Causation-Based T^2 Decomposition for Multivariate Process Monitoring and Diagnosis

JING LI

Arizona State University, Tempe, AZ 85287-5906

JIONGHUA JIN and JIANJUN SHI

University of Michigan, Ann Arbor, MI 48109-2117

Multivariate process monitoring and diagnosis is an important and challenging issue. The widely adopted Hotelling T^2 control chart can effectively detect a change in a system but is not capable of diagnosing the root causes of the change. The MTY approach makes efforts to improve the diagnosability by decomposing the T^2 statistic. However, this approach is computationally intensive and has a limited capability in root-cause diagnosis for a large dimension of variables. This paper proposes a causation-based T^2 decomposition method that integrates the causal relationships revealed by a Bayesian network with the traditional MTY approach. Theoretical analysis and simulation studies demonstrate that the proposed method substantially reduces the computational complexity and enhances the diagnosability, compared with the MTY approach.

Key Words: Bayesian Network; Causal Model; SPC.

With the wide adoption of sensors and sensor networks, multivariate process monitoring and diagnosis has become more and more important. Multivariate SPC, typically the Hotelling T^2 control chart (Alt (1985), Tracy et al. (1992)), provides an effective way to detect the mean shift of a random vector by generating out-of-control signals. However, it does not provide any information on which variable(s) in this vector causes the mean shift.

Dr. Li is an Assistant Professor in the Department of Industrial Engineering at Arizona State University. She is a Member of ASQ. Her email address is jinglz@asu.edu.

Dr. Jin is an Associate Professor in the Department of Industrial and Operations Engineering at the University of Michigan. She is a Member of ASQ. Her email address is jhjin@umich.edu.

Dr. Shi is a Professor in the Department of Industrial and Operations Engineering at the University of Michigan. He is a Senior Member of ASQ. His email address is shihang@umich.edu.

To improve the diagnosability of the Hotelling T^2 control chart, Mason et al. (1995) proposed decomposing the T^2 statistic into independent terms and charting those terms individually in order to identify the variables that significantly contribute to an out-of-control T^2 signal. This approach is referred to as the "MTY approach" in this paper. While the MTY approach is theoretically sound and appealing, it has some inherent deficiencies, especially when the dimension of variables is high. For example, it requires $p!$ different decompositions of the T^2 statistic for a system of p variables, which results in $p \times 2^{p-1}$ distinct terms being examined in the diagnosis. Although efforts were made to reduce the number of examined terms (Mason et al. (1997)), this number still far exceeds p, especially when there are multiple "faults" in the system. In this paper, a fault is defined as the mean shift of a variable.

In addition to the computational issue, the diagnosability of the MTY approach is also a concern. Although the decomposition of the T^2 statistic provides a way to identify the variables that significantly con-

tribute to an out-of-control T^2 signal, these variables may not be the root-cause variables of the fault in the system. For example, if variable X has a mean shift, this shift will propagate to X's downstream variables that are affected by X. In this case, both X and its downstream variables may be identified as significant variables in the T^2 decomposition. However, as the fault only occurs in X, the diagnostic decision concluding that not only X but also its downstream variables are the root causes of the fault is misleading. This situation becomes more severe as the number of variables and the number of faults increase.

This paper investigates how to effectively utilize the causal relationships among variables, represented by a Bayesian network (BN for short) (Korb and Nicholson (2003)), to reduce the computational complexity and improve the diagnosability of the MTY approach. The proposed method, called the "causation-based T^2 decomposition method," shares a common objective with the research in POBREP (process-oriented basis representation): integrating process knowledge and statistical data analysis. The early work in POBREP dated back to 1950, when Seder developed position-dimension (P-D) diagrams to describe patterns of dimensional errors at multiple locations on a part. Because these patterns were associated with specific causes, diagnosis of faults was visually achieved by qualitatively examining the patterns. Barton and Gonzalez-Barreto (1996) proposed the representation of a quality measurement vector as a linear combination of preidentified fault patterns plus a residual. The fault diagnosis was achieved through online monitoring of the coefficients of the linear model. This method is especially useful when the quality vector contains measurements of the same unit at different locations on a part or other types of profile data. Runger et al. (2007) further studied how to estimate those coefficients considering whether the process-oriented effects occur only as special causes or also as common causes of variation. Moreover, process- and product-design information was combined with statistical methods, such as principal components analysis and least squares estimation, to diagnose fixture faults in autobody assembly processes (Ceglarek and Shi (1996), Apley and Shi (1998)). SOV (stream-of-variation) methodologies were initially developed based on the state-space model representation of variation propagations in multistage manufacturing processes (Jin and Shi (1999), Shi (2006)).

The proposed causation-based T^2 decomposition method utilizes a different fault diagnosis mechanism from POBREP. POBREP methods focus on the analysis of multivariate *quality* variation patterns, which are the resultant effects of process faults (e.g., material impurities, tooling errors, or/and excessive temperature). However, POBREP methods do not specifically model the relationships among *process* variables and the interrelationships between process variables and quality variables. As a contrast, the causation-based T^2 decomposition method denotes each process and individual quality variable and their causal relationships explicitly, using a BN, and further investigates how to trace backward from certain quality or process problems (e.g., mean shift(s) in some quality or/and process variables) to the variables that are the root causes.

The adoption of BNs in this research is appealing mainly because it provides an effective model for describing causal influences. Although BNs are mathematically defined in terms of probabilistic independence statements, a connection can be made between this characterization and the notion of causal influences (Pearl and Verma (1991), Spirtes et al. (1993)). The causal interpretations facilitate the application of BNs to solve many real-world problems (Mani and Cooper (1999), Friedman et al. (2000), Li and Shi (2007), Li et al. (2006)). As an example, Figure 1 shows a BN structure of a hot forming process with one quality variable (X_5: final dimension of workpiece) and four process variables (X_1, temperature; X_2, material flow stress; X_3, tension in workpiece; and X_4, blank holding force, or BHF). A two-dimensional physical illustration of the hot forming process is given in Figure 2. Because the material flow stress (X_2) and the tension in a workpiece (X_3) directly affect the final workpiece dimension (X_5), where "directly" means that the causal influences are not mediated through other variables, X_2 and X_3 are connected to X_5 by directed arcs. The BHF (X_4) also affects the dimension of the workpiece (X_5), but only indirectly, i.e., through the tension in the work-

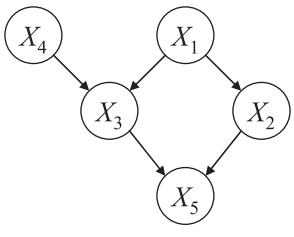

FIGURE 1. BN Structure of a Hot Forming Process.

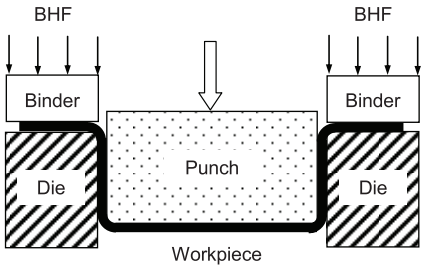

FIGURE 2. 2-D Illustration of the Hot Forming Process.

piece (X_3). Thus, there is no directed arc between X_4 and X_5. Similar interpretations can be applied to the causal relationships between other variables.

The rest of this paper is organized as follows. First, the key notations, concepts, and learning algorithms of Bayesian networks are introduced. Next, the proposed causation-based T^2 decomposition method is illustrated. Then, the hot forming process in Figures 1 and 2 is used as an example to demonstrate the proposed method. Finally, the conclusion is given.

Key Notations, Concepts, and Learning Algorithms of Bayesian Networks

A BN has two components: structure and parameters. This section introduces the key notations and concepts and the learning algorithms for each component.

BN Structure

The structure of a BN is a *directed acyclic graph* (DAG), i.e., a set of nodes $\{X_1, \ldots, X_p\}$ connected by directed arcs. A directed graph is *acyclic* if there is no directed path $X_i \rightarrow \cdots \rightarrow X_j$ such that $X_i = X_j$. The nodes represent random variables. If there is a directed arc from X_i to X_j, i.e., $X_i \rightarrow X_j$, then X_i is called a *parent* of X_j. If there is a directed path from X_j to X_k, i.e., $X_j \rightarrow \cdots \rightarrow X_k$, then X_k is called a *descendant* of X_j. In this paper, let $\mathbf{PA}(X_j)$ denote the set of variables that are all parents of X_j and let $\mathbf{DE}(X_j)$ denote the set of variables that are all descendants of X_j. For example, in Figure 1, $\mathbf{PA}(X_5) = \{X_2, X_3\}$ and $\mathbf{DE}(X_4) = \{X_3, X_5\}$. A DAG encodes the *Markov condition*, i.e., a variable is independent of its nondescendants given its parents. If the directions of the arcs in a DAG have causal interpretations, i.e., the parents of a variable are its direct causes, then the BN becomes a causal model, also called a *causal network*.

BN Parameters

By applying the chain rule of probabilities and utilizing the conditional independence relationships implied by the Markov condition, the joint distribution of all variables in a BN can be decomposed into a set of independent conditional probabilities, i.e.,

$$P(X_1, \ldots, X_p) = \prod_{j=1}^{p} P(X_j \mid \mathbf{PA}(X_j)). \qquad (1)$$

$P(X_j \mid \mathbf{PA}(X_j))$, $j = 1, \ldots, p$, are called parameters of the BN.

In this paper, it is assumed that all variables are continuous and a variable, X_j, given its parents of m_j variables denoted by $\mathbf{PA}(X_j) = \{PA_1(x_j), \ldots, PA_{m_j}(X_j)\}$, follows a Gaussian distribution, i.e.,

$$P(X_j \mid pa_1(X_j), \ldots, pa_{m_j}(X_j)) \sim N\left(\alpha_{0j} + \sum_{k=1}^{m_j} \alpha_{kj} pa_k(X_j), \sigma_j^2\right), \qquad (2)$$

where $pa_k(X_j)$, $k = 1, \ldots, m_j$, denotes the value of $PA_k(X_j)$; α_{0j}, α_{kj} are the coefficients that linearly link the values of X_j's parents with the mean of X_j; σ_j^2 is the variance of X_j; and σ_j^2 is not influenced by the values of X_j's parents. This parameterization of a BN is called *linear Gaussian*. If all variables in a BN have linear Gaussian conditional distributions, then the joint distribution is multivariate Gaussian (Lauritzen and Wermuth (1989)).

If the linear Gaussian BN has causal interpretations, an alternative representation to Equation (2) is often adopted (Korb and Nicholson (2003)), i.e.,

$$X_j = \alpha_{0j} + \sum_{k=1}^{m_j} \alpha_{kj} PA_k(X_j) + V_j, \qquad (3)$$

where V_j is the disturbance, $V_j \sim N(0, \sigma_j^2)$. In order to directly compare the relative strength of the causal influences from $PA_k(X_j)$, $k = 1, \ldots, m_j$, it is customary to use standardized variables (i.e., the variables having means equal to 0 and standard deviations equal to 1) in a linear Gaussian BN, denoted as Z_j. Then Equation (3) can be written as

$$Z_j = \sum_{k=1}^{m_j} p(PA_k(Z_j), Z_j) PA_k(Z_j) + E_j. \qquad (4)$$

The representation in Equation (4) is also adopted in path modeling (Wright (1921)), which has been extensively employed in the social sciences. $p(\cdot, \cdot)$ is called a *path coefficient*. $PA_k(Z_j)$ is the $PA_k(X_j)$

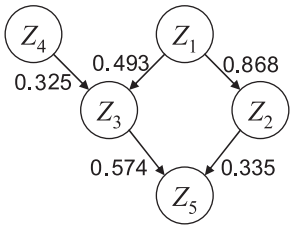

FIGURE 3. BN Structure and Parameters of the Hot Forming Process Using Standardized Variables.

after standardization. $E_j \sim N(0, \tilde{\sigma}_j^2)$ and $\tilde{\sigma}_j^2$ is given by (Weisberg (1985))

$$\tilde{\sigma}_j^2 = 1 - \sum_{k=1}^{m_j} p(PA_k(Z_j), Z_j)\rho(PA_k(Z_j), Z_j), \quad (5)$$

where $\rho(\cdot,\cdot)$ denotes the correlation coefficient between two variables. Figure 3 shows the structure of the BN in Figure 1 using standardized variables. The numbers beside the directed arcs are estimates of the parameters (i.e., the path coefficients) based on a particular dataset.

Learning of the BN

There are two steps in learning a BN: learning the structure and learning the parameters given the BN structure. The structure can be built from engineering knowledge or learning from data. One commonly used data-driven algorithm for structure learning is called PC (Peter and Clark; Spirtes et al. (1993)), which uses a series of statistical significance tests of conditional independence. For linear Gaussian BNs, PC uses the tests of partial correlation. Taking the BN structure in Figure 3 as an example, the PC algorithm is briefly introduced here.

- Start with a fully connected undirected graph in which each variable is linked with all other variables by undirected arcs.
- The undirected arc between one pair of variables is removed when conditional independence (or close to zero partial correlation) is found. For example, in Figure 3, the arc between Z_2 and Z_3 is removed because Z_2 is uncorrelated with Z_3 given their common cause Z_1; and the arc between Z_5 and Z_1 is removed because Z_5 is uncorrelated with Z_1 given Z_5's direct causes Z_2 and Z_3. Other undirected arcs can be removed similarly.
- For the undirected arcs that cannot be removed in the previous step, the following rule is used to orient the arcs: each triple of variables $Z_i - Z_j - Z_k$ is oriented as $Z_i \rightarrow Z_j \leftarrow Z_k$, if Z_i and Z_k are found to be independent (or uncorrelated) given a set of variables that do not contain Z_j. As a result, $Z_3 - Z_5 - Z_2$ is oriented as $Z_3 \rightarrow Z_5 \leftarrow Z_2$ because Z_3 and Z_2 are uncorrelated given $\{Z_1\}$, and $\{Z_1\}$ does not contain Z_5. Similarly, $Z_4 - Z_3 - Z_2$ is oriented as $Z_4 \rightarrow Z_3 \leftarrow Z_1$. Aside from this major rule in orientation, Meek (1995) introduced some supplementary rules, such as orienting the remaining undirected arcs in a way that no cycles are created in the BN. In addition, some orientations may be achieved through knowledge or first principles. For example, if Z_i is known to occur before Z_j and there is no feedback control in the system, then $Z_i \rightarrow Z_j$. Such knowledge helps orient $Z_1 - Z_2$ as $Z_1 \rightarrow Z_2$.

The parameters (i.e., the path coefficients) in a linear Gaussian BN can be estimated based on the sample correlation matrix of all variables in the BN, $\hat{\mathbf{\Sigma}}$, through multiple regressions (Kline (2005)), i.e.,

$$[\hat{p}(PA_1(Z_j), Z_j), \ldots, \hat{p}(PA_{m_j}(Z_j), Z_j)]^{\mathrm{T}} = \widehat{\mathrm{cor}}(\mathbf{PA}(Z_j), \mathbf{PA}(Z_j))^{-1}\widehat{\mathrm{cor}}(\mathbf{PA}(Z_j), Z_j)^{\mathrm{T}},$$

where $\widehat{\mathrm{cor}}(\cdot,\cdot)$ denotes the sample correlation matrix between two sets of variables. The path coefficients of the BN in Figure 3 were estimated in this way.

MLE (maximum likelihood estimation) can also be used to obtain path coefficient estimates that are equivalent to the solutions from multiple regressions. An advantage of MLE is that it can be integrated with Bayesian estimation, resulting in a Bayesian MAP (maximum A-posterior) approach that augments the likelihood with a prior that gives the initial belief about the parameters before seeing any data (Buntine (1996)). The Bayesian MAP approach facilitates the inclusion of prior knowledge into the parameters estimation and progressive updates of the parameters as more data become available.

Causation-Based T^2 Decomposition

The MTY approach generates $p!$ distinct decompositions. With the aid of the causal relationships implied by a BN, it is shown in this section that those decompositions can be classified into two categories (type A and type B) according to their diagnosability. The characteristics of each type of decomposition are discussed. Then, it is proven that all type-B decompositions converge to one decomposition, called "causation-based T^2 decomposition." Finally, the analytic results are summarized and the procedure of

applying the causation-based T^2 decomposition in monitoring and diagnosis is illustrated.

It is assumed in the causation-based T^2 decomposition method that the learned BN is free of learning and sampling errors, i.e., the structure reflects the true underlying causal relationships, and the path coefficients are obtained from the population correlation matrix $\mathbf{\Sigma}$, which is known.

Brief Review of the MTY Approach

Let $\mathbf{X}$ represent a p-dimensional random vector and $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma_X})$ with known mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma_X}$. The Hotelling's (1947) T^2 statistic is

$$T^2 = (\mathbf{X} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{\Sigma}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2, \tag{6}$$

where χ_p^2 represents a χ^2 distribution with p degrees of freedom. Without loss of generality, it is assumed in this paper that all variables are standardized, i.e., $\boldsymbol{\mu} = \mathbf{0}$ and $\mathbf{\Sigma}$ is used to denote the correlation matrix of standardized variables $\mathbf{Z}$. Then Equation (6) can be written as

$$T^2 = \mathbf{Z}^{\mathrm{T}} \mathbf{\Sigma}^{-1} \mathbf{Z}. \tag{7}$$

Mason et al. (1995) proposed a procedure (i.e., the MTY approach) to decompose the T^2 into independent components. This procedure starts with specifying an ordering of the p variables and then partitions the T^2 as

$$T^2 =^r T^2 = \sum_{j=1}^{p} {}^r T^2_{j \cdot 1,\ldots,j-1}. \tag{8}$$

In Equation (8), r serves as an indicator of the specific ordering. Because p variables have $p!$ different orderings, there are $p!$ distinct decompositions, denoted by ${}^rT^2$ ($r = 1, \ldots, p!$), which are all equal to the same T^2. ${}^rT_{j \cdot 1,\ldots,j-1}$ is the jth variable in ordering r, denoted by ${}^rZ_j$, adjusted by the mean and standard deviation of the conditional distribution of ${}^rZ_j$ given ${}^rZ_1, \ldots, {}^rZ_{j-1}$, i.e.,

$${}^rT_{j \cdot 1,\ldots,j-1} = \frac{{}^rZ_j - {}^r\mu_{j \cdot 1,\ldots,j-1}}{{}^r\sigma_{j \cdot 1,\ldots,j-1}}, \tag{9}$$

where

$${}^r\mu_{j \cdot 1,\ldots,j-1} = \sum_{i=1}^{j-1} {}^r\beta_i {}^rZ_i, \tag{10}$$

${}^r\beta_i$ ($i = 1, \ldots, j-1$) are the regression coefficients of ${}^rZ_j$ regressed on ${}^rZ_i$, and, according to Weisberg (1985),

$${}^r\sigma^2_{j \cdot 1,\ldots,j-1} = 1 - \sum_{i=1}^{j-1} {}^r\beta_i \rho({}^rZ_j, {}^rZ_i). \tag{11}$$

For example, if $p = 3$, there are $3! = 6$ decompositions of the T^2, as shown in (12):

$$\begin{aligned} T^2 =^1 T^2 &= T_1^2 + T_{2 \cdot 1}^2 + T_{3 \cdot 1,2}^2 \\ =^2 T^2 &= T_1^2 + T_{3 \cdot 1}^2 + T_{2 \cdot 1,3}^2 \\ =^3 T^2 &= T_2^2 + T_{3 \cdot 2}^2 + T_{1 \cdot 2,3}^2 \\ =^4 T^2 &= T_2^2 + T_{1 \cdot 2}^2 + T_{3 \cdot 1,2}^2 \\ =^5 T^2 &= T_3^2 + T_{1 \cdot 3}^2 + T_{2 \cdot 1,3}^2 \\ =^6 T^2 &= T_3^2 + T_{2 \cdot 3}^2 + T_{1 \cdot 2,3}^2. \end{aligned} \tag{12}$$

With the aid of the decompositions, it is possible to identify which variables significantly contribute to an out-of-control T^2 signal.

Classification of the Decompositions

A BN with causal interpretations provides an effective tool to identify how the fault in one variable propagates to other variables. Specifically, if a fault occurs in Z_j, it will affect all of Z_j's descendents but leave the nondescendents unaffected. This information can be utilized in diagnosing the fault by not only identifying the variables that significantly contribute to the out-of-control signal T^2 (i.e., Z_j and Z_j's descendents) but also exactly locating the root-cause variable of the fault (i.e., Z_j).

However, in the MTY approach, not all decompositions facilitate the exact root-cause diagnosis of the fault. Hawkins (1993) pointed out that, if a variable is regression adjusted by its downstream variables, the mean shift in this variable will be diluted. Here, the "downstream variables" are conceptually equivalent to the descendents in a BN. Because the ${}^rT_{j \cdot 1,\ldots,j-1}$ defined in Equation (9) is obtained by applying regression adjustment to ${}^rZ_j$ for ${}^rZ_1, \ldots, {}^rZ_{j-1}$, Hawkins" statement implies that, if $\{{}^rZ_1, \ldots, {}^rZ_{j-1}\}$ contains the descendents of ${}^rZ_j$, then the mean shift in ${}^rZ_j$ will be diluted in ${}^rT_{j \cdot 1,\ldots,j-1}$, i.e., using ${}^rT_{j \cdot 1,\ldots,j-1}$ may not precisely identify the mean shift in ${}^rZ_j$. Based on this consideration, the $p!$ decompositions in the MTY approach can be divided into two categories in terms of whether there exists a ${}^rT_{j \cdot 1,\ldots,j-1}$ in the decomposition that dilutes the mean shift in ${}^rZ_j$. These two categories of decompositions are defined as follows, where $\cap$ denotes the intersection of two sets and Φ denotes an empty set.

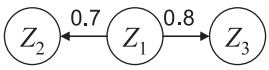

FIGURE 4. The Linear Gaussian BN of a Three-Variable System.

Definition 1. In decomposition ${}^rT^2$, if there exists a term ${}^rT_{j\cdot 1,\ldots,j-1}$ such that $\mathbf{DE}({}^rZ_j) \cap \{{}^rZ_1,\ldots,{}^rZ_{j-1}\} \neq \Phi$, i.e., $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$ contains at least one descendent of ${}^rZ_j$, ${}^rT^2$ is called *type-A decomposition*. Otherwise, if $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$ contains no descendent of ${}^rZ_j$, ${}^rT^2$ is called *type-B decomposition*.

According to Definition 1, a type-A decomposition must contain a term ${}^rT_{j\cdot 1,\ldots,j-1}$ in which $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$ includes some descendents of ${}^rZ_j$. If ${}^rZ_j$ has a mean shift, using ${}^rT_{j\cdot 1,\ldots,j-1}$ for diagnosis runs the risk of misdetecting the shift as the shift is diluted in ${}^rT_{j\cdot 1,\ldots,j-1}$. Thus, compared with type-B decompositions, type-A decompositions have lower diagnosability. This point is further illustrated through the following example.

An Illustrative Example

Consider a simple system of three variables, Z_1, Z_2, and Z_3, whose BN is shown in Figure 4. Among the six decompositions in Equation (12), ${}^1T^2$ and ${}^2T^2$ are type-B decompositions while ${}^3T^2$, ${}^4T^2$, ${}^5T^2$, and ${}^6T^2$ are type-A decompositions, according to Definition 1. In ${}^3T^2$, the term $T_{1\cdot 2,3}$ regresses Z_1 on its descendents Z_2 and Z_3. If Z_1 has a mean shift δ, the detected mean shift using $T_{1\cdot 2,3}$ is $E(T_{1\cdot 2,3}) = 0.52\delta$ (derivations are given in Appendix II). It can be seen that the mean shift in Z_1 is diluted in $T_{1\cdot 2,3}$. Similar phenomena can be found in other type-A decompositions. As a contrast, in the type-B decompositions, the mean shift in Z_j will not be diluted in ${}^rT_{j\cdot 1,\ldots,j-1}$, $j = 1, 2$. The results that support this conclusion are summarized in Table 1. For example, in type-B decomposition ${}^1T^2$ (Table 1(a)), if Z_2 has a mean shift δ, the detected mean shift using $T_{2\cdot 1}$ is $E(T_{2\cdot 1}) = 1.4\delta$. Thus, the mean shift in Z_2 is not diluted but enlarged in $T_{2\cdot 1}$. This enlargement increases the sensitivity in detecting the mean shift.

The above discussion and example demonstrate that type-A decompositions have lower sensitivity in detecting the mean shifts than type-B decompositions. Thus, to facilitate the root-cause diagnosis, type-B decompositions should be used.

TABLE 1. The Detected Mean Shifts in Type-B Decompositions

Variable with a mean shift δ	Mean shift detected by the terms in ${}^jT^2$ $(j = 1, 2)$
(a) ${}^1T^2 = T_1^2 + T_{2\cdot 1}^2 + T_{3\cdot 1,2}^2$	
Z_1	$E(T_1) = \delta$
Z_2	$E(T_{2\cdot 1}) = 1.4\delta$
Z_3	$E(T_{3\cdot 1,2}) = 1.7\delta$
(b) ${}^2T^2 = T_1^2 + T_{3\cdot 1}^2 + T_{2\cdot 1,3}^2$	
Z_1	$E(T_1) = \delta$
Z_2	$E(T_{2\cdot 1,3}) = 1.4\delta$
Z_3	$E(T_{3\cdot 1}) = 1.7\delta$

Causation-Based T^2 Decomposition

This section proves that all type-B decompositions converge to a unique decomposition called the causation-based T^2 decomposition.

Proposition 1. If ${}^rT^2$ is a type-B decomposition, then ${}^rT^2 = \sum_{j=1}^p {}^rT^2_{j\cdot\mathbf{PA}({}^rZ_j)}$.

A detailed proof of Proposition 1 is given in Appendix I.

Once a BN is built, the parents of a variable can be identified. In other words, the parents $\mathbf{PA}(Z_j)$ of Z_j do not change with different orderings of the variables. Thus, the left superscript r in Proposition 1 can be removed, indicating that all type-B decompositions converge to a unique decomposition that is the causation-based T^2 decomposition defined as follows.

Definition 2. The causation-based T^2 decomposition of a T^2 statistic is

$$T^2 = \sum_{j=1}^{p} T^2_{j\cdot\mathbf{PA}(Z_j)}. \tag{13}$$

To illustrate, the example in Figure 4 is revisited. The causation-based T^2 decomposition by Definition 2 is $T^2 = T_1^2 + T_{2\cdot 1}^2 + T_{3\cdot 1}^2$, to which the two type-B decompositions ${}^1T^2$ and ${}^2T^2$ converge according to Proposition 1, i.e., $T^2 =^1 T^2 =^2 T^2 = T_1^2 + T_{2\cdot 1}^2 + T_{3\cdot 1}^2$.

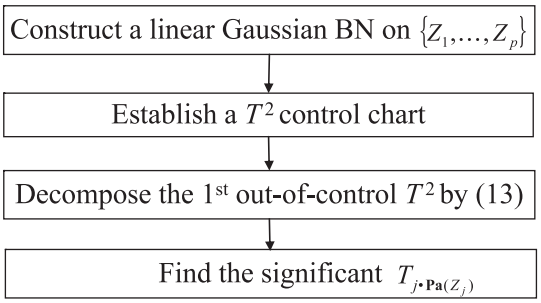

FIGURE 5. Flow Chart of Applying Causation-Based T^2 Decomposition to Monitoring and Diagnosis.

It is seen from Equation (13) that, to perform the causation-based T^2 decomposition, the direct causes (parents) $\mathbf{PA}(Z_j)$ of each variable Z_j must be known. Because these direct cause–effect (parents–child) relationships are explicitly represented by a BN, constructing a BN on the variables is a critical preceding step for the decomposition.

Monitoring and Diagnosis with Causation-Based T^2 Decomposition

Figure 5 shows the steps of applying the causation-based T^2 decomposition to a system of p variables $\{Z_1, \ldots, Z_p\}$. A linear Gaussian BN is first constructed to describe the causal relationships. Then a Hotelling T^2 control chart (Alt (1985)) is established using Equation (7). Once an out-of-control T^2 is generated, it is decomposed by the causation-based T^2 decomposition in Equation (13). Given that $\mathbf{\Sigma}$ is known, $T_{j\cdot\mathbf{PA}(Z_j)}$ follows the $N(0,1)$ distribution. Thus, one can compare each $|T_{j\cdot\mathbf{PA}(Z_j)}|$ with $z_{\alpha'/2}$ (the upper $\alpha'/2$ percentage point of the $N(0,1)$ distribution) to determine which $T_{j\cdot\mathbf{PA}(Z_j)}$ significantly contributes to the out-of-control T^2. A significant $T_{j\cdot\mathbf{PA}(Z_j)}$ implies that Z_j is highly likely to have a mean shift, and the sign of $T_{j\cdot\mathbf{PA}(Z_j)}$ indicates the mean shift direction.

Example

Simulation Setup

The diagnosability of the proposed causation-based T^2 decomposition method is demonstrated through the example in Figure 3. Because this system has five variables, there are five potential single-fault scenarios, each with only one variable having a mean shift, and 26 potential multiple-fault scenarios, each with more than one variable having mean shifts. A total of 31 fault scenarios are shown in the first six columns of Table 2, where δ denotes the magnitude of the mean shift in the unit of standard deviation.

The simulation study includes the following steps:

1. For every fault scenario s ($s = 1, \ldots, 31$), the following substeps (a)–(d) are performed.
   (a) A mean shift of three standard deviations (i.e., $\delta = 3$) is introduced to variable Z_j if Z_j has a mean shift in this fault scenario.
   (b) One dataset with M samples is generated, in which the introduced mean shift occurs at the first sample. A Hotelling T^2 control chart is constructed. $\chi^2_{5,\alpha} = \chi^2_{5,0.05} = 11.071$ is used as the control limit. (M is equal to the number of samples when the first out-of control T^2 is generated.)
   (c) The out-of-control T^2 is decomposed into five independent terms $T^2_{j\cdot\mathbf{PA}(Z_j)}$ ($j = 1, \ldots, 5$) based on Equation (13). Each $|T_{j\cdot\mathbf{PA}(Z_j)}|$ is compared with $z_{\alpha'/2}$. Here, the Bonferroni method is used to set $\alpha' = \alpha/5 = 0.01$ (i.e., $z_{\alpha'/2} = 2.576$) in order to control the overall false-alarm rate.
   (d) If $T_{j\cdot\mathbf{PA}(Z_j)}$ is significant, $d^r_j = 1$; otherwise, $d^r_j = 0$. Here, $j = 1, \ldots, 5$; r is used to number the simulation runs. Comparing the d^r_j ($j = 1, \ldots, 5$) with the truth of the fault scenario in substep (a) provides an assessment of the diagnostic performance.
2. Step 1 is repeated N times. d^r_j ($j = 1, \ldots, 5$; $r = 1, \ldots, N$) are collected.

For better understanding of the simulation procedure, one table from step 1 ($r = 15$) is shown in Table 2. N such tables are generated in step 2.

To compare the performance of the proposed causation-based T^2 decomposition method with the MTY approach, simulations are also conducted for the MTY approach. Because the MTY approach requires examining an excessive number of terms from the $p!$ decompositions, which is computationally involved and needs tremendous interpretation efforts, Mason et al. (1997) proposed a sequential scheme to reduce the computations to a reasonable number. Briefly, this scheme starts with computing the unconditional terms T_j ($j = 1, \ldots, p$). If any T_j is significant, the corresponding Z_j are removed and the T^2 statistic is calculated based on the rest of the variables. If this T^2 is in control, the computation

TABLE 2. Fault Scenarios and the Associated $T_{j\cdot\mathbf{PA}(Z_j)}$, $j = 1, \ldots, 5$ ($\delta = 3$, $r = 15$)

s	Z_1	Z_2	Z_3	Z_4	Z_5	$T_{1\cdot\mathbf{PA}(Z_1)}$	$T_{2\cdot\mathbf{PA}(Z_2)}$	$T_{3\cdot\mathbf{PA}(Z_3)}$	$T_{4\cdot\mathbf{PA}(Z_4)}$	$T_{5\cdot\mathbf{PA}(Z_5)}$	d_1^r	d_2^r	d_3^r	d_4^r	d_5^r
1	δ					2.256	1.922	−0.532	0.756	−2.210	0	0	0	0	0
2		δ				0.599	3.979	0.668	−0.375	−0.563	0	1	0	0	0
3			δ			1.742	−0.565	3.979	−0.842	1.557	0	0	1	0	0
4				δ		0.631	1.511	0.768	4.482	1.836	0	0	0	1	0
5					δ	1.458	1.577	−1.606	1.628	3.543	0	0	0	0	1
6	δ	δ				2.688	2.431	1.506	−0.323	1.060	1	0	0	0	0
7	δ		δ			3.967	1.620	5.079	−1.434	−0.307	1	0	1	0	0
8	δ			δ		2.844	9.291	−0.848	1.343	0.642	1	0	0	0	0
9	δ				δ	2.520	−0.179	0.877	0.019	5.635	0	1	0	0	1
10		δ	δ			0.246	3.280	3.708	−0.713	0.036	0	1	1	0	0
11		δ		δ		−1.096	5.022	0.530	1.420	−0.615	0	1	0	0	0
12		δ			δ	−1.609	3.169	0.765	−0.574	3.186	0	1	0	0	1
13			δ	δ		0.224	−2.305	3.720	1.756	−0.193	0	0	1	0	0
14			δ		δ	−0.571	−0.929	3.936	1.331	5.146	0	0	1	0	1
15				δ	δ	0.145	0.959	0.278	3.432	4.973	0	0	0	1	1
16	δ	δ	δ			3.571	3.573	3.404	0.654	1.014	1	1	1	0	0
17	δ	δ		δ		2.093	5.063	1.607	1.980	0.426	0	1	0	0	0
18	δ	δ			δ	3.637	3.083	0.251	−0.760	5.892	1	1	0	0	1
19	δ		δ	δ		2.724	0.080	5.546	3.449	−1.730	1	0	1	1	0
20	δ		δ		δ	3.928	−0.001	3.309	0.301	4.290	1	0	1	0	1
21	δ			δ	δ	4.460	0.715	−1.384	3.923	3.537	1	0	0	1	1
22		δ	δ	δ		1.150	4.019	3.886	5.605	−1.475	0	1	1	1	0
23		δ	δ		δ	1.197	5.168	3.893	−1.204	4.828	0	1	1	0	1
24		δ		δ	δ	−0.653	4.619	−0.516	3.745	3.561	0	1	0	1	1
25			δ	δ	δ	−1.620	−1.234	5.743	3.076	4.302	0	0	1	1	1
26	δ	δ	δ	δ		3.041	2.767	3.554	2.716	−0.520	1	1	1	1	0
27	δ	δ	δ		δ	3.699	5.502	2.242	1.512	4.001	1	1	0	0	1
28	δ	δ		δ	δ	1.397	4.130	−0.451	4.538	4.253	0	1	0	1	1
29	δ		δ	δ	δ	2.257	0.340	4.309	2.423	5.419	0	0	1	0	1
30		δ	δ	δ	δ	3.087	4.385	2.681	2.464	4.006	1	1	1	0	1
31	δ	δ	δ	δ	δ	3.260	5.910	3.604	1.459	5.677	1	1	1	0	1

ends; otherwise, $T_{j\cdot i}$ is computed for the remaining variables. The Z_j and Z_i corresponding to a significant $T_{j\cdot i}$ are removed and the T^2 statistic is recalculated. If this T^2 is out of control, higher order conditional terms are computed in a similar fashion until no variables remain for calculating the T^2 statistic. The diagnostic decision will be made based on all the removed variables after the entire computation is complete. Specifically, if a variable Z_j is removed when checking the unconditional terms, it indicates that Z_j has a mean shift. Otherwise, it indicates that the relationship between Z_j and other variables has changed.

Following this scheme, the 31 fault scenarios are revisited. The simulation uses the same datasets as those generated in steps 1∼2. The simulation procedure is similar to that in the causation-based T^2 decomposition except that substep (c) is replaced with substep (c′) below:

(c′) The out-of-control T^2 is analyzed by the scheme, i.e., each $|T_j|$ $(j = 1, \ldots, 5)$ is compared with $z_{\alpha'/2}$.

It should be pointed out that, in applying the scheme to the simulation datasets, the computation ends after significant T_j's are identified, i.e., no conditional terms need to be calculated. This is because only mean shifts are introduced to the variables in

TABLE 3. Comparison of Error Rates ${}^{\mathrm{I}}\overline{e}_j^s$ and ${}^{\mathrm{II}}\overline{e}_j^s$ ($j = 1, \cdots, 5$; $s = 1, \cdots, 31$; $N = 5000$)

						${}^{\mathrm{I}}e_j^s/{}^{\mathrm{II}}e_j^s$ (Causation-based T^2 decomposition)					${}^{\mathrm{I}}e_j^s/{}^{\mathrm{II}}e_j^s$ (MTY)				
s	Z_1	Z_2	Z_3	Z_4	Z_5	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
1	δ					0.084	0.013	0.015	0.018	0.015	0.084	0.466	0.217	0.018	0.233
2		δ				0.012	0.011	0.811	0.011	0.010	0.012	0.277	0.012	0.011	0.065
3			δ			0.013	0.014	0.029	0.012	0.010	0.013	0.012	0.221	0.012	0.230
4				δ		0.014	0.014	0.015	0.084	0.015	0.014	0.016	0.083	0.084	0.039
5					δ	0.012	0.010	0.013	0.013	0.005	0.012	0.010	0.012	0.013	0.317
6	δ	δ				0.329	0.054	0.010	0.009	0.013	0.329	0.003	0.135	0.009	0.494
7	δ		δ			0.319	0.008	0.100	0.011	0.013	0.319	0.311	0.015	0.011	0.782
8	δ			δ		0.297	0.011	0.013	0.302	0.011	0.297	0.330	0.485	0.302	0.334
9	δ				δ	0.331	0.005	0.011	0.009	0.018	0.331	0.303	0.141	0.009	0.021
10		δ	δ			0.011	0.059	0.121	0.009	0.011	0.011	0.338	0.340	0.009	0.556
11		δ		δ		0.013	0.054	0.010	0.326	0.010	0.013	0.331	0.056	0.326	0.163
12		δ			δ	0.012	0.057	0.012	0.009	0.025	0.012	0.336	0.009	0.009	0.072
13			δ	δ		0.008	0.010	0.101	0.322	0.012	0.008	0.010	0.063	0.322	0.401
14			δ		δ	0.008	0.009	0.133	0.010	0.021	0.008	0.010	0.330	0.010	0.015
15				δ	δ	0.009	0.009	0.009	0.331	0.017	0.009	0.010	0.058	0.331	0.154
16	δ	δ	δ			0.337	0.057	0.131	0.011	0.011	0.337	0.006	0.029	0.011	0.953
17	δ	δ		δ		0.337	0.059	0.011	0.331	0.009	0.337	0.007	0.453	0.331	0.709
18	δ	δ			δ	0.323	0.060	0.010	0.013	0.025	0.323	0.007	0.146	0.013	0.002
19	δ		δ	δ		0.329	0.007	0.126	0.339	0.008	0.329	0.306	0.001	0.339	0.897
20	δ		δ		δ	0.332	0.013	0.131	0.010	0.021	0.332	0.303	0.029	0.010	0.000
21	δ			δ	δ	0.327	0.011	0.009	0.342	0.023	0.327	0.307	0.459	0.342	0.004
22		δ	δ	δ		0.008	0.070	0.133	0.336	0.008	0.008	0.338	0.085	0.336	0.757
23		δ	δ		δ	0.012	0.059	0.128	0.011	0.021	0.012	0.342	0.341	0.011	0.001
24		δ		δ	δ	0.012	0.059	0.009	0.331	0.027	0.012	0.325	0.050	0.331	0.025
25			δ	δ	δ	0.013	0.012	0.122	0.332	0.023	0.013	0.011	0.078	0.332	0.003
26	δ	δ	δ	δ		0.332	0.057	0.125	0.339	0.013	0.332	0.005	0.002	0.339	0.988
27	δ	δ	δ		δ	0.337	0.059	0.124	0.013	0.024	0.337	0.007	0.027	0.013	0.000
28	δ	δ		δ	δ	0.338	0.058	0.011	0.335	0.020	0.338	0.008	0.462	0.335	0.000
29	δ		δ	δ	δ	0.336	0.008	0.125	0.334	0.019	0.336	0.301	0.003	0.334	0.000
30		δ	δ	δ	δ	0.010	0.058	0.132	0.330	0.022	0.010	0.335	0.087	0.330	0.000
31	δ	δ	δ	δ	δ	0.341	0.061	0.127	0.346	0.024	0.341	0.005	0.001	0.346	0.000

this study and there are no changes in the variable relationships.

Performance Indices and Results of Comparison

The following error rates are defined based on the simulation output d_j^r ($j = 1, \ldots, 5$; $r = 1, \ldots, N$) in order to compare the performances of the causation-based T^2 decomposition method and the MTY approach.

For a fault scenario $s \in \{1, \ldots, 31\}$, if Z_j truly has a mean shift, the error rate of misdetecting this mean shift, ${}^{\mathrm{II}}e_j^s$, is defined as ${}^{\mathrm{II}}e_j^s = 1 - \sum_{r=1}^{N} d_j^r/N$; if Z_j does not have a mean shift, the error rate of falsely identifying a mean shift, ${}^{\mathrm{I}}e_j^s$, is defined as ${}^{\mathrm{I}}e_j^s = \sum_{r=1}^{N} d_j^r/N$. Following these definitions, ${}^{\mathrm{I}}e_j^s$ and ${}^{\mathrm{II}}e_j^s$ ($j = 1, \ldots, 5$; $s = 1, \ldots, 31$) for the causation-based T^2 decomposition method and those for the MTY approach are computed, as shown in Table 3 ($N = 5000$). For example, for fault scenario 10, in which variables Z_2 and Z_3 have mean shifts, $[{}^{\mathrm{I}}e_1^{10}, {}^{\mathrm{II}}e_2^{10}, {}^{\mathrm{II}}e_3^{10}, {}^{\mathrm{I}}e_4^{10}, {}^{\mathrm{I}}e_5^{10}] = [0.011,\ 0.059,$

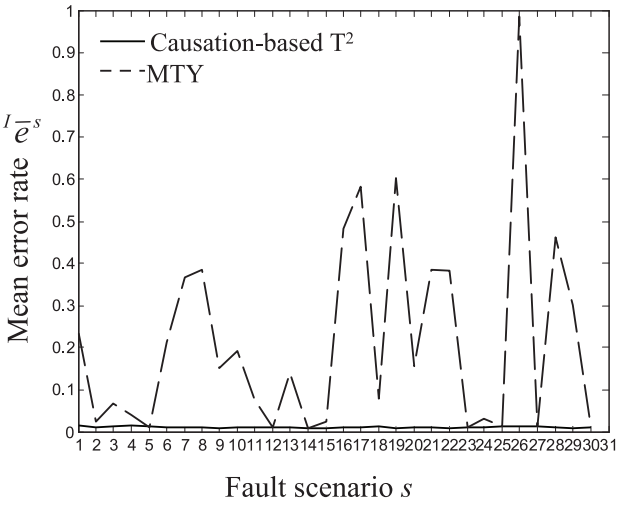

FIGURE 6. Comparison of Mean Error Rates ${}^{\mathrm{I}}\overline{e}^{s}$ ($s = 1, \ldots, 31$).

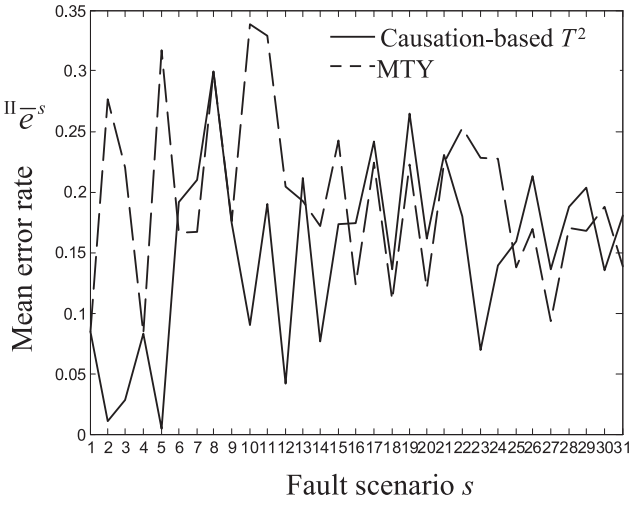

FIGURE 7. Comparison of Mean Error Rates ${}^{\mathrm{II}}\overline{e}^{s}$ ($s = 1, \ldots, 31$).

$0.121, 0.009, 0.011]$ in the causation-based T^2 decomposition method; and $[{}^{\mathrm{I}}e_1^{10}, {}^{\mathrm{II}}e_2^{10}, {}^{\mathrm{II}}e_3^{10}, {}^{\mathrm{I}}e_4^{10}, {}^{\mathrm{I}}e_5^{10}] = [0.011, 0.338, 0.340, 0.009, 0.556]$ in the MTY approach. Thus, in diagnosing fault scenario 10, the causation-based T^2 decomposition method has significantly lower error rates of misdetecting the true mean shifts in Z_2 and Z_3 and a significantly lower error rate of falsely identifying a mean shift in Z_5, compared with the MTY approach. Similar comparisons can be applied to other fault scenarios.

In addition to the aforementioned variable-by-variable comparison, it is of more interest to compare the two approaches in terms of false-positive and false-negative rates, which can be reflected by the two mean error rates defined as follows.

For each fault scenario s, let $\mathbf{\Gamma}^s$ be a set containing the indices of the variables that have mean shifts and $\overline{\mathbf{\Gamma}}^s$ be the complement of $\mathbf{\Gamma}^s$ with respect to $\{1, 2, 3, 4, 5\}$. Let $\dim(\cdot)$ denote the dimension of a set. For example, $\mathbf{\Gamma}^{10} = \{2, 3\}$, $\overline{\mathbf{\Gamma}}^{10} = \{1, 4, 5\}$, and $\dim(\mathbf{\Gamma}^{10}) = 2$. Furthermore, denote the mean error rate of misdetecting a mean shift by ${}^{\mathrm{II}}\bar{e}^{s}$, i.e., ${}^{\mathrm{II}}\bar{e}^{s} = \sum_{j=1,\, j\in\mathbf{\Gamma}^s}^{5} {}^{\mathrm{II}}e_j^s / \dim(\mathbf{\Gamma}^s)$; and denote the mean error rate of falsely identifying a mean shift by ${}^{\mathrm{I}}\bar{e}^{s}$, i.e., ${}^{\mathrm{I}}\bar{e}^{s} = \sum_{j=1,\, j\in\bar{\mathbf{\Gamma}}^s}^{5} {}^{\mathrm{I}}e_j^s / \dim(\overline{\mathbf{\Gamma}}^s)$. For example, ${}^{\mathrm{II}}\overline{e}^{10} = ({}^{\mathrm{II}}e_2^{10} + {}^{\mathrm{II}}e_3^{10})/2 = 0.09$ and ${}^{\mathrm{I}}\overline{e}^{10} = ({}^{\mathrm{I}}e_1^{10} + {}^{\mathrm{I}}e_4^{10} + {}^{\mathrm{I}}e_5^{10})/3 = 0.103$ in the causation-based T^2 decomposition method; and ${}^{\mathrm{II}}\overline{e}^{10} = 0.339$ and ${}^{\mathrm{I}}\overline{e}^{10} = 0.192$ in the MTY approach. Thus, in diagnosing fault scenario 10, the causation-based T^2 decomposition method has lower mean error rates than the MTY approach. Similarly, the mean error rates of the two approaches for other fault scenarios can be computed. The results are plotted in Figures 6 and 7. It can be seen that the ${}^{\mathrm{I}}\overline{e}^{s}$ ($s = 1, \ldots, 31$) of the causation-based T^2 decomposition method is no higher than that of the MTY approach in all the fault scenarios, while the ${}^{\mathrm{II}}\overline{e}^{s}$ ($s = 1, \ldots, 31$) of the two approaches are comparable. Because ${}^{\mathrm{I}}\overline{e}^{s}$ and ${}^{\mathrm{II}}\overline{e}^{s}$ reflect the performance of an approach in terms of false-positive and false-negative rates, the results in Figures 6 and 7 demonstrate that the causation-based T^2 decomposition method is significantly better than the MTY approach in lowering the false-positive rate.

Conclusion

Multivariate process monitoring and diagnosis is an important problem. While the widely adopted Hotelling T^2 control chart is effective in change detection, it alone cannot identify the root causes of the change. Thus, the MTY approach was developed to identify the variables significantly contributing to an out-of-control T^2 statistic. However, there are two major deficiencies in this approach. First, it is computationally expensive, as $p!$ different decompositions of the T^2 are needed. Second, the improvement in diagnosability is still limited.

This paper proposed a causation-based T^2 decomposition method, which integrated the causal relationships, represented by a BN, with the T^2 decomposition. The development of the method included several steps. First, based on the BN, the $p!$ decom-

positions were classified into two categories, type-A and type-B. It was found that diagnosis with typeA decompositions ran the risk of misdetecting the faults. Thus, type-B decompositions should be used to ensure diagnosability. Second, it was proved that all type-B decompositions converge to a unique decomposition, called the causation-based T^2 decomposition. Finally, the procedure of applying the causation-based T^2 decomposition to monitoring and diagnosis was illustrated.

Compared with the MTY approach, the proposed method has the advantages of significantly reduced computational complexity and enhanced diagnosability. In addition to the analytical study, an example of a hot forming process was also provided. Results showed that the causation-based T^2 decomposition method had a substantially lower false-positive error rate than the MTY approach.

There are a number of problems that need to be investigated in future research. For example, as an extension to Bayesian networks that are acyclic graphs, probabilistic networks (Buntine (1996)) can be adopted to represent manufacturing systems with a feedback loop, and the monitoring and diagnosis issues in such systems will be studied. Also, in the current Bayesian network setting, each node represents one process or quality variable. This can be extended by denoting one "feature" as a node. A feature may correspond to a raw variable or a parameter of a model on several variables, such as a process-oriented coefficient. Effective integration of the causation-based T^2 decomposition with PO-BREP modeling can address a broader range of complex systems.

Appendix I
Proof of Proposition 1

The proof of Proposition 1 includes two steps. The first step proves that, in a type-B decomposition, ${}^rT^2 = \sum_{j=1}^{p} {}^rT^2_{j\cdot 1,\ldots,j-1}$, the conditional set $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$ of the variable ${}^rZ_j$ includes all the parents of ${}^rZ_j$, $j = 1,\ldots,p$. This is Lemma 1 given below. The second step shows that the regression coefficients of ${}^rZ_j$ regressed on its nonparents in $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$ are equal to zero, $j = 1,\ldots,p$. Proving this utilizes the conclusion in Lemma 2 and the definition of ${}^rT_{j\cdot 1,\ldots,j-1}$ in Equations (9), (10), and (11).

Lemma 1. If ${}^rT^2$ is a type-B decomposition, then for any ${}^rT^2_{j\cdot 1,\ldots,j-1}$ $(j = 1,\ldots,p)$, $\mathbf{PA}({}^rZ_j) \subseteq \{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$.

Proof by contradiction: If there exists a ${}^rT^2_{j\cdot 1,\ldots,j-1}$ such that at least one parent of ${}^rZ_j$, denoted by ${}^rZ_l$, does not belong to $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$, i.e., $l \notin \{1,\ldots,j-1\}$ in ordering r, then $l \in \{j+1,\ldots,p\}$. Thus, according to Equation (8), ${}^rZ_j \in \{{}^rZ_1,\ldots,{}^rZ_{l-1}\}$ for ${}^rT^2_{l\cdot 1,\ldots,l-1}$, i.e., $\{{}^rZ_1,\ldots,{}^rZ_{l-1}\}$ contains a descendent of ${}^rZ_l$, implying that ${}^rT^2$ is a type-A decomposition, which conflicts with the given condition that ${}^rT^2$ is a type-B decomposition. □

Lemma 2. If Z_i is a nondescendent of Z_j, then the correlation between Z_i and Z_j, $\rho(Z_i, Z_j)$, can be written as

$$\rho(Z_i, Z_j) = \sum_{k=1}^{m_j} \rho(Z_i, PA_k(Z_j))p(PA_k(Z_j), Z_j).$$

Proof. It can be shown that

$$\begin{aligned}
&\rho(Z_i, Z_j)\\
&\quad = E(Z_iZ_j)\\
&\quad = E\left(Z_i\left(\sum_{k=1}^{m_j} p(PA_k(Z_j), Z_j)PA_k(Z_j) + E_j\right)\right)\\
&\quad = \sum_{k=1}^{m_j} \rho(Z_i, PA_k(z_j))p(PA_k(Z_j), Z_j) + E(Z_iE_j),
\end{aligned}$$

where the second equality holds by replacing Z_j with the righthand side of Equation (4). Because E_j is the disturbance to Z_j, it will impact Z_j and also Z_j's descendents through Z_j. But E_j cannot affect Z_i, which is a nondescendent of Z_j. Thus, $E(Z_iE_j) = 0$. □

Proof of Proposition 1 Using Lemmas 1 and 2

By Lemma 1, $\mathbf{PA}({}^rZ_j) \subseteq \{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$. Let $\overline{\mathbf{PA}}({}^rZ_j)$ denote the complement of $\mathbf{PA}({}^rZ_j)$ with respect to $\{{}^rZ_1,\ldots,{}^rZ_{j-1}\}$. Based on the definition of ${}^rT^2_{j\cdot 1,\ldots,j-1}$ in Equations (9), (10), and (11), it is known that to prove this proposition is to prove $\boldsymbol{\beta}_{\overline{Pa}} = \mathbf{0}$ in the regression below, where $\boldsymbol{\beta}_{Pa}$ and $\boldsymbol{\beta}_{\overline{Pa}}$ are the regression coefficients of the predictors $\mathbf{PA}({}^rZ_j)$ and $\overline{\mathbf{PA}}({}^rZ_j)$, respectively.

$${}^rZ_j = \boldsymbol{\beta}^{\mathrm{T}}_{Pa}\mathbf{PA}({}^rZ_j) + \boldsymbol{\beta}^{\mathrm{T}}_{\overline{\mathbf{PA}}}\overline{\mathbf{PA}}({}^rZ_j) + {}^r e_j. \quad \text{(A1)}$$

For a general regression model $Y = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X} + \varepsilon$ (assuming all variables are standardized), LSE gives $\mathbf{S}_{XX}\hat{\boldsymbol{\beta}} = \mathbf{S}_{XY}$, where $\mathbf{S}_{XX}$ is the sample correlation matrix of predictors $\mathbf{X}$, $\mathbf{S}_{XY}$ is a vector of sample correlations between predictors $\mathbf{X}$ and response Y,

and $\hat{\boldsymbol{\beta}}$ is the LSE estimate of regression coefficients $\boldsymbol{\beta}$. If the sample size is large enough to make the estimation errors negligible, $\mathbf{S}_{XX}$ and $\mathbf{S}_{XY}$ are replaced by their population counterparts $\boldsymbol{\Sigma}_{XX}$ and $\boldsymbol{\Sigma}_{XY}$, respectively, and $\boldsymbol{\beta}$ can be solved by $\boldsymbol{\Sigma}_{XX}\boldsymbol{\beta} = \boldsymbol{\Sigma}_{XY}$. Substituting this general LSE notation with the notation in the specific regression model of (A1) gives

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_{Pa} \\ \boldsymbol{\beta}_{\overline{Pa}} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \\ \mathbf{F} \end{bmatrix}, \tag{A2}$$

where

$$\begin{aligned} \mathbf{A} &= \mathrm{cor}(\mathbf{PA}({}^{r}Z_j), \mathbf{PA}({}^{r}Z_j)), \\ \mathbf{D} &= \mathrm{cor}(\overline{\mathbf{PA}}({}^{r}Z_j), \overline{\mathbf{PA}}({}^{r}Z_j)), \\ \mathbf{B} &= \mathrm{cor}(\mathbf{PA}({}^{r}Z_j), \overline{\mathbf{PA}}({}^{r}Z_j)), \\ \mathbf{G} &= \mathrm{cor}(\mathbf{PA}({}^{r}Z_j), {}^{r}Z_j), \end{aligned}$$

and

$$\mathbf{F} = \mathrm{cor}(\overline{\mathbf{PA}}({}^{r}Z_j), {}^{r}Z_j).$$

Let $f_h = \rho(\overline{PA}_h({}^{r}Z_j), {}^{r}Z_j)$, where $\overline{PA}_h({}^{r}Z_j) \in \overline{\mathbf{PA}}({}^{r}Z_j)$. Thus, $f_h \in \mathbf{F}$. Because ${}^{r}T^2$ is a type-B decomposition, $\overline{PA}_h({}^{r}Z_j)$ must be a nondescendent of ${}^{r}Z_j$. According to Lemma 2,

$$f_h = \sum_l \rho(\overline{PA}_h({}^{r}Z_j), PA_l({}^{r}Z_j)) p(PA_l({}^{r}Z_j), {}^{r}Z_j), \tag{A3}$$

where $PA_l({}^{r}Z_j) \in \mathbf{PA}({}^{r}Z_j)$.

By (A2),

$$\begin{aligned} f_h = &\sum_l \rho(\overline{PA}_h({}^{r}Z_j), PA_l({}^{r}Z_j)) \beta_{pa_l} \\ &+ \sum_{h'} \rho(\overline{PA}_h({}^{r}Z_j), \overline{PA}_{h'}({}^{r}Z_j)) \beta_{\overline{pa}_{h'}}, \end{aligned} \tag{A4}$$

where β_{pa_l} is the regression coefficient for the lth parent in $\mathbf{PA}({}^{r}Z_j)$ and $\beta_{\overline{pa}_{h'}}$ is the regression coefficient for the h'th nonparent in $\overline{\mathbf{PA}}({}^{r}Z_j)$.

Equalizing Equations (A3) and (A4) gives one solution to β_{pa_l} and $\beta_{\overline{pa}_{h'}}$, i.e.,

$$\beta_{pa_l} = p(PA_l({}^{r}Z_j), {}^{r}Z_j) \quad \text{and} \quad \beta_{\overline{pa}_{h'}} = 0. \tag{A5}$$

Similar to f_h, let $g_l = \rho(PA_l({}^{r}Z_j), {}^{r}Z_j)$. Thus, $g_l \in \mathbf{G}$. According to Lemma 2,

$$\begin{aligned} g_l = &\sum_{l'} \rho(PA_l({}^{r}Z_j), PA_{l'}({}^{r}Z_j)) \\ &\times p(PA_{l'}({}^{r}Z_j), {}^{r}Z_j). \end{aligned} \tag{A6}$$

By (A2),

$$\begin{aligned} g_l = &\sum_{l'} \rho(PA_l({}^{r}Z_j), PA_{l'}({}^{r}Z_j)) \beta_{pa_{l'}} \\ &+ \sum_{h} \rho(\overline{PA}_l({}^{r}Z_j), \overline{PA}_h({}^{r}Z_j)) \beta_{\overline{pa}_h}. \end{aligned} \tag{A7}$$

Equalizing Equations (A6) and (A7) also gives one solution to $\beta_{pa_{l'}}$ and $\beta_{\overline{pa}_h}$, i.e.,

$$\beta_{pa_{l'}} = p(PA_{l'}({}^{r}Z_j), {}^{r}Z_j) \quad \text{and} \quad \beta_{\overline{pa}_h} = 0, \tag{A8}$$

which is the same as the solution in Equation (A5), implying that Equations (A5) is a solution to Equation (A2). Assuming that there is no multicollinearity between the variables,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{D} \end{bmatrix}$$

is nonsingular and Equation (A5) is the unique solution to (A2). Thus, $\boldsymbol{\beta}_{\overline{Pa}} = \mathbf{0}$.

Appendix II
Calculation of $E(T_{j\cdot 1,\ldots,j-1})$ When Z_j Has a Mean Shift

This section shows how to compute $E(T_{1\cdot 2,3})$ when Z_1 has a mean shift δ. In a similar way, the $E(T_{j\cdot 1,\ldots,j-1})$ in Table 1 can be obtained.

Based on the definition in Equation (9), $T_{1\cdot 2,3} = (Z_1 - \mu_{1\cdot 2,3})/\sigma_{1\cdot 2,3}$, where $\mu_{1\cdot 2,3} = \beta_2 Z_2 + \beta_3 Z_3$ by Equation (10) and $\sigma^2_{1\cdot 2,3} = 1 - (\beta_2\rho(Z_1, Z_2) + \beta_3\rho(Z_1, Z_3))$ by Equation (11). Thus,

$$E(T_{1\cdot 2,3}) = \frac{E(Z_1) - (\beta_2 E(Z_2) + \beta_3 E(Z_3))}{\sqrt{1 - (\beta_2\rho(Z_1, Z_2) + \beta_3\rho(Z_1, Z_3))}}. \tag{A9}$$

In Equation (A9), it is given that $E(Z_1) = \delta$. Then $E(Z_2) = 0.7\delta$ and $E(Z_3) = 0.8\delta$ according to the BN in Figure 4. The path coefficients in the BN indicate that $\rho(Z_1, Z_2) = 0.7$, $\rho(Z_1, Z_3) = 0.8$, and $\rho(Z_2, Z_3) = 0.7 \times 0.8 = 0.56$. β_2, and β_3 can be solved by least squares estimation, i.e.,

$$\begin{aligned} \begin{bmatrix} \beta_2 \\ \beta_3 \end{bmatrix} &= \begin{bmatrix} 1 & \rho(Z_2, Z_3) \\ \rho(Z_2, Z_3) & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho(Z_1, Z_2) \\ \rho(Z_1, Z_3) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0.56 \\ 0.56 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0.7 \\ 0.8 \end{bmatrix} = \begin{bmatrix} 0.3671 \\ 0.5944 \end{bmatrix}. \end{aligned}$$

Inserting these numerical values into Equation (A9) gives $E(T_{1\cdot 2,3}) = 0.52\delta$. □

References

ALT, F. B. (1985). "Multivariate Quality Control". In *Encyclopedia of Statistical Sciences*, Kotz, S.; Johnson, N. L.; and Read, C. R., eds., 6, pp. 111–122.

Apley, D. W. and Shi, J. (1998). "Diagnosis of Multiple Fixture Faults in Panel Assembly". *ASME Journal of Manufacturing Science and Engineering* 120, pp. 793–801.

Barton, R. R. and Gonzalez-Barreto, D. R. (1996). "Process-Oriented Basis Representations for Multivariate Process Diagnostics". *Quality Engineering* 9, pp. 107–118.

Buntine, W. (1996). "A Guide to the Literature on Learning Probabilistic Networks from Data". *IEEE Transactions on Knowledge and Data Engineering* 8, pp. 195–210.

Ceglarek, D. and Shi, J. (1996). "Fixture Failure Diagnosis for Auto Body Assembly Using Pattern Recognition". *ASME Transactions, Journal of Engineering for Industry* 118, pp. 55–65.

Friedman, N., et al. (2000). "Using Bayesian Networks to Analyze Expression Data," *Journal of Computational Biology* 7(3–4), pp. 601–620.

Hawkins, D. M. (1991). "Multivariate Quality Control Based on Regression-Adjusted Variables". *Technometrics* 33, pp. 61–75.

Hawkins, D. M. (1993). "Regression Adjustment for Variables in Multivariate Quality Control". *Journal of Quality Technology* 25, pp. 170–182.

Hotelling, H. (1947). "Multivariate Quality Control". In *Techniques of Statistical Analysis*, Eisenhart, X. X.; Hastay, X. X.; and Wallis, X. X., eds., McGraw-Hill, New York.

Jin, J. and Shi, J. (1999). "State Space Modeling of Sheet Metal Assembly for Dimensional Control". *ASME Journal of Manufacturing Science and Engineering* 121, pp. 756–762.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. The Guilford Press, New York.

Korb, K.B. and Nicholson, A. E. (2003). *Bayesian Artificial Intelligence*. Chapman & Hall/CRC, London, UK.

Lauritzen, S. L. and Wermuth, N. (1989), "Graphical Models for Associations between Variables, Some of Which Are Qualitative and Some Quantitative". *Annals of Statistics* 17, pp. 31–57.

Li, J. and Shi, J. (2007). "Knowledge Discovery from Observational Data for Process Control Using Causal Bayesian Networks". *IIE Transactions* 39 (6), pp. 681–690.

Li, J.; Shi, J.; and Satz, D. (2006). "Modeling and Analysis of Disease and Risk Factors Through Learning Bayesian Network from Observational Data". Technical report.

Mani, S. and Cooper, G. F. (1999). "A Study in Causal Discovery from Population-Based Infant Birth and Death Records". *Proceedings of the AMIA Annual Fall Symposium, Philadelphia, PA*, pp. 315–319.

Mason, R. L.; Tracy, N. D.; and Young, J. C. (1995). "Decomposition of T2 for Multivariate Control Chart Interpretation". *Journal of Quality Technology* 27, pp. 99–108.

Mason, R. L.; Tracy, N. D.; and Young, J. C. (1997). "A Practical Approach for Interpreting Multivariate T2 Control Chart Signals". *Journal of Quality Technology* 29, pp. 396–406.

Meek, C. (1995). "Causal Inference and Causal Explanation with Background Knowledge". *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 403–410.

Pearl, J. and Verma, T. S. (1991). "A Theory of Inferred Causation". In *Principles of Knowledge Representation and Reasoning*. Proceedings of 2nd International Conference (KR '91), pp. 441–452.

Runger, G. C.; Barton, R. R.; Castillo, E. D.; and Woodall, W. H. (2007). "Optimal Monitoring of Multivariate Data for Fault Patterns". *Journal of Quality Technology* 39(2), pp. 159–172.

Seder, L. A. (1950). "Diagnosis with Diagrams—Part II". *Industrial Quality Control* 6, pp. 7–10.

Shi, J. (2006). *Stream of Variation Modeling and Analysis for Multistage Manufacturing Processes*. CRC Press Inc.

Spirtes, P.; Glymour, C.; and Scheines, R. (1993). *Causation, Prediction, and Search*. Springer-Verlag.

Tracy, N. D.; Young, J. C.; and Mason, R. L. (1992). "Multivariate Control Charts for Individual Observations". *Journal of Quality Technology* 24, pp. 88–95.

Weisberg, S. (1985). *Applied Linear Regression*. John Wiley, New York.

Wright, S. (1921). "Correlation and Causation". *Journal of Agricultural Research* 20, pp. 557–585.