gamma}$. The objective is to select the additional sensor so that the resulting $\boldsymbol{\gamma}$ causes the system to be diagnosable (i.e. $\mathbf{G}_n + \mathbf{A}$ non-singular). Note that the elements of $\boldsymbol{\gamma}$ represent the effects of the p sources on the additional sensor.

Now let $\{\lambda_i, \mathbf{z}_i\}$, $i = 1, 2, \dots, p + 1$, denote the eigenvalue/eigenvector pairs of \mathbf{G}_n , arranged so that the eigenvalues are in ascending order, and let $m = p + 1 - \text{rank}(\mathbf{G}_n)$.

Thus, m is the dimension of the null space of \mathbf{G}_n , and there are exactly m null eigenvalues ($\lambda_i=0: i=1,2,\dots,m$). We refer to the associated eigenvectors $\{\mathbf{z}_i: i=1,2,\dots,m\}$ as null eigenvectors, which we take to be an orthonormal set. If γ is such that $\mathbf{z}^T \mathbf{A} \mathbf{z} = 0$ for any null eigenvector \mathbf{z} , then \mathbf{G}_{n+1} will be singular, because $\mathbf{z}^T \mathbf{G}_{n+1} \mathbf{z} = \mathbf{z}^T \mathbf{G}_n \mathbf{z} + \mathbf{z}^T \mathbf{A} \mathbf{z} = 0 + 0 = 0$. Consequently, the strategy will involve selecting the additional sensor so that $\mathbf{z}^T \mathbf{A} \mathbf{z} > 0$ for all null eigenvectors of \mathbf{G}_n .

To accomplish this, alternative expressions for $\mathbf{z}^T \mathbf{A} \mathbf{z}$ will be useful. Let $\{\eta_i, \mathbf{w}_i\}$, $i=1,2,\dots,p$, denote the eigenvalue/eigenvector pairs of $\tilde{\Gamma}^T \tilde{\Gamma}$, arranged so that the eigenvalues are in descending order, and let $r = \text{rank}(\tilde{\Gamma}^T \tilde{\Gamma})$. Furthermore, given any vector $\mathbf{z} = [\boldsymbol{\alpha}^T \ \beta]^T$ with $\boldsymbol{\alpha}$ an arbitrary p -length column vector and β an arbitrary scalar, $\mathbf{z}^T \mathbf{A} \mathbf{z}$ can be expressed as (the derivation is included in the appendix)

$$\mathbf{z}^T \mathbf{A} \mathbf{z} = \|\mathbf{h}(\mathbf{z}, \gamma)\|^2, \quad (8)$$

where the $(r+1)$ -length column vector $\mathbf{h}(\mathbf{z}, \gamma)$ is defined as a function of \mathbf{z} and γ via

$$\mathbf{h}(\mathbf{z}, \gamma) = \begin{bmatrix} \mathbf{H}(\boldsymbol{\alpha})\gamma \\ \boldsymbol{\alpha}^T(\gamma \circ \gamma) + \beta \end{bmatrix}, \quad (9)$$

and the $r \times p$ matrix $\mathbf{H}(\boldsymbol{\alpha})$ is defined as a function of $\boldsymbol{\alpha}$ via

$$\mathbf{H}(\boldsymbol{\alpha}) = \begin{bmatrix} \sqrt{2\eta_1}(\boldsymbol{\alpha} \circ \mathbf{w}_1)^T \\ \vdots \\ \sqrt{2\eta_r}(\boldsymbol{\alpha} \circ \mathbf{w}_r)^T \end{bmatrix}. \quad (10)$$

These concepts lead to the following guideline for locating an additional sensor to ensure diagnosability, the proof of which is also in the appendix. An example at the end of this section illustrates the use of the guideline.

Guideline 1: The system is diagnosable after adding an additional sensor if and only if γ is such that $\{\mathbf{h}(\mathbf{z}_i, \gamma): i=1,2,\dots,m\}$ are linearly independent, where $\{\mathbf{z}_i: i=1,2,\dots,m\}$ is the set of null eigenvectors of \mathbf{G}_n . In order to ensure this, first calculate the null eigenvectors $\{\mathbf{z}_i = [\boldsymbol{\alpha}_i^T \ \beta_i]^T: i=1,2,\dots,m\}$ of \mathbf{G}_n and the r non-null eigenvalues and eigenvectors of $\tilde{\Gamma}^T \tilde{\Gamma}$. Second, calculate numerical values for $\{\mathbf{H}(\boldsymbol{\alpha}_i): i=1,2,\dots,m\}$ and inspect the resulting expressions for $\{\mathbf{h}(\mathbf{z}_i, \gamma): i=1,2,\dots,m\}$ from (9), as a function of the to-be-determined γ , in order to identify the conditions that γ must satisfy for diagnosability. The required conditions for γ provide insight into where the sensor must be placed. For $m=1$, we only require that at least one element of $\mathbf{h}(\mathbf{z}_1, \gamma)$ differs from zero. For $m=2$, we require that $\mathbf{h}(\mathbf{z}_1, \gamma)$ and $\mathbf{h}(\mathbf{z}_2, \gamma)$ are not collinear.

For $m>2$ it may be less straightforward to use the above guidelines. It may even be impossible to find a single sensor that will result in diagnosability. However, the following guideline suggests an iterative procedure in which additional sensors are added successively until the system is diagnosable. The proof of Guideline 2 is in the appendix.

Guideline 2: Recall that m is the dimension of the null space of the $(p+1) \times (p+1)$ matrix \mathbf{G}_n and that $\text{rank}(\mathbf{G}_n) = (p+1) - m$. After adding an additional sensor with γ denoting the last row of $\tilde{\Gamma}$, we have

$$\text{rank}(\mathbf{G}_{n+1}) = \text{rank}(\mathbf{G}_n) + \text{rank}[\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)].$$

If we can find a γ such that $\mathbf{h}(\mathbf{z}_i, \gamma)$ differs from the zero vector for at least one $i \in \{1, 2, \dots, m\}$, then $\text{rank}(\mathbf{G}_{n+1})$ is strictly greater than $\text{rank}(\mathbf{G}_n)$. Thus, for each additional sensor that we add, we increase the rank of the Gram matrix by at least one (assuming at least one $\mathbf{h}(\mathbf{z}_i, \gamma)$ differs from zero) and possibly more. The effects of γ on $\{\mathbf{h}(\mathbf{z}_i, \gamma): i=1, 2, \dots, m\}$ can be ascertained as described under Guideline 1. This allows us to eventually arrive at a diagnosable sensor layout by iteratively adding sensors, provided that there exists some diagnosable layout.

Example 1: In a continuation of the example presented in section 2, suppose we originally considered only the first three variation patterns shown in figure 2 and implemented a three-sensor layout consisting of sensors 1, 2, and 4 in figure 1(c). The diagnostic matrix for this layout is

$$\mathbf{\Gamma} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

from which we can verify that the Gram matrix is non-singular. Now suppose we discover that patterns 4 and 5 in figure 2 may possibly affect the liftgate build quality and decide to include them in the model. With the original $n=3$ sensors, the diagnostic matrix and Gram matrix are

$$\tilde{\mathbf{\Gamma}} = \begin{bmatrix} -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \text{ and } \mathbf{G}_n = \begin{bmatrix} (\tilde{\mathbf{\Gamma}}^T \tilde{\mathbf{\Gamma}}) \circ (\tilde{\mathbf{\Gamma}}^T \tilde{\mathbf{\Gamma}}) & \tilde{\mathbf{\Gamma}}_1^T \tilde{\mathbf{\Gamma}}_1 & \vdots & \tilde{\mathbf{\Gamma}}_p^T \tilde{\mathbf{\Gamma}}_p \\ \tilde{\mathbf{\Gamma}}_1^T \tilde{\mathbf{\Gamma}}_1 & \cdots & \tilde{\mathbf{\Gamma}}_p^T \tilde{\mathbf{\Gamma}}_p & n \end{bmatrix} = \begin{bmatrix} 4 & 0 & 1 & 0 & 0 & 2 \\ 0 & 4 & 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 1 & 3 \end{bmatrix}.$$

The system is no longer diagnosable because the rank of \mathbf{G}_n is 4. Thus, $m=2$, and the corresponding null eigenvectors are $\mathbf{z}_1 = [0.35, 0.35, 0, 0.35, 0.35, -0.71]^T$ and $\mathbf{z}_2 = [0, 0, 0, 0.71, -0.71, 0]^T$. Furthermore, $r = \text{rank}(\tilde{\mathbf{\Gamma}}^T \tilde{\mathbf{\Gamma}}) = 3$, and the non-null eigenvalues and eigenvectors of $\tilde{\mathbf{\Gamma}}^T \tilde{\mathbf{\Gamma}}$ are $\{\eta_1, \eta_2, \eta_3\} = \{3, 2, 2\}$, $\mathbf{w}_1 = [0.58, -0.58, -0.58, 0, 0]^T$, $\mathbf{w}_2 = [0, 0, 0, 0.71, -0.71]^T$, and $\mathbf{w}_3 = [0.71, 0.71, 0, 0, 0]^T$.

Guideline 1 can be used to determine an appropriate location for adding a sensor. Since $m=2$, the requirement is that $\mathbf{h}(\mathbf{z}_1, \gamma)$ and $\mathbf{h}(\mathbf{z}_2, \gamma)$ are not collinear, where $\boldsymbol{\gamma}^T = [\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5]$ denotes the new row of $\mathbf{\Gamma}$, the elements of which depend on where the additional sensor is placed. Substituting the preceding values for the eigenvalues and eigenvectors into equation (9) gives

$$\mathbf{h}(\mathbf{z}_1, \gamma) \begin{bmatrix} 0.5\gamma_1 - 0.5\gamma_2 \\ 0.5\gamma_4 - 0.5\gamma_5 \\ 0.5\gamma_1 + 0.5\gamma_2 \\ 0.35\gamma_1^2 + 0.35\gamma_2^2 + 0.35\gamma_4^2 + 0.35\gamma_5^2 - 0.71 \end{bmatrix} \text{ and } \mathbf{h}(\mathbf{z}_2, \gamma) \begin{bmatrix} 0 \\ \gamma_4 + \gamma_5 \\ 0 \\ 0.71\gamma_4^2 - 0.71\gamma_5^2 \end{bmatrix}.$$

In order for $\mathbf{h}(\mathbf{z}_2, \gamma)$ to differ from zero, we cannot have both γ_4 and γ_5 equal to zero, which would be the case if we placed the sensor on one of the body sides (recall that γ_4 and γ_5 are the effects of pattern 4 and pattern 5 on the additional sensor).

Consequently, if the system is to be diagnosable with the additional sensor, it must be placed on the roof. For a roof sensor, $\gamma_1 = \gamma_2 = \gamma_3 = 0$, and $\gamma_4 = 1$, so that the preceding expressions reduce to

$$\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma}) = \begin{bmatrix} 0 \\ 0.5 - 0.5\gamma_5 \\ 0 \\ 0.35\gamma_5^2 - 0.35 \end{bmatrix} \quad \text{and} \quad \mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma}) = \begin{bmatrix} 0 \\ 1 + \gamma_5 \\ 0 \\ 0.71 - .71\gamma_5^2 \end{bmatrix}.$$

If the additional roof sensor is placed near sensor 5 in figure 1(c), then γ_5 will be close to 1, in which case $\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma})$ is zero. If the sensor is located near sensor 4 in figure 1(c), then γ_5 will be close to -1 , in which case $\mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma})$ is zero. Consequently, we should not place the additional roof sensor near the locations of sensors 4 and 5 in figure 1(c). This suggests placing the sensor near the middle of the roof, similar to where sensor 3 was positioned. In this case, $\gamma_5 = 0$, $\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma}) = [0, 0.5, 0, -0.35]^T$, and $\mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma}) = [0, 1, 0, 0.71]^T$. Because these two vectors are not collinear, Guideline 1 states that the system is diagnosable with this sensor layout. It can be verified that the $\mathbf{\Gamma}$ matrix for this layout (see equation (5)) results in a non-singular Gram matrix (equation (6)).

4. Adding sensors to improve accuracy

Up to this point, we have only considered diagnosability in the sensor layout guidelines. Diagnosability means only that the Gram matrix is non-singular. A particular sensor layout may result in a non-singular but poorly conditioned Gram matrix, in which case the system will be diagnosable but the estimation accuracy may be unacceptably poor.

Laying out a sensor system to satisfy some accuracy requirements is a more difficult problem. The estimation accuracy for the $(p+1)$ -length vector of parameters $\boldsymbol{\sigma} = [\sigma_1^2, \dots, \sigma_p^2, \sigma^2]^T$ depends in a complex manner on the particular set of true values for the component variances. We first consider estimation accuracy for the special case that $\sigma_1 = \sigma_2 = \dots = \sigma_p = 0$. For the WMLS algorithm in this case, which is an asymptotic MLE, the error covariance matrix for the estimate of $\boldsymbol{\sigma}$ can be approximated as (Ding *et al.* 2005)

$$\boldsymbol{\Sigma}_\sigma \approx \frac{2}{N} \sigma^4 \mathbf{G}^{-1},$$

which is proportional to the inverse of the Gram matrix. The sum of the estimation error variances for all $p+1$ components is equal to the trace of $\boldsymbol{\Sigma}_\sigma$, which is proportional to

$$\text{trace}(\mathbf{G}^{-1}) = \text{trace} \left(\left[\sum_{i=1}^{p+1} \lambda_i \mathbf{z}_i \mathbf{z}_i^T \right]^{-1} \right) = \text{trace} \left(\sum_{i=1}^{p+1} \frac{1}{\lambda_i} \mathbf{z}_i \mathbf{z}_i^T \right) = \sum_{i=1}^{p+1} \frac{1}{\lambda_i},$$

where $\{\lambda_i, \mathbf{z}_i\}$, $i=1, 2, \dots, p+1$, denote the eigenvalue/eigenvector pairs of \mathbf{G} arranged so that the eigenvalues are in ascending order. The estimation accuracy will clearly deteriorate if any of the eigenvalues is much smaller than 1 (i.e. $1/\lambda_i$ is much larger than 1).

When $\sigma_1 \cdots \sigma_p$ are not zero, the expression for Σ_σ becomes (Ding *et al.* 2005)

$$\Sigma_\sigma \approx \frac{2}{N} [\{tr(\Sigma_y^{-1} \cdot V_i \cdot \Sigma_y^{-1} \cdot V_j)\}_{i,j=1}^{p+1}]^{-1},$$

where $\{a_{ij}\}_{i,j=1}^{p+1}$ is a $(p+1) \times (p+1)$ matrix with a_{ij} as its (i, j) element, $V_i = \Gamma_i \Gamma_i^T$ for $i = 1, \dots, p$ and $V_{p+1} = \mathbf{I}$. One may verify that the matrix $\{tr(\Sigma_y^{-1} \cdot V_i \cdot \Sigma_y^{-1} \cdot V_j)\}_{i,j=1}^{p+1}$ is \mathbf{G}^* in equation (4). Because \mathbf{G} and \mathbf{G}^* are Gram matrices for the same quantities but under a slightly different inner product, one is singular or poorly conditioned if and only if the other one is (the proof of this claim is provided in appendix E). Hence, good (poor) accuracy associated with the system when $\sigma_1 = \dots = \sigma_p = 0$ will translate to good (poor) accuracy associated with the system when $\sigma_1, \dots, \sigma_p$ differ from zero. In light of this and for reasons of simplicity (e.g. to avoid having to specify specific values of $\sigma_1, \dots, \sigma_p$ on which to base the system design), we recommend selecting the additional sensors based on \mathbf{G} , rather than \mathbf{G}^* .

As in the previous section, suppose we have a current sensor layout with n sensors and want to add an additional sensor to improve accuracy. Instead of \mathbf{G}_n being singular with m null eigenvalues, we now consider the case that there are a set of m' small eigenvalues $\{\lambda_i: i = 1, 2, \dots, m'\}$, and $\{z_i: i = 1, 2, \dots, m'\}$ are the associated eigenvectors. In light of the preceding paragraph, the objective is to select the additional sensor in order to increase the small eigenvalues of $\mathbf{G}_{n+1} = \mathbf{G}_n + \mathbf{A}$. The exact relationship between γ and the eigenvalues of \mathbf{G}_n and \mathbf{G}_{n+1} is complex and inconvenient to use. However, a necessary condition to increase the eigenvalues of \mathbf{G}_{n+1} , and consequently improve the estimation accuracy, is to increase the value of $\|\mathbf{h}(z_i, \gamma)\|^2$ ($i = 1, 2, \dots, m'$). This follows by noting that the smallest eigenvalue of \mathbf{G}_{n+1} is

$$\begin{aligned} \lambda_{\min}(\mathbf{G}_{n+1}) &= \min_{\|z\|=1} z^T \mathbf{G}_{n+1} z = \min_{\|z\|=1} z^T \mathbf{G}_n z + z^T \mathbf{A} z \leq z_i^T \mathbf{G}_n z_i + z_i^T \mathbf{A} z_i \\ &= \lambda_i + \|\mathbf{h}(z_i, \gamma)\|^2: \quad i = 1, 2, \dots, m'. \end{aligned}$$

This leads to the following guideline for adding an additional sensor to improve accuracy.

Guideline 3: First, find the set of small eigenvalues $\{\lambda_i: i = 1, 2, \dots, m'\}$ and the associated eigenvectors $\{z_i\}_{i=1, \dots, m'}$. Then, select the additional sensor so that the corresponding new row γ^T of the diagnostic matrix results in values of $\lambda_i + \|\mathbf{h}(z_i, \gamma)\|^2$, $i = 1, 2, \dots, m'$, that are as large as possible.

Deciding a set of eigenvalues to be selected is a focused topic in statistical analysis such as principal component analysis (PCA). One of the widely used methods is the scree plot (i.e. a Pareto plot) suggested by Johnson and Wichern (2002, page 441), where the eigenvalues are ordered and plotted, and people will look for an elbow (bend) in the scree plot. In PCA, eigenvalues are arranged in descending order because people mean to select the group of the largest eigenvalues. Here we arrange the eigenvalues in ascending order to select the group of the smallest group. To get a quantitative sense, we also recommend calculating the ratio of λ_{i+1}/λ_i for those eigenvalues smaller than 1 (whose amplification on output variance is $1/\lambda_i > 1$), and then select the set of small eigenvalues where the largest λ_{i+1}/λ_i ratio occurs.

Example 2: Reconsider the liftgate example, and suppose we have a current sensor layout with the $n = 5$ sensors numbered 1, 2, 4, 5, and 7 in figure 1(c). Considering all five patterns shown in figure 2, the diagnostic matrix and Gram matrix for this layout are

$$\tilde{\Gamma} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \text{ and}$$

$$\mathbf{G}_n = \begin{bmatrix} \left(\tilde{\Gamma}^T \tilde{\Gamma} \right) \circ \left(\tilde{\Gamma}^T \tilde{\Gamma} \right) & \left. \begin{array}{c} \tilde{\Gamma}_1^T \tilde{\Gamma}_1 \\ \vdots \\ \tilde{\Gamma}_p^T \tilde{\Gamma}_p \end{array} \right\} \\ \hline \tilde{\Gamma}_1^T \tilde{\Gamma}_1 & \cdots & \tilde{\Gamma}_p^T \tilde{\Gamma}_p \end{bmatrix} = \begin{bmatrix} 9 & 1 & 1 & 0 & 0 & 3 \\ 1 & 9 & 1 & 0 & 0 & 3 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 4 & 2 \\ 3 & 3 & 1 & 2 & 2 & 5 \end{bmatrix}.$$

The diagonal elements of \mathbf{G}_n^{-1} (which are proportional to the variances of the estimated elements of σ) are $\{0.19, 0.19, 1.50, 0.50, 0.50, 1.00\}$. Thus, the estimates of σ_3^2 and σ^2 are the least accurate. The eigenvalues of \mathbf{G}_n are $\{0.45, 0.97, 4.00, 5.67, 8.00, 12.91\}$. Using a scree plot or the ratio of λ_{i+1}/λ_i , it is obvious that the set of small eigenvalues include the first two. Then, we have $m' = 2$, $\lambda_1 = 0.45$, and $\lambda_2 = 0.97$. The corresponding eigenvectors are $\mathbf{z}_1 = [0.12, 0.12, 0.64, 0.33, 0.33, -0.59]^T$ and $\mathbf{z}_2 = [0.22, 0.22, -0.75, 0.28, 0.28, -0.42]^T$. Furthermore, $r = \text{rank}(\tilde{\Gamma}^T \tilde{\Gamma}) = 5$, and the non-null eigenvalues and eigenvectors of $\tilde{\Gamma}^T \tilde{\Gamma}$ are $\{\eta_1, \eta_2, \eta_3, \eta_4, \eta_5\} = \{4.56, 2.00, 2.00, 2.00, 0.44\}$, $\mathbf{w}_1 = [0.66, -0.66, -0.37, 0, 0]^T$, $\mathbf{w}_2 = [0, 0, 0, 0, 1]^T$, $\mathbf{w}_3 = [0, 0, 0, 1, 0]^T$, $\mathbf{w}_4 = [0.71, 0.71, 0, 0, 0]^T$, and $\mathbf{w}_5 = [0.26, -0.26, 0.93, 0, 0]^T$. Substituting these into equation (9) gives

$$\mathbf{h}(\mathbf{z}_1, \gamma) = \begin{bmatrix} 0.23\gamma_1 - 0.23\gamma_2 - 0.72\gamma_3 \\ 0.66\gamma_5 \\ 0.66\gamma_4 \\ 0.16\gamma_1 + 0.16\gamma_2 \\ 0.03\gamma_1 - 0.03\gamma_2 + 0.55\gamma_3 \\ 0.12\gamma_1^2 + 0.12\gamma_2^2 + 0.64\gamma_3^2 + 0.33\gamma_4^2 + 0.33\gamma_5^2 - 0.59 \end{bmatrix},$$

$$\mathbf{h}(\mathbf{z}_2, \gamma) = \begin{bmatrix} 0.44\gamma_1 - 0.44\gamma_2 + 0.84\gamma_3 \\ 0.56\gamma_5 \\ 0.56\gamma_4 \\ 0.32\gamma_1 + 0.32\gamma_2 \\ 0.05\gamma_1 - 0.05\gamma_2 - 0.65\gamma_3 \\ 0.22\gamma_1^2 + 0.22\gamma_2^2 - 0.75\gamma_3^2 + 0.28\gamma_4^2 + 0.28\gamma_5^2 - 0.42 \end{bmatrix}.$$

In order to decide which pattern should have most effect on the additional sensor, we recommend doing so in accordance with the element of γ that has the largest sum

of squares of coefficients. Here, the sum of squares of coefficients of γ_3 is clearly the largest, suggesting that pattern 3 should have as large an effect as possible on the additional sensor. This is accomplished by placing the additional sensor as high up on the body side as possible. If we choose a location that corresponds to sensor 6 in figure 1(c), then $\gamma = [1, 1, 1, 0, 0]^T$, and the result is $\|\mathbf{h}(\mathbf{z}_1, \gamma)\| = 1.02$ and $\|\mathbf{h}(\mathbf{z}_2, \gamma)\|^2 = 2.06$. With this value of γ , it can be verified that the Gram matrix \mathbf{G}_{n+1} after adding the additional sensor is

$$\mathbf{G}_{n+1} = \begin{bmatrix} 16 & 0 & 0 & 0 & 0 & 4 \\ 0 & 16 & 4 & 0 & 0 & 4 \\ 0 & 4 & 4 & 0 & 0 & 2 \\ 0 & 0 & 0 & 4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 4 & 2 \\ 4 & 4 & 2 & 2 & 2 & 6 \end{bmatrix}.$$

The diagonal elements of \mathbf{G}_{n+1}^{-1} are $\{0.10, 0.10, 0.40, 0.40, 0.40, 0.60\}$, compared to the diagonal elements $\{0.19, 0.19, 1.50, 0.50, 0.50, 1.00\}$ of \mathbf{G}_n^{-1} discussed earlier. Thus, the additional sensor reduces the estimation error variances for σ_3^2 and σ^2 by 73% and 40%, respectively.

5. Sensor placement in a three-station assembly process

The examples in the previous sections were rather simple ones in which the $\mathbf{\Gamma}$ matrix can be readily determined given the sensor locations, and all elements of $\mathbf{\Gamma}$ are either ± 1 or 0. In this case, a numerical optimization algorithm is largely unnecessary. Actual assembly processes are often far more complicated, however, consisting of multiple assembly stations and sequences of operations. In this section we apply the guidelines from the previous sections in a three-station panel assembly example that, although still relatively simple, is not as transparent as the previous examples. In situations like this, it might be advantageous to use the guidelines in conjunction with numerical optimization. The example demonstrates the use of the guidelines and how they can facilitate a numerical search.

5.1 The assembly process

The three-station assembly process shown in figure 3, which was also considered in Ding *et al.* (2002b), Zhou *et al.* (2003a), and Apley and Ding (2005), represents a segment of an automotive body assembly with the geometry of each part simplified to a rectangle. The assembly process welds four parts together in two stations (Stations I and II). In Station I, Parts 1 and 2 are joined, and the resulting sub-assembly is joined with Parts 3 and 4 in Station II. In an assembly station, each part (or subassembly) is located in a fixture using a pin that mates with a hole in the part and second pin that mates with a slot in the part. A pin/hole combination constrains two degrees-of-freedom and a pin/slot constrains only one degree-of-freedom. Together, a pin/hole and pin/slot completely constrain all three degrees-of-freedom of the part within the x - z plane. The active holes and slots at each station are shown darkened. Holes and slots that are not darkened are not used in that

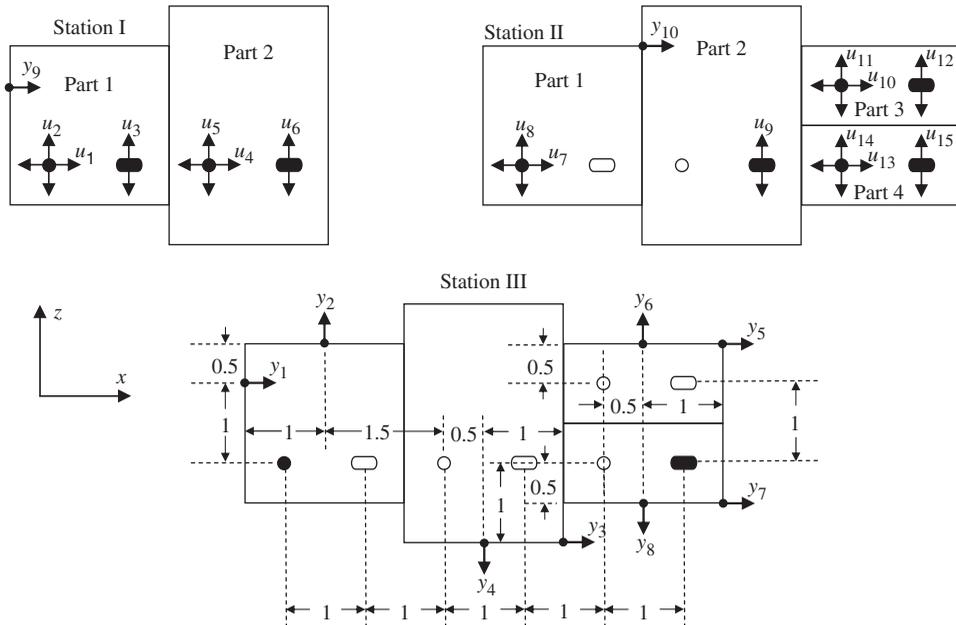


Figure 3. A three station assembly system where Parts 1 and 2 are joined in Station I; the Part 1–2 subassembly is joined with Parts 3 and 4 in Station II; and the final assembly is measured in Station III.

particular station. Station III is a measurement station in which no assembly takes place, although measurement is not restricted to Station III. The dimensions shown in the Station III figure have been scaled for convenience, and have no physical meaning.

The variation sources that we will consider are deviations of the pin/hole combinations in the x - and z -directions and deviations of the pin/slot combinations in the z -direction for the two assembly stations. Because there are a total of five active pin/hole combinations and five active pin/slot combinations in Stations I and II, there are a total of 15 potential variation sources ($p = 15$). These are indicated by the arrows labelled u_1 through u_{15} in the Station I and Station II figures. Deviations in the positive x - and positive z -directions are taken to be positive.

In order to estimate the variance components associated with the 15 variation sources, suppose that a total of eight sensors ($n = 8$) are installed in Station III, with two sensors (one measuring the x -deviation and one measuring the z -deviation) placed on each of the four parts. The locations of the eight sensors are labelled y_1 through y_8 , where the direction of the arrows indicates whether the x -coordinate or z -coordinate is being measured (temporarily ignore the additional sensors denoted by y_9 and y_{10}).

For this initial sensor layout, we first check diagnosability to verify whether it is possible to estimate $p + 1 = 16$ variance components (including that of sensor noise) using $n = 8$ sensors. The Γ matrix can be calculated using a simple computer program based on any of the systematic multi-station modelling procedures described

in Carlson *et al.* (2000), Jin and Shi (1999), or Ding *et al.* (2000), which yields (the bar overscore indicates a repeating digit):

$$\Gamma = \begin{bmatrix} 0 & 0.\bar{6} & -1 & 0 & 0 & 0.\bar{3} & 0 & 0.\bar{13} & -0.\bar{3} & 0 & 0 & 0 & 0 & 0 & 0.2 \\ 0 & -0.\bar{3} & 0.5 & 0 & 0 & -0.\bar{16} & 0 & -0.\bar{06} & 0.\bar{16} & 0 & 0 & 0 & 0 & 0 & -0.1 \\ -1 & 0.\bar{3} & 0 & 1 & -1 & 0.\bar{6} & 0 & -0.\bar{13} & 0.\bar{3} & 0 & 0 & 0 & 0 & 0 & -0.2 \\ 0 & -0.\bar{16} & 0 & 0 & 0.5 & -0.\bar{3} & 0 & -0.\bar{3} & 0.\bar{83} & 0 & 0 & 0 & 0 & 0 & -0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -0.3 & 0 & 1 & 0.5 & -0.5 & 0 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & -0.9 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0.1 & 0 & 0 & 0 & 0 & 1 & -0.5 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1 & 0 & 0 & 0 & 0 & 0 & 0.5 & -0.4 \end{bmatrix}.$$

It can easily be verified that the resulting Gram matrix for the above Γ is singular with rank 15, so that the system is not diagnosable.

5.2 Ensuring diagnosability

To ensure diagnosability, we need to place additional sensors. Because $m = 1$ in this case, Guideline 1 implies that we only need $h(z_1, \gamma)$ to differ from the zero vector, where γ is the 15×1 vector that represents the effects of the 15 variation sources on the additional sensor. The vector $h(z_1, \gamma)$, which can be calculated using equations (9) and (10) as described in the previous examples, is

$$h(z_1, \gamma) = \begin{bmatrix} -0.0128\gamma_1 - 0.0128\gamma_4 \\ 0.101\gamma_1 + 0.101\gamma_4 \\ -0.0333\gamma_1 - 0.0333\gamma_4 \\ 0.0988\gamma_1 + 0.0988\gamma_4 \\ 0.2882\gamma_1 + 0.2882\gamma_4 \\ 0.3139\gamma_1 + 0.3139\gamma_4 \\ -0.8928\gamma_1 - 0.8928\gamma_4 \\ 0.7071\gamma_1^2 - 0.7071\gamma_4^2 \end{bmatrix}.$$

The insight gained from this is that we need to place the additional sensor so that $\gamma_1 \neq -\gamma_4$. Recall that γ_1 and γ_4 represent the effects of the variation sources u_1 and u_4 on the additional sensor. Because u_1 and u_4 are associated with x -direction displacement of Parts 1 and 2, respectively, in Station I, it does not help to place the sensor on Part 3 or 4 in Station II or III. Thus, the additional sensor must measure the x -direction on either Part 1 or Part 2.

Because Parts 1 and 2 appear in all three stations, we need to further decide in which station to place the sensor. If the sensor is placed in Station II or III, then we will have $\gamma_1 = -\gamma_4$. This is because in Stations II and III, Parts 1 and 2 have already been joined, at which time a positive deviation of u_1 becomes indistinguishable from a negative deviation of u_4 . Consequently, the only choice is to place the additional sensor in Station I, measuring either the x -direction of Part 1 or the x -direction of Part 2. If the sensor is placed on Part 1 then $\gamma_1 = 1$ and $\gamma_4 = 0$. If the sensor is placed on Part 2 then $\gamma_1 = 0$ and $\gamma_4 = 1$. Either placement results in a diagnosable sensor layout according to Guideline 1.

Suppose that we place the additional sensor at y_9 on Part 1, which is the same position as y_1 but in Station I. It can be verified that the Gram matrix is now full rank.

In fact, after adding y_9 , we can even remove y_1 and still retain diagnosability with an eight-sensor system (y_2 to y_9). After adding a sensor, we recommend iteratively removing each column of Γ and checking whether the resulting Gram matrix is still full-rank, in order to determine whether any of the original sensors can be removed (or perhaps relocated to a more advantageous position). This can be checked quickly and easily with a computer program.

5.3 Improving accuracy

Starting with the diagnosable eight-sensor system (y_2 to y_9) from the previous step, the next step is to determine whether accuracy is sufficient. With sensors y_2 to y_9 , we have

$$\tilde{\Gamma} = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.\bar{3} & 0.5 & 0 & 0 & -0.1\bar{6} & 0 & -0.0\bar{6} & 0.1\bar{6} & 0 & 0 & 0 & 0 & 0 & -0.1 \\ -1 & 0.\bar{3} & 0 & 1 & -1 & 0.\bar{6} & 0 & -0.1\bar{3} & 0.\bar{3} & 0 & 0 & 0 & 0 & 0 & -0.2 \\ 0 & -0.1\bar{6} & 0 & 0 & 0.5 & -0.\bar{3} & 0 & -0.\bar{3} & 0.8\bar{3} & 0 & 0 & 0 & 0 & 0 & -0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -0.3 & 0 & 1 & 0.5 & -0.5 & 0 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & -0.9 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0.1 & 0 & 0 & 0 & 0 & 1 & -0.5 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.1 & 0 & 0 & 0 & 0 & 0 & 0.5 & -0.4 \end{bmatrix}.$$

The diagonal elements of \mathbf{G}_n^{-1} are $\{0.68, 4.75, 5.79, 3.58, 8.59, 35.21, 0.55, 30.31, 3.09, 3.61, 9.42, 12.43, 2.15, 6.75, 0.81, 0.61\}$, which indicate that the estimates of σ_6^2 and σ_8^2 are the least accurate. The eigenvalues of \mathbf{G}_n are $\{0.023, 0.031, 0.050, 0.103, 0.163, 0.221, 0.234, 0.310, 0.629, 1.109, 1.347, 1.916, 2.626, 2.932, 5.020, 11.132\}$, from which one can observe that there are a few small eigenvalues.

Same as in Example 2, calculating the ratio of λ_{i+1}/λ_i or using the scree plot, we find that the set of small eigenvalues includes the first three eigenvalues (the largest $\lambda_{i+1}/\lambda_i = 2.06$ happens at $i = 3$), in which case we set $m' = 3$. In order to use Guideline 3, we first use equations (9) and (10) to calculate

$$\mathbf{h}(z_1, \gamma) = \begin{bmatrix} 0.0987\gamma_5 + 0.1578\gamma_6 \\ -0.0964\gamma_5 - 0.1528\gamma_6 \\ -0.1885\gamma_2 - 0.1637\gamma_3 + 0.3575\gamma_5 + 0.596\gamma_6 \\ 0.1560\gamma_5 + 0.2336\gamma_6 \\ -0.0994\gamma_3 - 0.4733\gamma_5 - 0.6834\gamma_6 \\ -0.1239\gamma_2^2 + 0.1259\gamma_3^2 - 0.4028\gamma_5^2 + 0.8961\gamma_6^2 \end{bmatrix},$$

$$\mathbf{h}(z_2, \gamma) = \begin{bmatrix} -0.2149\gamma_8 - 0.1244\gamma_9 \\ 0.1107\gamma_8 + 0.1213\gamma_{10} + 0.1048\gamma_{12} \\ 0.6086\gamma_8 + 0.1039\gamma_9 + 0.1133\gamma_{10} \\ 0.1117\gamma_8 + 0.1578\gamma_{10} + 0.1084\gamma_{12} \\ -0.1254\gamma_8 \\ -0.9578\gamma_8^2 + 0.1417\gamma_9^2 + 0.1645\gamma_{10}^2 - 0.1756\gamma_{12}^2 \end{bmatrix},$$

$$\mathbf{h}(\mathbf{z}_3, \boldsymbol{\gamma}) = \begin{bmatrix} 0.1297\gamma_{11} - 0.1486\gamma_{14} \\ -0.1628\gamma_9 + 0.2996\gamma_{11} + 0.3173\gamma_{12} \\ -0.1607\gamma_{10} - 0.4087\gamma_{12} + 0.1448\gamma_{14} \\ 0.136\gamma_9 - 0.1501\gamma_{10} + 0.42\gamma_{11} + 0.1129\gamma_{15} \\ -0.2091\gamma_{10} + 0.1965\gamma_{11} - 0.423\gamma_{12} - 0.1449\gamma_{14} - 0.0992\gamma_{15} \\ -0.1157\gamma_{12} \\ -0.1331\gamma_8^2 + 0.1855\gamma_9^2 - 0.2179\gamma_{10}^2 + 0.5743\gamma_{11}^2 \\ +0.6851\gamma_{12}^2 + 0.2683\gamma_{14}^2 + -0.0971\gamma_{15}^2 \end{bmatrix}.$$

Note that we have cleaned up the expressions for these three \mathbf{h} vectors by ignoring terms with very small coefficients (our threshold was 10% of the largest coefficient, or 0.0958). The immediate goal is to make the norm of these three \mathbf{h} vectors as large as possible.

The above expressions indicate that the major contributors to $\|\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma})\|$ are γ_5 and γ_6 (again based on the magnitude of their coefficients), the major contributors to $\|\mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma})\|$ are γ_8 and γ_9 , and the major contributors to $\|\mathbf{h}(\mathbf{z}_3, \boldsymbol{\gamma})\|$ are γ_{11} and γ_{12} . By inspection of figure 3, it is not possible to find a single sensor that will simultaneously make $\|\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma})\|$, $\|\mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma})\|$, and $\|\mathbf{h}(\mathbf{z}_3, \boldsymbol{\gamma})\|$ large. The reason is that in order for γ_{11} and γ_{12} to differ from zero, the sensor must be located on Part 3. Any sensor located on Part 3, however, will not be strongly affected by u_5, u_6, u_8 , or u_9 (i.e. will not result in large values for $\gamma_5, \gamma_6, \gamma_8$, and γ_9). On the other hand, figure 3 indicates that it may be possible to make $\gamma_5, \gamma_6, \gamma_8$, and γ_9 large (thereby making $\|\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma})\|$ and $\|\mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma})\|$ larger simultaneously) with a single additional sensor. This suggests the following strategy for adding *two* additional sensors in order to increase the norm of the three \mathbf{h} vectors. We locate the first sensor to make $\gamma_5, \gamma_6, \gamma_8$, and γ_9 as large as possible, and we locate the second sensor to make γ_{11} and γ_{12} as large as possible.

In order to make $\gamma_5, \gamma_6, \gamma_8$, and γ_9 large, the first additional sensor should be placed on Parts 1 or 2. Because Parts 1 and 2 appear on all three stations, we also need to decide in which station to place the sensor. We cannot choose Station I, because u_8 and u_9 would have no effect on the sensor (i.e. $\gamma_8 = \gamma_9 = 0$). Moreover, fixture layouts are usually designed so that the effects of faults that occur in one station are not amplified in subsequent downstream stations. Hence, we can most likely make $\gamma_5, \gamma_6, \gamma_8$, and γ_9 larger by placing the first additional sensor in Station II, rather than the downstream Station III.

Consequently, we restrict our search to Parts 1 or 2 on Station II. It is not difficult to see that placing a sensor further away from the locating point generally results in a larger γ . On a rectangular part, this translates to placing the additional sensor at a corner. We can therefore narrow down the set of candidate sensor locations to 16 possibilities (either an x - or a z -direction sensor at one of the eight corners on Parts 1 and 2).

Calculating the resulting \mathbf{G}_{n+1} matrix for a total of 16 different sensor locations is relatively easy with the aid of a numerical program in (for example) MATLAB. This would reveal that placing an additional x -direction sensor on the top-right corner of Part 1 in Station II (i.e., γ_{10} in figure 3) maximizes the smallest eigenvalue of \mathbf{G}_{n+1} among the 16 candidate locations. For the resulting nine-sensor system

(y_2 to y_{10}), the smallest eigenvalue of \mathbf{G}_{n+1} is 0.0584, which is 2.5 times larger than the smallest eigenvalue of \mathbf{G}_n . The diagonal elements of \mathbf{G}_{n+1}^{-1} are $\{0.3842, 1.4438, 0.5864, 3.2294, 2.6017, 3.6385, 0.1688, 8.7818, 1.7787, 2.6091, 8.3464, 10.8194, 1.6370, 6.2866, 0.7517, 0.3734\}$. Comparing this with the diagonal elements of \mathbf{G}_n^{-1} , the additional sensor reduces the estimation error variances for σ_6^2 and σ_8^2 by 95.1% and 71.0%, respectively. After adding y_{10} , we could then focus on making $\|\mathbf{h}(z_3, \gamma)\|$ large (via making γ_{11} and γ_{12} large) by placing the second additional sensor on Part 3 in Station III. Alternatively, we could start anew after y_{10} is added by recalculating the new \mathbf{h} vector(s) corresponding to any small eigenvalue(s) of the new \mathbf{G}_{n+1} .

5.4 Comparison with numerical search

Others (e.g. Khan *et al.* 1998, 1999, 2000, Wang and Nagarkar, 1999, Ding *et al.* 2003, Liu *et al.* 2005) have used numerical search methods to find the additional sensor location that maximizes the minimum eigenvalue of \mathbf{G}_{n+1} (equivalent to E-optimality) over the feasible design space. For this three-station problem, the feasible design space might be the two-dimensional space over the panel surfaces of each part in each station. The pool of possible candidate sensor locations is generated by discretising each panel with a grid of resolution 0.01. With this resolution, Part 2, for example, has 60 000 candidate locations for either an x -direction or a z -direction sensor. The best single sensor location found by an exhaustive search turned out to coincide with the one recommended above based on the guidelines, namely y_{10} shown in figure 3. Consequently, one might use the guidelines as a method of substantially narrowing the search space, in order to reduce computational expense and improve the robustness of a numerical optimization procedure.

6. Conclusions

This paper has developed guidelines that aid in placing additional sensor(s) in an existing sensor layout, in order to ensure diagnosability and/or improve variance components estimation accuracy. The guidelines were illustrated with examples from autobody assembly. In the three-station example, the recommended sensor location using the guidelines turned out to coincide with the ‘optimal’ location produced by an exhaustive search, and the additional sensor substantially improved the estimation accuracy of the variance components.

In more complex manufacturing systems, the sensor placement guidelines could be used in conjunction with the numerical optimization strategies of (for example) Wang and Nagarkar (1999), Khan and Ceglarek (2000), or Liu *et al.* (2005), which search over all possible candidate sensor layouts. The guidelines could be used to judiciously select the set of candidate layouts over which the optimization algorithm must search. Alternatively, the guidelines could be used to provide a set of ‘good’ sensor locations that the optimization routine could fine-tune by searching only locally, in the neighbourhoods of the candidate locations.

There are also situations in which one might wish to use the guidelines by themselves, in lieu of a numerical search routine. There are typically many engineering constraints on where sensors can be located and what features can be measured. For example, if the panels in figure 3 truly were perfectly flat with no distinguishing

features, then the only way a sensor could measure an x - or z -direction displacement would be to locate it on an edge or on one of the holes/slots. Furthermore, some edges (either abrupt sheared edges or contoured flanges) are not formed with enough precision to yield an accurate measurement of the panel displacement. Although a process engineer may have a clear conception of where sensors can or cannot be located, it may be quite difficult to meticulously encode all of these constraints into an optimization algorithm. The fact that this is not a black and white issue further complicates the matter. A sensor located on one panel feature may yield a panel displacement measurement that provides some information, but that is not as accurate as if the sensor were located on a different feature. When using the guidelines of this paper to select a good sensor location, a process engineer could easily take into account a myriad of other qualitative factors based on experience and engineering judgment that would be otherwise difficult to incorporate quantitatively into an optimization routine.

Acknowledgements

The authors gratefully acknowledge financial support from the NSF grants DMI-0354824, DMI-0217481, DMI-0348150, and from the State of Texas Advanced Technology Program under grant 000512-0237-2003. The authors also appreciate the editor and referees for their valuable comments and suggestions.

Appendices (proofs of various results)

Appendix A. Derivation of equation (7)

$$\begin{aligned}
 \mathbf{G}_{n+1} &= \left[\begin{array}{c|c} & \Gamma_1^T \Gamma_1 \\ \hline (\Gamma^T \Gamma) \circ (\Gamma^T \Gamma) & \vdots \\ & \Gamma_p^T \Gamma_p \\ \hline \Gamma_1^T \Gamma_1 & \cdots & \Gamma_p^T \Gamma_p & | & n+1 \end{array} \right] \\
 &= \left[\begin{array}{c|c} & \tilde{\Gamma}_1^T \tilde{\Gamma}_1 + \gamma_1^2 \\ \hline (\tilde{\Gamma}^T \tilde{\Gamma} + \gamma\gamma^T) \circ (\tilde{\Gamma}^T \tilde{\Gamma} + \gamma\gamma^T) & \vdots \\ & \tilde{\Gamma}_p^T \tilde{\Gamma}_p + \gamma_p^2 \\ \hline \tilde{\Gamma}_1^T \tilde{\Gamma}_1 + \gamma_1^2 & \cdots & \tilde{\Gamma}_p^T \tilde{\Gamma}_p + \gamma_p^2 & | & n+1 \end{array} \right] \\
 &= \left[\begin{array}{c|c} & \tilde{\Gamma}_1^T \tilde{\Gamma}_1 + \gamma_1^2 \\ \hline (\tilde{\Gamma}^T \tilde{\Gamma}) \circ (\tilde{\Gamma}^T \tilde{\Gamma}) + 2(\tilde{\Gamma}^T \tilde{\Gamma}) \circ (\gamma\gamma^T) + (\gamma\gamma^T) \circ (\gamma\gamma^T) & \vdots \\ & \tilde{\Gamma}_p^T \tilde{\Gamma}_p + \gamma_p^2 \\ \hline \tilde{\Gamma}_1^T \tilde{\Gamma}_1 + \gamma_1^2 & \cdots & \tilde{\Gamma}_p^T \tilde{\Gamma}_p + \gamma_p^2 & | & n+1 \end{array} \right] \\
 &= \mathbf{G}_n + \mathbf{A}.
 \end{aligned}$$

Appendix B. Derivation of equation (8)

Given that rank $(\tilde{\Gamma}^T \tilde{\Gamma})$ is r , it follows that $\eta_i = 0$ for $i = r + 1, r + 2, \dots, p$, and inserting the spectral decomposition $\tilde{\Gamma}^T \tilde{\Gamma} = \sum_{i=1}^r \eta_i \mathbf{w}_i \mathbf{w}_i^T$ into the expression of \mathbf{A} in equation (7) gives

$$\mathbf{A} = \begin{bmatrix} 2\sum_{i=1}^r \eta_i (\mathbf{w}_i \circ \boldsymbol{\gamma})(\mathbf{w}_i \circ \boldsymbol{\gamma})^T + (\boldsymbol{\gamma} \circ \boldsymbol{\gamma})(\boldsymbol{\gamma} \circ \boldsymbol{\gamma})^T & \boldsymbol{\gamma} \circ \boldsymbol{\gamma} \\ (\boldsymbol{\gamma} \circ \boldsymbol{\gamma})^T & 1 \end{bmatrix}, \tag{B1}$$

where we have used the fact that $(\mathbf{a}\mathbf{a}^T) \circ (\mathbf{b}\mathbf{b}^T) = (\mathbf{a} \circ \mathbf{b})(\mathbf{a} \circ \mathbf{b})^T$ for any vectors \mathbf{a} and \mathbf{b} of equal dimension. With $\mathbf{z} = [\boldsymbol{\alpha}^T \ \beta]^T$ and \mathbf{A} as in the foregoing expression,

$$\begin{aligned} \mathbf{z}^T \mathbf{A} \mathbf{z} &= 2\boldsymbol{\alpha}^T \sum_{i=1}^r \eta_i (\mathbf{w}_i \circ \boldsymbol{\gamma})(\mathbf{w}_i \circ \boldsymbol{\gamma})^T \boldsymbol{\alpha} + (\boldsymbol{\alpha}^T (\boldsymbol{\gamma} \circ \boldsymbol{\gamma}))^2 + 2\beta \boldsymbol{\alpha}^T (\boldsymbol{\gamma} \circ \boldsymbol{\gamma}) + \beta^2 \\ &= 2 \sum_{i=1}^r \eta_i [\boldsymbol{\alpha}^T (\mathbf{w}_i \circ \boldsymbol{\gamma})]^2 + [\boldsymbol{\alpha}^T (\boldsymbol{\gamma} \circ \boldsymbol{\gamma}) + \beta]^2 \\ &= \sum_{i=1}^r [\sqrt{2\eta_i} (\boldsymbol{\alpha} \circ \mathbf{w}_i)^T \circ \boldsymbol{\gamma}]^2 + [\boldsymbol{\alpha}^T (\boldsymbol{\gamma} \circ \boldsymbol{\gamma}) + \beta]^2 \\ &= \|\mathbf{H}(\boldsymbol{\alpha})\boldsymbol{\gamma}\|^2 + [\boldsymbol{\alpha}^T (\boldsymbol{\gamma} \circ \boldsymbol{\gamma}) + \beta]^2 = \|\mathbf{h}(\mathbf{z}, \boldsymbol{\gamma})\|^2. \end{aligned} \quad \square$$

Appendix C. Proof of Guideline 1

The system is nondiagnosable (\mathbf{G}_{n+1} is singular) iff there exists a nontrivial vector \mathbf{z} such that $\mathbf{z}^T \mathbf{G}_{n+1} \mathbf{z} = \mathbf{z}^T \mathbf{G}_n \mathbf{z} + \mathbf{z}^T \mathbf{A} \mathbf{z} = 0$. Both \mathbf{G}_n and \mathbf{A} are positive semi-definite (\mathbf{G}_n because it is a Gram matrix, and \mathbf{A} because of equation (8)). Thus, $\mathbf{z}^T \mathbf{G}_{n+1} \mathbf{z} = 0$ iff $\mathbf{z}^T \mathbf{G}_n \mathbf{z}$ and $\mathbf{z}^T \mathbf{A} \mathbf{z}$ are both zero. Now $\mathbf{z}^T \mathbf{G}_n \mathbf{z}$ is zero iff \mathbf{z} lies in the null space of \mathbf{G}_n , and the null space of \mathbf{G}_n is the span of its null eigenvectors. It follows that the system is non-diagnosable iff there exists a set of coefficients $\{\delta_i; i = 1, 2, \dots, m\}$, not all zero, such that

$$\begin{aligned} 0 &= (\delta_1 \mathbf{z}_1 + \delta_2 \mathbf{z}_2 + \dots + \delta_m \mathbf{z}_m)^T \mathbf{A} (\delta_1 \mathbf{z}_1 + \delta_2 \mathbf{z}_2 + \dots + \delta_m \mathbf{z}_m) \\ &= \|\mathbf{h}(\delta_1 \mathbf{z}_1 + \delta_2 \mathbf{z}_2 + \dots + \delta_m \mathbf{z}_m, \boldsymbol{\gamma})\|^2 \\ &= \|\mathbf{h}(\mathbf{z}_1, \boldsymbol{\gamma})\delta_1 + \mathbf{h}(\mathbf{z}_2, \boldsymbol{\gamma})\delta_2 + \dots + \mathbf{h}(\mathbf{z}_m, \boldsymbol{\gamma})\delta_m\|^2, \end{aligned}$$

where the second equality follows from equation (8), and the last equality follows from the linearity of $\mathbf{h}(\mathbf{z}, \boldsymbol{\gamma})$ with respect to \mathbf{z} . There exists a set of coefficients such that preceding equation holds iff $\{\mathbf{h}(\mathbf{z}_i, \boldsymbol{\gamma}); i = 1, 2, \dots, m\}$ are linearly dependent, which completes the proof. □

Appendix D. Proof of Guideline 2

Let $\mathcal{N}(\bullet)$ denote the null space of a matrix and $\dim(\bullet)$ the dimension of a linear vector space. Because $\mathbf{G}_{n+1} = \mathbf{G}_n + \mathbf{A}$, and \mathbf{G}_n and \mathbf{A} are both positive semi-definite, $\mathcal{N}(\mathbf{G}_{n+1})$ is the intersection of $\mathcal{N}(\mathbf{G}_n)$ and $\mathcal{N}(\mathbf{A})$. Thus, $\mathcal{N}(\mathbf{G}_{n+1}) = \{\mathbf{z} \in \text{span}\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m\}; \mathbf{z}^T \mathbf{A} \mathbf{z} = 0\}$. As in the proof of Guideline 1, consider

linear combinations of the form $\mathbf{z} = \delta_1 \mathbf{z}_1 + \delta_2 \mathbf{z}_2 + \dots + \delta_m \mathbf{z}_m = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m] \boldsymbol{\delta}$ with $\boldsymbol{\delta} = [\delta_1, \delta_2, \dots, \delta_m]^T$. For any such \mathbf{z} , it also follows from the proof of Guideline 1 that $\mathbf{z}^T \mathbf{A} \mathbf{z} = 0$ iff $[\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)] \boldsymbol{\delta} = \mathbf{0}$. Thus, $\mathcal{N}(\mathbf{G}_{n+1}) = \{[\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m] \boldsymbol{\delta} : [\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)] \boldsymbol{\delta} = \mathbf{0}\}$, the dimension of which is equal to the dimension of $\mathcal{N}([\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)])$. Because the rank of any matrix is equal to the number of columns minus the dimension of its null space, it follows that

$$\begin{aligned} \text{rank}(\mathbf{G}_{n+1}) &= (p+1) - \dim(\mathcal{N}(\mathbf{G}_{n+1})) \\ &= (p+1) - \dim(\mathcal{N}([\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)])) \\ &= (p+1) - (m - \text{rank}[\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)]) \\ &= \text{rank}(\mathbf{G}_n) + \text{rank}[\mathbf{h}(\mathbf{z}_1, \gamma), \mathbf{h}(\mathbf{z}_2, \gamma), \dots, \mathbf{h}(\mathbf{z}_m, \gamma)]. \quad \square \end{aligned}$$

Appendix E. Proof of claim related to estimation accuracy

The following is a proof that if the system is designed to have good accuracy in the simplifying scenario where $\sigma_1 = \dots = \sigma_p = 0$, then it will also have good accuracy in the more general scenario where $\sigma_1, \dots, \sigma_p$ are not all zero. Consider the set of $p+1$ matrices $\{\{\boldsymbol{\Gamma}_i \boldsymbol{\Gamma}_i^T : i=1, 2, \dots, p\}, \mathbf{I}\}$. \mathbf{G} is the Gram matrix for the set $\{\{\boldsymbol{\Gamma}_i \boldsymbol{\Gamma}_i^T : i=1, 2, \dots, p\}, \mathbf{I}\}$ under the matrix inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{A} \mathbf{B}^T)$. Similarly, \mathbf{G}^* is the Gram matrix for the *same* set $\{\{\boldsymbol{\Gamma}_i \boldsymbol{\Gamma}_i^T : i=1, 2, \dots, p\}, \mathbf{I}\}$ under the matrix inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\boldsymbol{\Sigma}_y^{-1/2} \mathbf{A} \boldsymbol{\Sigma}_y^{-1} \mathbf{B}^T \boldsymbol{\Sigma}_y^{-1/2})$, where $\boldsymbol{\Sigma}_y = \sum_{i=1}^p \sigma_i^2 \boldsymbol{\Gamma}_i \boldsymbol{\Gamma}_i^T + \sigma^2 \mathbf{I}$. A Gram matrix for a set is singular if and only if the members of the set are linearly dependent. Because this does not depend on which inner product is chosen to define the Gram matrix, it follows that \mathbf{G}^* is singular if and only if \mathbf{G} is singular. One may use similar, but more tedious, arguments to show that \mathbf{G}^* is poorly conditioned if and only if \mathbf{G} is poorly conditioned. This claim also relies on the observation that $\boldsymbol{\Sigma}_y$ is well-conditioned, even if $\boldsymbol{\Gamma}$ is close to singular, because of the $\sigma^2 \mathbf{I}$ term (the eigenvalues of the positive definite matrix $\boldsymbol{\Sigma}_y$ are bounded below by σ^2). \square

References

- Anderson, T.W., Asymptotic efficient estimation of covariance matrices with linear structure. *Annals of Statistics*, 1973, **1**, 135–141.
- Apley, D.W. and Lee, H.Y., Identifying spatial variation patterns in multivariate manufacturing processes: a blind separation approach. *Technometrics*, 2003, **45**, 187–198.
- Apley, D.W. and Ding, Y., A characterisation of diagnosability conditions for variance components analysis in assembly operations. *IEEE Trans. Autom. Sci. Eng.*, 2005, **2**, 101–110.
- Apley, D.W. and Shi, J., A factor-analysis method for diagnosing variability in multivariate manufacturing processes. *Technometrics*, 2001, **43**, 84–95.
- Apley, D.W. and Shi, J., Diagnosis of multiple fixture faults in panel assembly. *ASME. J. Manuf. Sci. Eng.*, 1998, **120**, 793–801.
- Atkinson, A.C. and Donev, A.N., *Optimum Experimental Designs*, 1992 (Oxford University Press: New York, NY).
- Barton, R.R. and Gonzalez-Barreto, D.R., Process-oriented basis representations for multivariate process diagnosis. *Qual. Eng.*, 1996, **9**, 107–118.

- Bohme, J.F., Estimation of special parameters of correlated signals in wavefields. *Signal Proc.*, 1986, **11**, 329–337.
- Camelio, A.J., Hu, S.J. and Ceglarek, D.J., Modeling variation propagation of multi-station assembly systems with compliant parts. *ASME J. Mech. Des.*, 2003, **125**, 673–681.
- Carlson, J.S., Lindkvist, L. and Soderberg, R., Multi-fixture assembly system diagnosis based on part and subassembly measurement data, in *Proceedings of the 2000 ASME Design Engineering Technical Conference*, Baltimore, MD, 10–13 September 2000.
- Ceglarek, D. and Shi, J., Dimensional variation reduction for automotive body assembly. *Manuf. Rev.*, 1995, **8**, 139–154.
- Ceglarek, D. and Shi, J., Fixture failure diagnosis for the autobody assembly using pattern recognition. *ASME J. Eng. for Indust.*, 1996, **118**, 55–66.
- Ceglarek, D., Shi, J. and Wu, S.M., A knowledge-based diagnostic approach for the launch of the auto-body assembly process. *ASME J. Eng. Indust.*, 1994, **116**, 491–499.
- Chang, M. and Gossard, D.C., Computational method for diagnosis of variation-related assembly problem. *Int. J. Prod. Res.*, 1998, **36**, 2985–2995.
- D'Assumpcao, H.A., Some new signal processors for arrays of sensors. *IEEE Trans. Inform. Theor.* 1980, **26**, 441–453.
- Ding, Y., Ceglarek, D. and Shi, J., Modeling and diagnosis of multi-station manufacturing processes: Part I. State space model, in *Proceedings of the 2000 Japan/USA Symposium on Flexible Automation*, Ann Arbor, MI, 23–26 July 2000.
- Ding, Y., Ceglarek, D. and Shi, J., Fault diagnosis of multi-station manufacturing processes using state space approach. *ASME J. Manuf. Sci. Eng.*, 2002a, **124**, 313–322.
- Ding, Y., Shi, J. and Ceglarek, D., Diagnosability analysis of multi-station manufacturing processes. *ASME J. Dynam. Syst., Measure., Cont.*, 2002b, **124**, 1–13.
- Ding, Y., Kim, P., Ceglarek, D. and Jin, J., Optimal sensor distribution for variation diagnosis for multi-station assembly processes. *IEEE Trans. Robot. Autom.*, 2003, **19**, 543–556.
- Ding, Y., Gupta, A. and Apley, D.W., Singularity issues in fixture fault diagnosis for multi-station assembly processes. *ASME J. Manuf. Sci. Eng.*, 2004, **126**, 200–210.
- Ding, Y., Zhou, S. and Chen, Y., A comparison of process variance estimation methods for in-process dimensional measurements and control. *ASME J. Dynam. Syst., Measure.*, 2005, **127**, 69–79.
- Djurdjanovic, D. and Ni, J., Linear state space modeling of dimensional machining errors. *Trans. NAMRI/SME*, 2001, **XXIX**, 541–548.
- Djurdjanovic, D. and Ni, J., Bayesian approach to measurement scheme analysis in multistation machining systems. *J. Eng. Manuf.*, 2003, **217**, 1117–1130.
- Djurdjanovic, D. and Ni, J., Measurement scheme synthesis in multi-station machining systems. *ASME J. Manuf. Sci. Eng.*, 2004, **126**, 178–188.
- Hu, S.J. and Wu, S., Identifying sources of variation in automobile body assembly using principal component analysis. *Trans. NAMRI/SME*, 1992, **XX**, 311–316.
- Jin, J. and Shi, J., State space modeling of sheet metal assembly for dimensional control. *ASME J. Manuf. Sci. Eng.*, 1999, **121**, 756–762.
- Johnson, R.A. and Wichern, D.W., *Applied Multivariate Statistical Analysis*, 5th ed., 2002 (Prentice-Hall: New Jersey, Inc.).
- Khan, A., Ceglarek, D. and Ni, J., Sensor location optimisation for fault diagnosis in multi-fixture assembly systems. *ASME J. Manuf. Sci. Eng.*, 1998, **120**, 781–792.
- Khan, A., Ceglarek, D., Shi, J., Ni, J. and Woo, T.C., Sensor optimisation for fault diagnosis in single fixture systems: A methodology. *ASME J. Manuf. Sci. Eng.*, 1999, **121**, 109–121.
- Khan, A. and Ceglarek, D., Sensor optimisation for fault diagnosis in multi-fixture assembly systems with distributed sensing. *ASME J. Manuf. Sci. Eng.*, 2000, **122**, 215–226.
- Liu, Q., Ding, Y. and Chen, Y., Optimal coordinate sensor placements for estimating mean and variance components of variation sources. *IIE Trans.*, 2005, **37**, 877–889.
- Mantripragada, R. and Whitney, D.E., Modeling and controlling variation propagation in mechanical assemblies using state transition models. *IEEE Trans. Robot. Autom.*, 1999, **15**, 124–140.

- Qin, S.J., Statistical process monitoring: Basics and beyond. *J. Chemometrics*, 2003, **17**, 480–502.
- Rao, C.R. and Kleffe, J., *Estimation of Variance Components and Applications*, 1988 (North-Holland: Amsterdam).
- Rong, Q., Ceglarek, D. and Shi, J., Dimensional fault diagnosis for compliant beam structure assemblies. *ASME J. Manuf. Sci. Eng.*, 2000, **122**, 773–780.
- Schott, J.R., *Matrix Analysis for Statistics*, 1997 (John Wiley & Sons: New York, NY).
- Searle, S.R., Casella, G. and McCulloch, C.E., *Variance Components*, 1992 (John Wiley & Sons: New York, NY).
- Shalon, D., Gossard, D., Ulrich, K. and Fitzpatrick, D., Representing geometric variations in complex structural assemblies on CAD systems, in *Proceedings of the 19th Annual ASME Advances in Design Automation Conference*, 1992, **DE-Vol. 44-2**, 121–132.
- Wang, Y. and Nagarkar, S. R., Locator and sensor placement for automated coordinate checking fixtures. *ASME J. Manuf. Sci. Eng.*, 1999, **121**, 709–719.
- Zhou, S., Ding, Y., Chen, Y. and Shi, J., Diagnosability study of multistage manufacturing processes based on linear mixed-effects models. *Technometrics*, 2003a, **45**, 312–325.
- Zhou, S., Huang, Q. and Shi, J., State space modeling for dimensional monitoring of multi-stage machining process using differential motion vector. *IEEE Trans. Robot. Autom.*, 2003b, **19**, 296–308.