Diagnostic Tools for Multivariable Model-Based Control Systems

Parthasarathy Kesavan and Jay H. Lee*

Department of Chemical Engineering, Auburn University, Auburn, Alabama 36849

In this paper we propose a set of diagnostic tools that can be used for detecting and estimating the causes for performance degradation in multivariable model-based control systems. Two simple diagnostic tests are proposed, on the basis of prediction error analysis, to detect the performance degradation and bifurcate the potential causes for the degradation. These tests are useful for narrowing down the list of potential causes for performance degradation. For detailed fault estimation, we propose to use a stochastic model based on Gaussian-sum statistics. The optimal estimation for this model requires a bank of a growing number of filters and is computationally intractable. To propose a practical solution, we combine the different suboptimal approaches in the literature. The results of the diagnosis can be used for alerting the operators about the potential faults and for adapting the controllers.

1. Introduction

The demand for higher efficiency in today's process industries has resulted in the use of advanced model-based multivariable control techniques like model predictive control (MPC) as shown in the recent survey article by Qin (1996). However, a common scenario is that these control systems work well when commissioned, but their performances degrade over time and operators shut them off eventually. To sustain the full benefits of these control systems in dynamic plant environments, they need to be *maintained* on a constant basis (Studebaker, 1995). On the other hand, as the control systems became more complex, the task of maintenance has gotten beyond a normal plant operator's capability. Most advanced control systems like MPC require the presence of experts to diagnose the cause of performance degradation and retune if needed. Given the lack of such personnel in usual plant situations, it is desirable that the maintenance tasks be automated to a great extent. At the very least, systematic tools are needed to guide the engineers to the corrections necessary to recover the performance.

While intensive research over the last few decades has yielded a plethora of control system design techniques, only recently has the importance of the controller maintenance been recognized by the community and have several researchers given serious look at the problem. For single-loop controllers, the autocorrelation test (Box and Jenkins, 1976), closed-loop potential (CLP) (Devries and Wu, 1978), and performance index (PI) (Harris, 1989) have been suggested as tools for assessing the controller performance. They utilize the minimum variance controller performance as a benchmark for comparison. The cross correlation test has also been suggested for distinguishing between plant model error and disturbance model error (Stanfelj *et al.*, 1993). This test utilizes the cross correlation between setpoint variations and IMC error and can be used to diagnose which model needs correction. Note that, unlike the monitoring tools like the autocorrelation test, CLP and PI, this diagnostic test requires setpoint perturbations to the system. These methods have generally been well received by the industries (Kozub, 1996). More recently, Tyler and Morari (1995) have suggested the use of constraints on the FIR coefficients to assess the performance. These constraints can be obtained from the minimum variance criterion or from other desired specifications such as closed-loop rise time, settling time, etc. Apart from the above time domain methods, frequency domain approaches have also been used for monitoring and detection (Pryor, 1982; Kendra and Cinar, 1995). Pryor used the power spectrum for detecting periodic disturbances. Kendra and Cinar used the sensitivity and complimentary sensitivity functions for detecting the performance degradation. The multivariate analogues to the univariate performance indices have been developed recently by Huang *et al.* (1995) and Harris et al. (1995).

Although the above developments are encouraging, maintaining the performance of general multivariable control systems remains a challenge. First, the above performance monitoring methods for single-loop controllers are difficult to generalize for multivariable control systems (e.g., MPC) that are based on more general and time-varying objectives. For these controllers, output trends during normal operations tend to be greatly affected by the present operating conditions and optimization criteria (used in the plant optimization). In addition, while appropriate actions against a performance degradation depend on the cause of the degradation, the problem of diagnosis remains largely untouched. In light of the fact that one is often faced with a wide variety of potential causes, a more general and formal approach is clearly needed for the problem. An effective solution should be applicable to a large class of multivariable control systems and be capable of addressing a wide variety of potential events leading to performance changes. In addition, the framework should allow the user to combine the operating data and experience (as they are gathered) with the model in a synergistic manner for gradual improvement in the efficiency.

In this paper, we outline one potential approach to the problem. Our focus will be on *multivariable model-based* control systems. Like the previous work in this area, our approach is statistical and draws upon tools available in statistical analysis and probability theory. The nature of the problem is such that the cause of performance degradation has to be quickly detected and identified from a wide array of potential causes on the basis of the limited amount of data. To accomplish this in an effective manner, data from the past plant operation and process knowledge (that are often available) must be carefully integrated with the present

* To whom all correspondence should be addressed: phone (334)844−2060, fax (334)844−2063, e-mail: jhl@eng.auburn.edu.

S0888-5885(96)00665-3 CCC: $14.00

operating data. Statistics and probability theory provide a formal framework for doing this.

The approach that we propose here is to first perform some simple tests to narrow down the list of the potential causes and then possibly follow it up with a detailed fault estimation. For the former, we propose two new diagnostic tools on the basis of prediction error analysis. Like the performance monitoring tests discussed above, these tools involve simple statistical tests (e.g., the χ^2 test) on the prediction error and are very easy to apply. However, they do require a full dynamic model, which can be the same model used by the model-based controller. For detailed fault quantification, we first propose a fault model employing *Gaussian-sum* statistics. We argue why Gaussian-sum statistics are well suited to modeling fault signals. When the fault model is added into the process model, the resulting problem is a state estimation problem (for a Gaussian-sum system) which has been well studied in the literature (for example, Buxbaum and Haddad, 1969; Jaffer and Gupta, 1971). The optimal estimator for the resulting system can be derived, but is computationally intractable as it requires a growing number of filters. To propose a practical solution, we combine different approximations available in the literature into a sub-optimal algorithm that endows the user with the ability to trade-off accuracy for computational ease.

The results from the diagnosis can be used in two ways: First, it can be used to guide the control engineer to an appropriate corrective action, if needed. Second, the information can be processed into quantities (e.g., a minimum mean-square estimate of states and parameters) that can be fed back to the model-based controller for performance improvement.

This paper is organized as follows: In section 2, we formulate the detection/diagnosis problem. In section 3, we propose the tools for detection, classification, and screening of different faults. Specific methods are presented, and their implications are discussed. The detailed fault quantification scheme is developed in section 4. We also discuss how the results from the diagnosis can be used in modifying the controller. Comparisons between our approach and some of the well-known approaches to fault diagnosis are then made. The proposed methods are applied to a heat exchanger example in section 5. Some concluding remarks are made in section 6.

2. Problem Definition

2.1. Scope. When an operator observes an unusual trend in the plant signals, his immediate question would be whether the trend will last and, if yes, what is causing it. Normally the operator is faced with an extensive list of potential causes including the following:

(a) Process Parameter Changes. Process parameters can change due to phenomena like catalyst poisoning and heat exchanger fouling or due to nonlinearity. Disturbances that occur infrequently but in large magnitude can also be included in this class.

(b) Disturbance Parameter Changes. Disturbance characteristics affect the appropriate form of a feedback law. Disturbances can change from sustained types to drifting or oscillatory types and vice versa.

(c) Actuator/Sensor Problems. Many of the output signals of a model-based control system are setpoints to the lower-level loops. Models used for control move computation are identified by perturbing these setpoints; therefore, inner-loop dynamics are included in the model. Control valves can exhibit significant nonideal behavior like stiction and hysteresis. Inner loops are often retuned by different operators, and this can also change the loop dynamics significantly. Sensors can also develop a bias, and a drift, and its accuracy can decrease over time. Finally, transmitters can show gain variations and biases. In an interactive system, faults in a single actuator or sensor can induce unusual behavior in all input and output signals. In addition, closed-loop dynamics can be much more sensitive to errors in some actuators/sensors than in others. These make the fault isolation problem nontrivial.

(d) Actuator Saturation. When a large disturbance occurs or a large setpoint change is made, one or more actuators can saturate and this can cause the outputs to deviate significantly from their setpoints.

Note that an appropriate reaction by the operator depends on the root cause of the problem. For instance, in the case of input saturation, the operator has no choice but to wait unless he/she has the freedom to replace the actuator(s). On the other hand, instrumentation problems are often best dealt with through replacement or recalibration. They may also require some temporary modifications to the control system. Sustained changes in the process parameters or disturbance characteristics require retuning of the control system. It is also possible that changes are short-lived, and the control system is best left alone.

In summary, it is desirable to judge whether a significant change has occurred to the process and to extract the root cause for the unusual trend. Some of this knowledge may already be embedded into the model used for control. Others may become available as more data are gathered and operating experience is obtained. A model used for control and performance monitoring/diagnosis should be flexible enough that this type of gradual model improvement is possible.

2.2. Mathematical Model. We will use the following model structure for further development:

$$x_{k+1} = f_1(x_k,u_k,p_k) + f_2(p_k)w_k \tag{1}$$

$$y_k = g_1(x_k,p_k) + g_2(p_k)v_k \tag{2}$$

In the above, x is the state vector (that includes the disturbance states), u is the known input vector, and y is the measured output vector. w and v are state and output noise vectors that account for the stochastic disturbance and measurement error. f_1 and f_2 are operators that relate the state and input vectors at 1 sample time to the state vector at the next sample time. In the case of nonlinear ODE models, they represent solution operators and may be defined only implicitly.

p is an important vector variable that requires some comments. It includes the parameters that are normally constant, but shows jumps or drifts during an abnormal operating condition. For this reason, it will be referred to as the "fault parameter" vector throughout this paper. The fault parameter vector can include process parameters, disturbance parameters, noise parameters, and those disturbances that occur infrequently but in large magnitude and lead to abnormal situations. The range of values for these parameters can be either continuous or discrete. The above model is quite general and can effectively describe most of the faults discussed earlier.

2.3. Objective. The objective is to keep track of which parameters (if any) have changed and by how much, on the basis of the output data. In other words, we would like to track jumps or drifts in p_k on the basis of y_k. This is shown pictorially in Figure 1. Clearly,

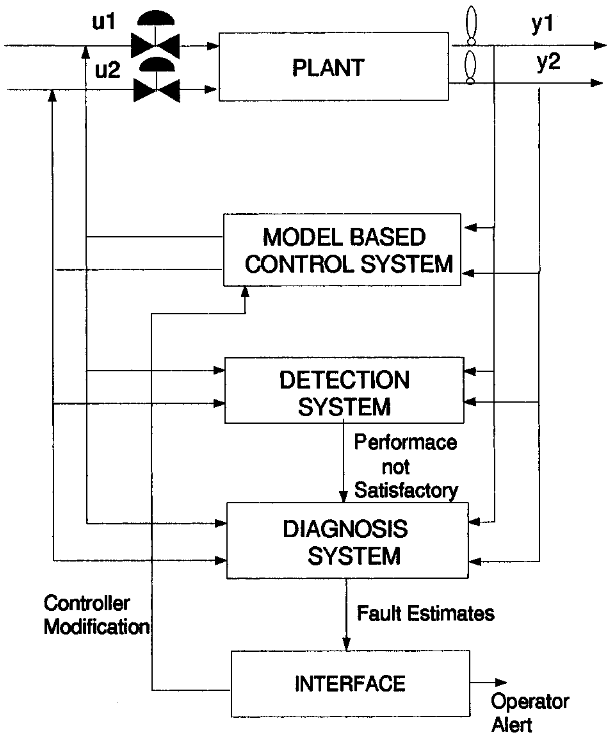

Figure 1. Structure of the proposed scheme for obtaining improved performance and guiding engineers to possible maintenance action.

this is a nonlinear estimation problem, but of a nonconventional kind. As an example, we note that the fault parameter behaviors are nonstationary and are poorly represented by usual stochastic models such as stationary Gaussian processes. This is important as the efficiency of the estimation is one of the key objectives, *i.e.*, it should not require an excessive amount of online data to converge to the new values of p_k when a jump occurs.

3. Monitoring/Diagnostic Tools for Classification and Screening

In this section, we propose simple statistical testing schemes that can be used to detect unusual data trend and obtain preliminary ideas about the potential causes responsible for it. Closed-loop input and output data are used in the testing schemes. The tests are much more informative than previously proposed tests of comparable complexity, but we emphasize that because the tests are based on the prediction error analysis, they are restricted to the case where a full dynamic model is available and used in control computation (as in model predictive control). This is in contrast to methods like the minimum variance test that can be performed with only partial process information and can be used to benchmark the performance of any controller.

3.1. Diagnostic Tool No. 1: Prediction Error Monitoring Chart. 3.1.1. Method and Justification. For model-based control systems, monitoring the prediction error signal provides useful insights into the cause for closed-loop performance degradation. This is because certain factors degrading control performance affect the model prediction error while others do not. The examples of the former kind include unmeasured actuator faults, parameter changes, sensor faults, etc. The examples of the latter kind are input saturation, measured disturbances, improper setting of control weights, etc. (for example, see Figure 4 and Figure 12 in the recent paper by Kozub, 1996). Thus monitoring the prediction error allows the control engineer to differentiate between these two types of causes. The above method, though simple, is effective in obtaining some preliminary ideas about the nature of the problem and provides useful information that can be used in further diagnostic schemes.

3.1.2. Computation of Prediction Errors. For the state space model described in eqs 1 and 2, the innovation term in the state observer (used by the model-based controller) represents the prediction error. It is well-known that the Kalman filter is the optimal state estimator for linear systems with Gaussian external inputs. For nonlinear systems, even though the Gaussian characteristics are lost due to the nonlinear dynamics, it is still common to use the Kalman filter on the basis of the model being relinearized at each time step (called the extended Kalman filter (EKF)). The prediction error and its covariance obtained from the EKF for the system of (1) and (2) (considering p_k to be constant under normal operating conditions) are given by

$$x_{k+1|k} = f_1(x_{k|k},u_k,p_k) \tag{3}$$

$$x_{k+1|k+1} = x_{k+1|k} + P_{k+1|k}G_{k+1}^{\mathrm{T}}(\Sigma_{\epsilon_{k+1}})^{-1}\epsilon_{k+1} \tag{4}$$

$$P_{k+1|k} = A_kP_{k|k}A_k^{\mathrm{T}} + f_2(p_k)Qf_2(p_k)^{\mathrm{T}} \tag{5}$$

$$P_{k+1|k+1} = (I - P_{k+1|k}G_{k+1}^{\mathrm{T}}(\Sigma_{\epsilon_{k+1}})^{-1}G_{k+1})P_{k+1|k} \tag{6}$$

where

$$\epsilon_{k+1} = y_{k+1} - g_1(x_{k+1|k},p_k) \tag{7}$$

$$\Sigma_{\epsilon_{k+1}} = G_{k+1}P_{k+1|k}G_{k+1}^{\mathrm{T}} + g_2(p_k)Rg_2(p_k)^{\mathrm{T}} \tag{8}$$

and

$$A_k = \left.\frac{\partial f_1}{\partial x}\right|_{x=x_{k|k},u_k,p_k} \tag{9}$$

$$G_{k+1} = \left.\frac{\partial g_1}{\partial x}\right|_{x=x_{k+1|k}} \tag{10}$$

Here, ϵ_{k+1} denotes the prediction error at time $k + 1$, and $\Sigma_{\epsilon_{k+1}}$ denotes its covariance. Use of other advanced state estimation techniques are certainly possible. For example, the moving horizon estimation proposed by Robertson *et al.* (1996) gives the advantage of using a nonlinear model without linearization and the ability to incorporate the bounds on the disturbance and noise signal directly into the estimation algorithm, but at the expense of substantial increase in the computational cost. The prediction error and its covariance matrix can be used along with multivariate statistical monitoring tools such as the χ^2 test, CUSUM test, EWMA test, etc. (MacGregor, 1988) to detect any abnormality.

3.1.3. Example. In this example, the usefulness of monitoring both the prediction error and the output error is demonstrated. The process is a shell and tube heat exchanger (Jacobsen and Skogestad, 1994). The nonlinear dynamic model of the plant is given by

$$\tau_{\mathrm{c}}\frac{\mathrm{d}T_{\mathrm{c}}}{\mathrm{d}t} = \frac{q_{\mathrm{c}}}{q_{\mathrm{c}}^*}(T_{\mathrm{ci}} - T_{\mathrm{c}}) + \alpha_{\mathrm{c}}(T_{\mathrm{H}} - T_{\mathrm{c}}) \tag{11}$$

$$\tau_{\mathrm{H}}\frac{\mathrm{d}T_{\mathrm{H}}}{\mathrm{d}t} = \frac{q_H}{q_H^*}(T_{\mathrm{H}i} - T_{\mathrm{H}}) + \alpha_H(T_{\mathrm{H}} - T_{\mathrm{c}}) \tag{12}$$

Table 1. Steady State Operating Conditions of the Heat Exchanger

variable	value	units
$q_c = q_H$	10	L/min
T_{ci}	25	°C
T_{Hi}	100	°C
T_c	59	°C
T_H	66	°C
$\alpha = (UA/\rho c_p)$	5	m^3/min
V	75	L

Table 2. Parameter Values of the Linear Model Around the Steady State

parameter	value
K_{11}	−2.087
K_{12}	1.74
K_{21}	−1.74
K_{22}	2.087
K_{d1}	0.54
K_{d2}	0.45
τ	7.5

where q^* denotes the nominal (steady state) flow and

$$\tau_c = \frac{V_c}{q_c^*}; \quad \alpha_c = \frac{UA}{\rho_c q_c^* c_{Pc}} \tag{13}$$

$$\tau_H = \frac{V_H}{q_H^*}; \quad \alpha_H = \frac{UA}{\rho_H q_H^* c_{PH}} \tag{14}$$

For simplicity, we choose $\alpha_c = \alpha_H$ and $q_c^* = q_H^*$. The outlet temperatures, T_c and T_H, are the controlled variables, and the feed flow rates, q_c and q_H, are the manipulated variables. The inlet coolant temperature, T_{ci}, is considered to be an unmeasured disturbance to the process. The states and parameters corresponding to the steady state operating conditions are given in Table 1.

To make the situation more realistic, we assume that the above nonlinear model is not known. Instead, the plant is controlled using a model predictive controller (MPC) with the following linear model of the plant:

$$\begin{bmatrix} T_c \\ T_H \end{bmatrix} = \frac{1}{(\tau s + 1)} \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} q_c \\ q_H \end{bmatrix} + \frac{1}{(\tau s + 1)} \begin{bmatrix} K_{d1} \\ K_{d2} \end{bmatrix} \tag{15}$$

The values of the linear process gains and time constants are given in Table 2. The parameters of the model were obtained through identification around the steady state operating condition. The model prediction and the corresponding prediction error were obtained from the Kalman filter on the basis of the state space representation of the above model (see eqs 33 and 34). The controller parameters (refer to the MPC toolbox (Morari and Ricker, 1994)) are chosen as $p = m = 5$, $uwt = [0.1\ 0.1]^T$, $ywt = [1\ 1]^T$, and $[0\ 0]^T \le u \le [15\ 15]^T$.

Case 1. In this case, there was a large load disturbance of 12° in T_{ci} (which is an unmeasured disturbance) at $t = 15$. For monitoring, we used the χ^2 testing scheme. Figure 2 shows the multivariate χ^2 monitoring charts for both the output error and prediction error with the "q in a row" threshold scheme (with $q = 5$). From Figure 2a we see that the recent output error trend is significantly higher than the normal operating condition values, indicating a degradation in the performance of the multivariable control system. However, Figure 2b shows that the prediction error still lies within the bounds, suggesting that the model is able to predict the plant output reasonably well. This in turn eliminates the possible causes such as unmeasured actuator faults, parameter changes, sensor faults, etc.

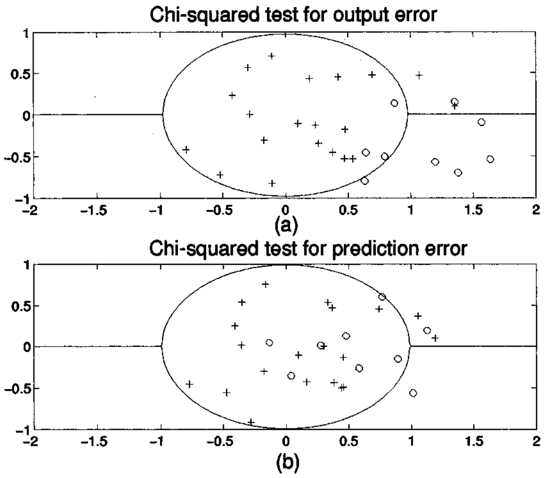

Figure 2. χ^2 monitoring charts for the case of valve saturation. Here "+" represents the data from $t = 1$ to $t = 20$, and "o" represents the data from $t = 21$ to $t = 29$.

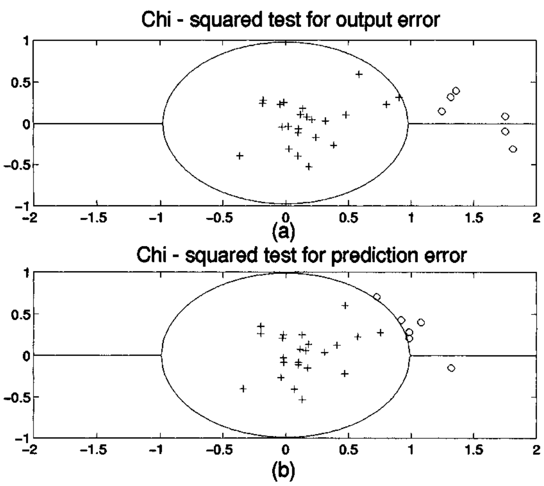

Figure 3. χ^2 monitering test for the case of stuck valve. Here "+" represents the data from $t = 1$ to $t = 20$, and "o" represents the data from $t = 21$ to $t = 29$.

from the list of potential causes. A closer look at the inputs indicated that the cold water flow rate had hit the constraint, causing the observed deviation. In this case, to improve the performance, either the valve has to be resized to allow larger flow rates or the disturbance has to be eliminated at the source.

Case 2. In this case the T_{ci} (unmeasured) changed in a ramp fashion from 25° (steady state value) at $t = 15$ to 35° at $t = 21$. Thus there was a 10° rise in the unmeasured disturbance. In addition, actuator 1 got stuck at $t = 10$ and hence was not implementing the input suggested by the controller. Figure 3 shows the monitoring plots of the output error and the prediction error. From the plots we see that there are significant deviations in both the output and prediction errors. Diagnosis of the performance degradation is more difficult in this case due to the wide array of possible reasons. These include faults in different actuators, sensors, disturbance dynamics change, changes in process parameters, etc. The diagnosis algorithm for detecting the most likely cause is discussed in the following sections.

3.2. Diagnostic Tool No. 2: Controller Detuning. 3.2.1. Method and Justification. A way to further reduce the list of potential faults consists of detuning or shutting down the controller (for example, changing the input weight in a predictive controller or switching the controller to the manual mode) and examining its effect on the prediction error trend. Although this test is intrusive, it should not cause any additional operational problems since detuning or shutting down the controller is the most common response for an operator when he/she senses an abnormal condition. Examining the effect of detuning helps in detecting the process model error regardless of the presence or absence of other types of faults.

Let us use a simple linear SISO process to explain the utility of the above method. The process is represented as

$$y_k = G(q)(u_k + u_{d_k}) + H(q)e_k + s_k \tag{16}$$

where G and H are the plant and disturbance transfer functions, u_k is the input specified by the controller, u_{d_k} is the input error signal due to an actuator fault, e_k is the combined unmeasured disturbance and noise signal, and s_k is the output error signal due to a sensor fault. For model prediction, we assume that the following model is used:

$$\hat{y}_k = \tilde{G}(q)u_k + \tilde{H}(q)\tilde{e}_k \tag{17}$$

where $\tilde{e}_k$ is an independent and identically distributed noise sequence. For the optimal predictor based on the above model, the one step prediction error is

$$\epsilon_k = \tilde{H}^{-1}(q)(y_k - \tilde{G}(q)u_k) \tag{18}$$

Using eqs 16 and 18, the prediction equation can be rewritten as

$$\epsilon_k = \tilde{H}^{-1}(q)(G(q) - \tilde{G}(q))u_k + \tilde{H}^{-1}(q)H(q)e_k + \tilde{H}^{-1}(q)G(q)u_{d_k} + \tilde{H}^{-1}(q)s_k \tag{19}$$

Assuming a control law of the form

$$u_k = -G_c(q)y_k \tag{20}$$

is used, (16) becomes

$$u_k = \frac{-G_c(q)H(q)}{1 + G_c(q)G(q)}e_k - \frac{G_c(q)G(q)}{1 + G_c(q)G(q)}u_{d_k} - \frac{G_c(q)}{1 + G_c(q)G(q)}s_k \tag{21}$$

Substituting (21) into (18) and simplifying, we get

$$\epsilon_k = \tilde{H}^{-1}(q)H(q)e_k - \left(\frac{G_c(q)}{1 + G_c(q)G(q)}e_k - \frac{\tilde{H}^{-1}(q)G_c(q)G(q)}{1 + G_c(q)G(q)}u_{d_k} - \frac{\tilde{H}^{-1}(q)G_c(q)}{1 + G_c(q)G(q)}s_k\right)\Delta G + \tilde{H}^{-1}(q)G(q)u_{d_k} + \tilde{H}^{-1}(q)s_k \tag{22}$$

where $\Delta G = G - \tilde{G}$ is the process model error. It is clear from the above equation that if $\Delta G = 0$, the prediction error will be unaffected by the controller detuning. Thus one can look at the autocorrelation function (ACF) or other appropriate functions of the prediction error before and after changing of the controller parameters. If there is no appreciable change, one can then rule out the possibility of the process parameter change. Though we have used an SISO process to make our point, the same argument holds for general multivariable processes (with a Kalman filter as the predictor). This screening tool can be very useful when the process model contains many parameters that must be reidentified. It can help avoid painstaking reidentification efforts which may be unnecessary in the first place. Another way of testing for process model error is to use setpoint perturbations as suggested by Stanfelj *et al.* (1993), but this may be less desirable in typical plant situations.

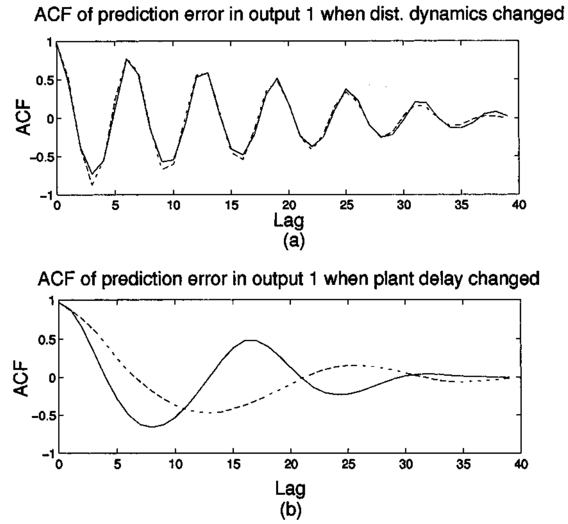

Figure 4. ACFs obtained before and after controller detuning. Here the solid line represents the ACF before controller detuning, and the dashed line represents the ACF after controller detuning.

3.2.2. Example. Consider the nonlinear model given by eqs 11 and 12 as the plant controlled by MPC using a linear model given by eq 15. A Kalman filter was used for state estimation, and the prediction error was computed as discussed in the previous section.

Case 1. In this case, let us consider the situation of change in the disturbance dynamics. The inlet coolant temperature (T_{ci}) became oscillatory at $t = 15$ sample time instants. This created oscillations in the plant outputs. In response to this the bandwidth of the controller was reduced by detuning at $t = 60$ sample time instants. This was done by increasing the input weight in MPC from [0.1 0.1] to [2 2]. Figure 4a shows the autocorrelation function (ACF) of the prediction error before and after the detuning. Since they are almost identical and fall on top of each other, it is difficult to distinguish between the two. This indicates the absence of significant error in the plant model.

Case 2. In this case, let us consider the situation of a plant parameter change. A time delay of the 4 sample time instants developed in the plant at $t = 10$ sample instants. The unmeasured disturbance (T_{ci}) had a step change of 10 at $t = 15$ sample time instants. The closed-loop response became oscillatory due to the modeling error caused by the delay change. Again, at $t = 60$, the controller was detuned by the same amount as in case 1. Figure 4b shows the ACFs of the prediction error before and after detuning. The difference between the solid and broken lines (notation as in case 1) indicates the possibility of the plant parameter change.

4. Fault Quantification

The tools introduced in the previous section provide useful insights into the nature of a performance degradation. For instance, one can determine if the cause lies in the process model or elsewhere. If the process model is found to be the source of the error, new identification can be performed. In many cases, however, reidentifying the entire model may be unnecessary as only a part of the model (e.g., the model for a particular input channel) may have gone bad. Oftentimes, there is prior knowledge as to which set of parameters are most likely to change and in what manner. Utilizing this prior knowledge can greatly reduce the identification efforts and make automatic model adaptation realizable.

In this section, we discuss a way to identify and quantify fault(s) from a number of potential causes, using state estimation. Given the usual scenario of a large number of possible faults and the limited number of measurements, proper definition and utilization of prior statistics about the different faults are very important for efficient estimation of the fault. In this work, we show how to incorporate the prior statistics effectively into the fault model. The resulting state estimation problem has been well studied, and a large body of literature exists on the subject (for example, Buxbaum and Haddad, 1969; Jaffer and Gupta, 1971). Unfortunately, the optimal estimator requires a growing number of filters and is therefore computationally intractable. We develop a suboptimal estimation scheme by combining the different approximations available in the literature.

4.1. Statistical Model for Fault Parameter Vector. In Bayesian estimation, we need to specify the prior probability distribution of the fault parameter vector sequence. The efficiency of the estimation will depend critically on how well this is done. We model the fault parameters using the stochastic difference equations

$$\begin{aligned} z_{k+1} &= A_f z_k + B_f e_k \\ p_k &= C_f z_k + D_f e_k \end{aligned} \tag{23}$$

where e_k is an *independent, identically distributed* vector sequence. Each element of e_k corresponds to a particular fault, and therefore e_k will be referred to as the "fault signal vector" throughout this paper. It is important to understand that $e_k \neq 0$ implies that a particular fault/parameter change has occurred at $t = k$. To facilitate the exposition, we assume that $A_f = I$ and $D_f = 0$ throughout this paper (B_f and C_f can be used to model the correlated changes). It is easy to see that this choice corresponds to the case where parameters show jumps, *i.e.*, whenever e_k is nonzero, one or more elements of p_k jump to another value. It is straightforward to model other types of changes such as impulses or drifts by choosing different A_f.

Once A_f, B_f, C_f, and D_f are fixed, the distribution of the fault signal determines the prior probability distribution assigned to the fault parameter vector sequence. In order to assign appropriate statistics of e_k, we need to examine the characteristics of faults. The common features are as follows:

(a) Typical faults occur infrequently, but in large magnitude when they do. (b) Single faults are much more likely than simultaneous faults. An important point to recognize here is that the common practice of modeling the external signal (e_k in our case) as a joint Gaussian signal and using the corresponding optimal filter is not appropriate here. The main problem with assigning joint Gaussian statistics to e_k is that the bell-shaped curves are ill-suited to model the above-listed characteristics of a typical fault signal. First of all, it is difficult to find a bell-shaped curve that fits the characteristics of "infrequent, but large in magnitude". Clearly, one has to give up one or the other or settle with a compromise. In addition, it is difficult to shape joint Gaussian density functions to reflect the fact that individual faults are much more likely than simultaneous faults. Finally, it is impossible to model fault signals as signals having discrete ranges or multimodal distributions. These observations are demonstrated for the two-dimensional case in Figure 5. In the figure, one can clearly see the difficulty of choosing a bell-shaped curve that fits the typical probability density of a fault signal. If Gaussian statistics are insisted, one must resort to using a distribution of large covariance. This essentially increases the search space and decreases the efficiency of the estimator. A practical implication is that it may take inordinate amount of measurements before the filter estimates converge to the true value. The wide prior search space may also cause the filter to distribute the errors among different fault signals, thereby making the isolation of the true fault impractical. In addition, due to the large covariance, the estimates will be very sensitive to measurement errors.

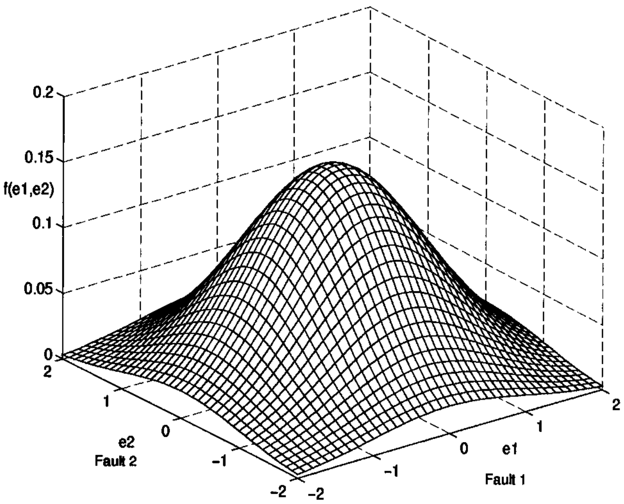

Figure 5. Joint Gaussian density vs density functions of a typical fault signal.

A better way to model the fault signal is by using a weighted sum of Gaussian densities. As a first step in obtaining an appropriate Gaussian sum, one specifies the probabilities corresponding to different fault scenarios as listed below:

case 1: occurrence of no fault with probability (w.p.) $= p_1$

case 2: occurrence of fault scenario 2 w.p. $= p_2$

$\vdots$

case N: occurrence of fault scenario N w.p. $= p_N$

where N is the total number of fault scenarios that are considered. Fault scenarios include all the individual and simultaneous faults that can possibly occur. p_i is the prior probability that is assigned to the corresponding fault scenario and can be decided on the basis of preliminary diagnosis, past operating records, and past experience. Once the probabilities for individual fault scenarios are assigned, each fault scenario is modeled according to (23). The fault signal e_k for the ith scenario is a Gaussian i.i.d. sequence of mean $\bar{e}_i$ and covariance Σ_i^e. Hence, when $\bar{e}_i$ and Σ_i^e are assigned, the mean and variances for those elements of e_k corresponding to the ith fault are changed from their normal values.

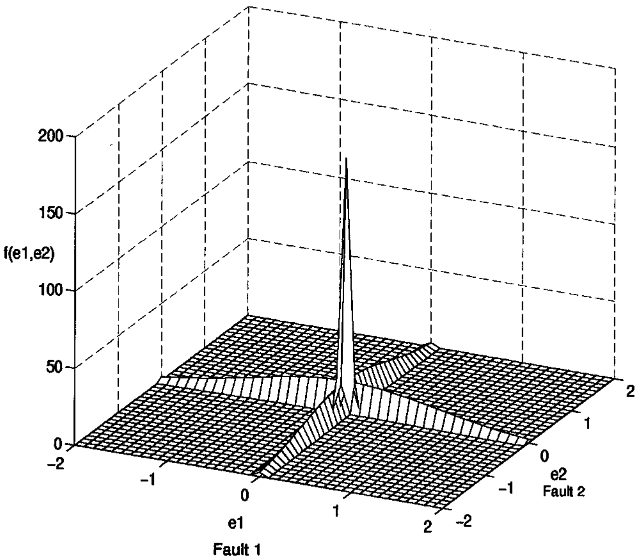

Figure 6. Non-Gaussian density obtained by incorporating prior knowledge.

The overall prior density function is then given by the sum of the different Gaussian density functions weighted by the corresponding prior probabilities. Thus

$$e_k \approx \sum_{j=1}^{N} p_j N(\bar{e}_j, \Sigma_j^e) \tag{24}$$

where $N(\bar{e},\Sigma)$ denotes a joint Gaussian density function with mean $\bar{e}$ and covariance Σ. Note that the above results in a non-Gaussian probability density function for e_k.

Example. Consider the case where there exist two possible faults. Each fault is independent, and the probability for their occurrence is 0.1. Then, at each time instant, the possible fault scenarios could be (1) no fault w.p. = 0.81, (2) occurrence of the first fault alone w.p. = 0.09, (3) occurrence of the second fault w.p. = 0.09, and (4) occurrence of both the faults w.p. = 0.01. The fault signal vectors for the four fault scenarios are modeled using Gaussian noise of mean zero and the following covariance matrices: Σ_0^e = diag[$1e^{-2}$ $1e^{-2}$], Σ_1^e = diag[1 $1e^{-2}$], Σ_2^e = diag[$1e^{-2}$ 1], and Σ_3^e = diag[1 1]. The resulting non-Gaussian joint probability density function is shown in Figure 6. From this figure we see that while the probabilities for individual faults especially at larger magnitudes are significant, the joint probabilities for simultaneous faults are very low. A bell-shaped curve cannot approximate this density closely. Although not demonstrated here, the mean of various Gaussian distributions can be adjusted to create a multimodal distribution, if needed.

4.2. Optimal Estimation. The fault model of (23) where e_k is defined through (24) can be augmented to the process models (1) and (2). The next step is to perform the state estimation. The optimal state estimation result is available in the literature (for example, Buxbaum and Haddad, 1969), but is summarized here for the sake of completeness as most readers should be unfamiliar with the result.

In order to handle the non-Gaussian prior statistics modeled as a weighted sum of a finite number of Gaussian sequences, we branch into different filters at each instant of time, each branch corresponding to a particular fault scenario as shown in Figure 7. Let γ_k be an indicator variable that can be used to indicate the occurrence of a particular fault at time k ($\gamma_k \epsilon \{1, 2, ..., N\}$). Thus $\gamma_k = i$ indicates that the ith fault scenario has occurred at the kth time instant. Let us denote the initial branching time as the zeroth time instant. Then, we define the *fault sequence* as $\Gamma_k = \{\gamma_0, \gamma_1, ..., \gamma_k\}$. We also use S^k to denote the set of all possible fault sequences at time k (hence, $\Gamma_k \epsilon S^k$). Note that each fault sequence Γ_k leads to a particular Gaussian system, and the standard Kalman filter provides the optimal estimate for each branch.

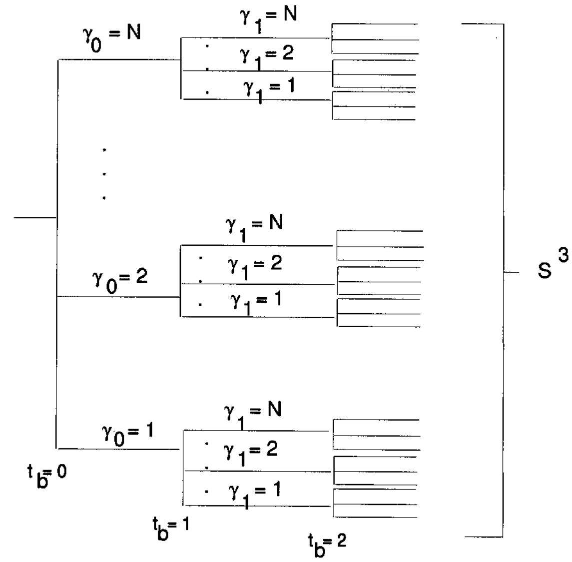

Figure 7. Branching of the filters in optimal multiple filter approach.

The question is how to combine the estimates from different branches into one. In Bayesian estimation, the posterior probability of each fault sequence ($P_b(\Gamma_k|Y_k)$) is first computed according to

$$P_b(\Gamma_k|Y_k) = \frac{\rho(y_k|\Gamma_k,Y_{k-1})P_b(\Gamma_k|Y_{k-1})}{\sum_{\Gamma_k}\rho(y_k|\Gamma_k,Y_{k-1})P_b(\Gamma_k|Y_{k-1})} \tag{25}$$

where P_b is the probability, $Y_k = \{y_1, y_2, ..., y_k\}$, and $\rho(y_k|\Gamma_k,Y_{k-1})$ is the likelihood of the current measurement given a particular fault sequence and the previous set of measurements. The likelihood function $\rho(y_k|\Gamma_k,Y_{k-1})$ can be obtained by using the estimate, $X_{k|k-1}$ (where X denotes the combined state x and fault parameter p vector), and covariance, $P_{k|k-1}$, that result from the Kalman filter for Γ_k, as follows:

$$\rho(y_k|\Gamma_k,Y_{k-1}) = \frac{1}{\sqrt{(2\pi)^{n_y}\det(\Sigma_{\epsilon_k})}}\exp\{-0.5(y_k - g_1(X_{k|k-1}(\Gamma_k)))^{\mathrm{T}}\Sigma_{\epsilon_k}^{-1}(y_k - g_1(X_{k|k-1}(\Gamma_k)))\} \tag{26}$$

where

$$\Sigma_{\epsilon_k} = G_k P_{k|k-1} G_k^{\mathrm{T}} + g_2(p_{k|k-1}) R g_2(p_{k|k-1})^{\mathrm{T}} \tag{27}$$

The conditional probability $P_b(\Gamma_k|Y_{k-1})$ is given by

$$P_b(\Gamma_k|Y_{k-1}) = P_b(\Gamma_{k-1}|Y_{k-1})P_b(\gamma_k) \tag{28}$$

where $P_b(\gamma_k)$ is the prior probability corresponding to the assumed fault at time k.

Once the posterior probability for each branch is computed, the optimal estimate can be obtained by (1) choosing the estimate for the branch with the maximum probability, leading to the maximum aposteriori estimate, or (2) computing the weighted sum of each estimate (weighed with the posterior probability), leading to the minimum mean-square estimate.

Note that one can run the multiple filters at all time instances and use the estimates of the faults to decide on the presence or absence of the faults. *This excludes the necessity of using additional monitoring tools such as those discussed in section* 3. On the other hand, use of simple monitoring/diagnostic tools introduced earlier eliminates a number of potential scenarios and obviates the need for running the multiple filters at all time instants, thereby reducing the on-line computational requirement.

4.3. Practical Implementation of Multiple Filters. The optimal multiple filter (MF) approach as discussed earlier requires branching into different filters at each instant of time. This causes an exponential growth in the number of filters, with time as shown in Figure 7 (for example, in M time steps after the initial branching, the number of filters reach N^M). Thus the optimal estimation becomes computationally intractable. In addition, this results in too many similar fault sequences, making the convergence of the probability slow.

Several suboptimal approaches such as the maximum likelihood method (Buxbaum and Haddad, 1969), fusion or pseudo-Baye's algorithm (Buxbaum and Haddad, 1969; Jaffer and Gupta, 1971; Tugnait and Haddad, 1979), hypothesis test algorithm (Buxbaum and Haddad, 1969), infrequent branching algorithm (Willsky, 1976), etc. have been proposed to contain the growth of the filters. The usefulness of a particular algorithm is problem-dependent. In this paper, we propose a suboptimal algorithm which combines several different approximations. Parameters for each approximation can be adjusted to choose a right balance between accuracy and computational ease for a given problem.

4.3.1. Infrequent Branching. The MF diagnostic scheme is started only when the prediction error exhibits a statistically significant deviation from its normal value. At this time branching into different filters is done. After the branching takes place, the main reason for the quick increase in the number of filters in the optimal estimator is the branching into different possible cases at *each* instant of time as shown in Figure 7. However, on the basis of the characteristics of faults, we know that faults are infrequent in nature. In addition, the exact timing of the fault occurrence should not matter much in most cases. Hence, after the initial branching, further branching is done only after M time steps. Between the branching time instances, Σ^e is kept at a constant value, the choice of which involves a trade-off between efficient estimation of the faults and sensitivity to measurement noise. Thus the number of filters after 1 horizon of M time steps are only N^2 as compared to N^M for those in the optimal estimation. The fault sequence Γ_k is altered to reflect this. Call the new fault sequence corresponding to the infrequent branching as $\tilde{\Gamma}_k$. Thus after 1 horizon length, $\tilde{\Gamma}_k$ is a member of only S^2 (note that, in optimal filtering, after 1 horizon length of M time instants, Γ_M is a member of S^M). The choice of M represents a trade-off between computational complexity and the ability to quickly diagnose the occurrence of more than one fault within a short span of time.

4.3.2. Maximum Likelihood Method. After each branching, we keep monitoring the posterior probabilities of the different sequences. If the ratio of the maximum posterior probability to the second maximum posterior probability is greater than 20:1 for l consecutive time instants, then that fault sequence is the most likely sequence. We retain this sequence and terminate the other sequences (hence the name maximum likelihood method). Notice that this is similar to the q in a row concept widely used in statistical process control (SPC) monitoring charts. We observed that the choice of l (for q in a row) is not crucial, especially for linear systems. Since in nonlinear systems we have only approximate estimates of the states and covariance, in order to add robustness to the suboptimal algorithm, the q in a row concept can be used. As a rule of thumb, the value of l can be chosen as 5 in most cases. Now, when the most likely sequence is chosen and the other sequences are cut off, we have only one filter. The prediction error is again reduced because of the new fault and state estimates. We now start monitoring the prediction error again to detect the occurrence of further faults.

4.3.3. Branch Cutting. If the maximum likelihood condition is unable to cut the branches, the number of filters tend to increase with the branching. In order to curtail the number of filters, we proceed toward reducing the filter dimension by eliminating the fault sequences that have very low probability of occurrence. To demonstrate this, let us consider the simple case when each fault has probability of occurrence α at each time instant. Let $\delta_{i,j}\epsilon\{0,1\}$ be a variable denoting the absence or presence of the jth fault at the ith branching time instant. Note that $\Sigma_j\delta_{i,j}$ gives the total number of faults that have occurred at the ith branching instant. The probability of the total number of faults being less than or equal to m in say n branching time instances is given by (Robertson *et al.*, 1995):

$$P_{\mathrm{b}}\Big(\sum_{i=0}^{n}\sum_{j=1}^{N_{\mathrm{f}}}\delta_{i,j} \le m\Big) = \sum_{k=0}^{m}\binom{nN_{\mathrm{f}}}{k}(1-\alpha)^{(nN_{\mathrm{f}}-k)}\alpha^{k} \quad (29)$$

where N_{f} is the number of faults. By choosing the probability to be greater than a threshold ζ, we can decide on the value of m (the total number of faults to be allowed within the given number of branching n). The resulting number of sequences is given by

$$\sum_{k=0}^{m}\binom{nN_{\mathrm{f}}}{k} \quad (30)$$

4.3.4. Fusion. Finally, to bound the number of filters, we employ the fusion or the pseudo-Baye's algorithm given in the literature (Buxbaum and Haddad, 1969; Jaffer and Gupta, 1971; Tugnait and Haddad, 1979). It involves combining those filters which make identical assumptions about the occurrence of the fault (and thus about the state noise density function) in the last, say d branching instants. The rationale behind this approach is that the Kalman filters corresponding to different initial estimates and covariance matrices converge to the same estimates and covariance asymptotically (Buxbaum and Haddad, 1969). Now, let d be the fusion horizon, *i.e.* after d branchings the fusion should take place (note that since the branching occurs only once in M sample instants, dM sample time instants would have passed before the fusion starts). The fault sequence $\tilde{\Gamma}_d = \{\gamma_0, ..., \gamma_d\}\epsilon S^d$. By fusing the filters which differ only in the assumption made at 0th instant, we get rid of γ_0. Hence, the new fault sequence after fusion is given by $\tilde{\Gamma}_d^1 = \{\gamma_1, ..., \gamma_d\}\epsilon S^d$. The

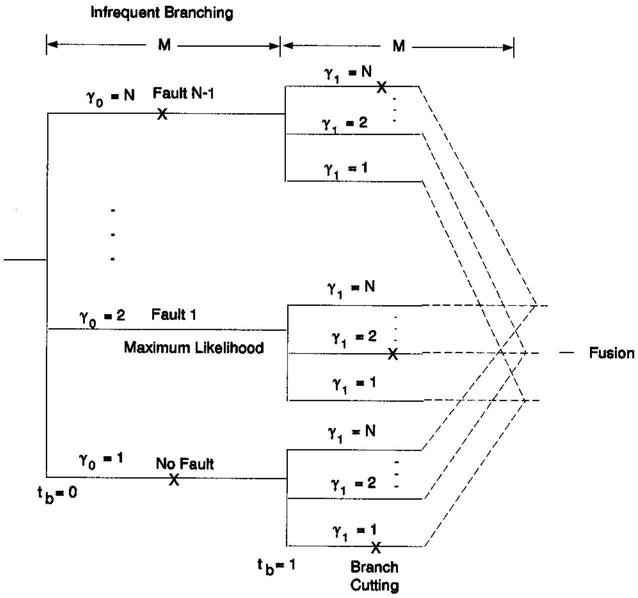

Figure 8. Branching and fusion in the suboptimal multifilter approach.

equations for the state estimates and covariance for the fused sequences are given by

$$\hat{X}_{k|k}(\tilde{\Gamma}_d^1) = \sum_{\gamma_0} \frac{\hat{X}_{k|k}(\tilde{\Gamma}_d^1,\gamma_0) P_b(\tilde{\Gamma}_d^1,\gamma_0|Y_k)}{\sum_{\gamma_0} P_b(\tilde{\Gamma}_d^1,\gamma_0|Y_k)} \tag{31}$$

$$P_{k|k}(\tilde{\Gamma}_d^1) = \sum_{\gamma_0} (P_{k|k}(\tilde{\Gamma}_d^1,\gamma_0) + (\hat{X}_{k|k}(\tilde{\Gamma}_d^1,\gamma_0) - \hat{X}_{k|k}(\tilde{\Gamma}_d^1))(\hat{X}_{k|k}(\tilde{\Gamma}_d^1,\gamma_0) - \hat{X}_{k|k}(\tilde{\Gamma}_d^1))^{\mathrm{T}}) \frac{P_b(\tilde{\Gamma}_d^1,\gamma_0|Y_k)}{\sum_{\gamma_0} P_b(\tilde{\Gamma}_d^1,\gamma_0|Y_k)} \tag{32}$$

where $\hat{X}_{k|k}(\tilde{\Gamma}_d^1,\gamma_0)$ denotes the state estimate obtained from the filter corresponding to the fault sequence $(\tilde{\Gamma}_d^1,\gamma_0)$ and $P_{k|k}(\tilde{\Gamma}_d^1,\gamma_0)$ is the corresponding covariance of the estimate. Note that if there were N^{d+1} sequences before fusion, then we will have only N^d sequences after fusion. The choice of fusion horizon (after which the estimates are to be fused) depends on the computational complexity and the amount of memory required by the filters to estimate the states.

The overall suboptimal algorithm is shown in Figure 8.

4.4. Controller Modification. The fault/parameter estimates can be used to make appropriate changes in the control algorithm (if possible) to handle the faults. For example, in the case of actuator fault such as a bias (due to a change in actuator zero or valve leak, etc.), the bias value can be passed on to the controller to modify the input moves. In the case of a failed (stuck) valve, the controller should be informed about the unavailability of this input for control purposes (loss in degree of freedom). In the MPC framework, this can be done by either setting the δ constraint corresponding to this input as zero or increasing the corresponding input weight. When the parameter changes, the new parameter estimates from the diagnosis scheme can be passed on to the controller. Thus the model used for control is adapted to reflect the true process output, resulting in a better performance. At this point, note that changes in the disturbance dynamics, a common scenario in process industries, can also be handled in this framework by adapting the disturbance model parameters. If additional tests are necessary for parameter identification, the specifics of the needed tests can be provided. In the case of sensor faults such as bias, the estimate of the bias can be used by the controller to take a corrective action until the sensor is recalibrated.

4.5. Comparison to Other Approaches. 4.5.1. Parameter Estimation Techniques. One of the oldest and most widely used techniques for fault detection is to estimate the parameters of the process using a single Kalman filter or an extended Kalman filter (Himmelblau, 1978; Isermann, 1984; Watanabe and Himmelblau, 1983). If the new parameter estimates are significantly different from the old estimates, a fault is said to have occurred. However, when the number of possible faults in the system is large (typical situation in practice), application of these techniques for estimation leads to slow convergence and distribution of the error among different faults as discussed earlier. This makes fault isolation practically impossible.

4.5.2. Parallel Filter Algorithm. The MF approach should be distinguished from the parallel filter approach that has been used in the past for disturbance detection (Ku *et al.*, 1994; King and Giles, 1990). The parallel filter approach considers a fixed number of scenarios and assumes that a *single scenario* will be able to explain the entire set of measurements. In order to identify the fault scenario, a Kalman filter corresponding to each fault is run with *constant state noise covariance matrices*. For example, a Gaussian density of covariance Σ_i can be used to model the ith fault after the branching time instant and so on. Posterior probability for the filter corresponding to each fault is calculated as before, and the filter with the maximum probability is chosen. This method is useful when one is interested in identifying the fault when it occurs and correcting it before proceeding further. This is unsuitable for detecting faults that occur consecutively, since there is no provision for switching between faults. In addition, since the parallel filter has high filter gains at every instant of time (constant covariance matrices), the estimates can be very sensitive to measurement noise as compared to MF estimates. The insensitivity of the MF estimates to noise is due to the time-varying covariance matrix used by the filters constituting it (again due to the non-Gaussian density assumption). Finally, the parallel filter approach can be interpreted as a special case of MF approach with the branching horizon size M to be chosen as infinity. Clearly, the MF approach is a much more general framework.

4.5.3. Generalized Likelihood Ratio (GLR) Algorithm. The GLR algorithm for detecting and estimating faults and jump parameter changes was proposed by Willsky and Jones (1974). This method determines the maximum likelihood estimate (MLE) of the time of occurrence of the fault (θ) and the magnitude of the fault (v). This is done by first determining the MLE of v for all possible θ (from $\theta = 1$ to $\theta = k$) by using the measurements from $t = \theta$ to $t = k$, where k is the current time instant. Then the MLEs, $\hat{\theta}$ and $\hat{v}$, are chosen as those that maximize the conditional density of the measurements given θ and the corresponding $\hat{v}$. A ratio between the likelihood obtained by using the MLEs and the null hypothesis (no fault assumption) is taken. If the ratio exceeds a threshold limit, then a fault is deemed to have occurred. Note that a full implementation of the GLR involves a growing bank of filters due to the necessity for estimating v corresponding to different possible θ. The following remarks can be made about GLR in comparison to the MF approach: Firstly, unlike the MF approach, the GLR algorithm does not utilize the prior knowledge that is available about the faults and parameters. Due to this, the GLR method could require a much large of number of measurements for isolating the actual fault when compared to those for MF. Secondly, the GLR method also has the same disadvantage as the parallel filter approach in detecting consecutive faults since the switching between faults is not considered.

4.5.4. Other Methods. Other analytical redundancy approaches such as the detection filter approach (Willsky, 1976) and the local filter approach (Fathi *et al.*, 1993) have also been suggested in the literature. The detection filter approach isolates faults to particular directions and estimates them. This approach, however, is restricted to linear time invariant systems. The local filter approach constructs several local filters at different fault locations. A sequential probability ratio test then determines the presence or absence of a particular fault. This test was used in conjunction with a knowledge-based system. In multivariable systems with sufficient interaction, it is difficult to construct independent local filters corresponding to different faults. Other well-studied approaches include designing a four parameter controller for combined fault detection and control (Ajbar and Kantor, 1993; Nett *et al.*, 1988). In these approaches, a fault detection scheme for input/output faults is designed together with the controller since their performances tend to be interdependent. However, the detection scheme in these methods is restricted to a linear structure. This can be very limiting in light of the fact that the nature of most fault estimation problems is nonlinear. In addition, fault signals are assumed to be additive and norm-bounded. In other words, not only are the types of faults restricted but the types of prior knowledge that can be incorporated are also limited. These limitations make it practically impossible to design an effective diagnosis scheme for general multivariable systems. For linear deterministic systems, the unknown input observer has been proposed by Patton *et al.* (1989). However, the extension of this design to general nonlinear multivariable systems has found limited success due to the complex nature of the problem.

4.6. Practical Considerations. We would like to make the following remarks for the user in practical situations: (a) A practical issue that one might face when implementing the MF method on-line is the effect of nonlinearity of the process. When the controlled process is highly nonlinear and a linear model around a particular operating point is used for control purposes, the occurrence of a sudden large change/fault can move the plant to a totally different operating regime. If the controller is in the *auto* mode, the modeling error due to the operating regime change could cause the inputs to fluctuate wildly, making the estimation of the linear model parameters very difficult. Under these circumstances, it might be better to put the controller in the *manual* mode and run the diagnosis algorithm with prespecified input variations to identify the parameters/faults.

(b) Model errors are always present, but difficult to handle rigorously. Perhaps the biggest obstacle is the usual lack of precise quantitative information on model errors. Traditionally, some state noise is added to the model to counteract the model error effect. This adds robustness to the algorithm in general. This is seen in the heat exchanger example where a linear model developed around a particular operating point is used with some state noise under the normal conditions.

(c) The efficiency of the algorithm is intimately related to the prior knowledge used in decision making. For instance, if disturbances are modeled as independent signals entering the output channels (as often done in practice), one would have a very difficult time distinguishing between the sensor errors and disturbances *i.e.*, one has to rely entirely on the signal shape. On the other hand, if correlations among disturbances in different output channels are modeled, one would have an increased chance of discriminating between the two.

(d) The occurrence of a new *unmodeled* fault, in general, leads to a distribution of the posterior probabilities among several of the modeled faults, resulting in indecisiveness. When this happens, the user can get an idea that a new type of fault has occurred. In this case, the algorithm serves as a starting point for the user by eliminating all the modeled faults as not being responsible for the performance degradation.

5. Example: Heat Exchanger

5.1. Simulation. The plant and model equations for this example are given by eqs 11, 12, and 15, respectively. The inlet temperature of the coolant (T_{ci}) was regarded as an unmeasured disturbance input to the plant. The state space representation of (15) can be written as

$$\begin{bmatrix} \dfrac{dT_c}{dt} \\ \dfrac{dT_H}{dt} \end{bmatrix} = \begin{bmatrix} \dfrac{-1}{\tau} & 0 \\ 0 & \dfrac{-1}{\tau} \end{bmatrix} \begin{bmatrix} T_c \\ T_H \end{bmatrix} + \begin{bmatrix} \dfrac{K_{11}}{\tau} & \dfrac{K_{12}}{\tau} & \dfrac{K_{d1}}{\tau} \\ \dfrac{K_{21}}{\tau} & \dfrac{K_{22}}{\tau} & \dfrac{K_{d2}}{\tau} \end{bmatrix} \begin{bmatrix} q_c \\ q_H \\ T_{ci} \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad (33)$$

$$y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} T_c \\ T_H \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad (34)$$

where $[w_1\ w_2]^T$ are Gaussian white noise sequences representing the state noise and $[v_1\ v_2]^T$ represents the measurement noise. The discrete version of the above model can be represented as

$$x_{k+1} = f(x_k, \theta_k, u_k + u_{d_k}) + w_k \quad (35)$$

$$y_k = g(x_k) + s_k + v_k \quad (36)$$

where $x = [T_c\ T_H]^T$, θ, u, and y represents the model states, plant parameters (steady state gains and time constant), and inputs and outputs of the plant, respectively. f represents the solution operator for the differential equations over 1 sampling period. In this case g is linear. u_d and s are fault parameters and are included in the model to account for the input and output faults, respectively.

The unmeasured disturbance input (T_{ci}) can be modeled as

$$T_{ci_k} = \frac{1}{(1 - q^{-1})(1 - \beta q^{-1})} w'_{k-1} \quad (37)$$

where q represents the difference operator such that, $q^{-1}w'_k = w'_{k-1}$, w' is a white noise sequence and β is a disturbance parameter such that $0 \le \beta \le 1$. Note that when $\beta = 0$, the disturbance model is an integrated white noise model which is useful in modeling random step type disturbances. When $\beta = 1$, the disturbance model is a double-integrated white noise model which is useful in modeling drifting type disturbances. When

$0 < \beta < 1$, the model is an integrated white noise filtered through first-order dynamics. Hence varying the disturbance parameter β provides the flexibility in adapting the model for change in disturbance dynamics. The state space representation of the disturbance model can be written as

$$x_{d_k} = \beta x_{d_{k-1}} + w'_{k-1} \tag{38}$$

$$T_{\mathrm{ci}_k} = T_{\mathrm{ci}_{k-1}} + x_{d_{k-1}} \tag{39}$$

The above disturbance states can be augmented with the model states in (35). In terms of the notations in (1) and (2), the fault parameter vector p, input u, and output y are given by $p = [\theta^{\mathrm{T}}\ \beta\ u_d^{\mathrm{T}}\ s^{\mathrm{T}}]^{\mathrm{T}}$. $u = [q_{\mathrm{c}}\ q_H]^{\mathrm{T}}$ and $y = [T_{\mathrm{c}}\ T_{\mathrm{H}}]^{\mathrm{T}}$.

Now, the potential faults that we consider in this example are a decrease in the heat transfer coefficient due to fouling, a change in disturbance dynamics, faults in the two actuators, and faults in the two sensors. There are 5 parameters in θ (4 process steady state gains and 1 time constant). When the heat transfer coefficient changes (parameter α in (11) and (12)), the steady state gains change in a correlated manner. Examining the expressions for the steady state gains obtained by linearizing the nonlinear plant (11) and (12), we can obtain a preliminary idea about the correlation between the gain changes in response to the change in the heat transfer coefficient (α). For instance, if K_{11} decreases, K_{22} increases by the same amount and vice versa. The same is true for K_{12} and K_{21}.

On the basis of these considerations, we model the fault parameter vector as (compare with equation 23 with $A_{\mathrm{f}} = C_{\mathrm{f}} = I$)

$$p_{k+1} = p_k + B_{\mathrm{f}} e_k \tag{40}$$

where

$$B_{\mathrm{f}} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} & & & \\ & 1 & & \\ & & I_2 & \\ & & & I_2 \end{bmatrix} \tag{41}$$

Hence, e_k has eight elements of which the first three correspond to the heat transfer coefficient change, the fourth corresponds to the disturbance dynamics change, the fifth and sixth correspond to the actuator faults, and the final two correspond to the sensor faults. We note that even though all five elements of the parameter vector θ are triggered by the same change and therefore are completely correlated, we chose to model it as three independent changes instead because of the nonlinearity.

The above fault parameter model can now be augmented with the plant and disturbance states to obtain the overall model of the system. Denote the overall state vector as $X = [x^{\mathrm{T}}\ [x_d\ T_{\mathrm{ci}}]\ p^{\mathrm{T}}]^{\mathrm{T}}$. Note that considering the plant parameters (θ) and disturbance parameter (β) as states (included in p) makes the overall model nonlinear. Hence we use an EKF for estimating the states under normal conditions and multiple EKFs for fault diagnosis. The state noise covariance matrix Q of the augmented state X is

$$Q = \begin{bmatrix} 0.0025 I_2 & & & \\ & 0.4 & & \\ & & 1e^{-10} & \\ & & & \Sigma^e \end{bmatrix}$$

Under normal operating conditions, since the parameters were treated as constant, we set $\Sigma^e \approx 0$.

Monitoring tool no. 1 was used to detect abnormalities. Since the q in a row scheme was used in the monitoring tool to detect the abnormality, once the abnormality is detected, we go H ($H \geq q$) time steps backward and start branching into multiple filters for detailed diagnosis. At the branching instances, the fault signal e was modeled as a Gaussian sum with the prior probabilities and covariance matrices shown in Table 3, and between the branching time intervals, the covariance was kept constant at Σ_i^e. Note that by dividing the covariance by the horizon length M, we distribute the covariance to the entire branching horizon. This method was useful in detecting faults that occur within the branching horizon.

For monitoring tool no. 1 the value of q was taken as 5 and H (the number of time steps we go backward in time once abnormality is detected) was taken as 10. For diagnosis with the suboptimal multiple filter (MF), the parameters of the suboptimal filter were taken as $M =$ 5 sample time instants, $l = 5$, $\zeta = 0.995$, and $d = 3$. The value of m from eq 29 was 2. We also do not consider the occurrence of simultaneous faults within a single horizon (5 sample time instants). Hence, the resulting number of sequences from the branch cutting will only be

$$\sum_{k=0}^{2} \binom{3 \times 6}{k} - 3\binom{6}{2} = 127 \tag{42}$$

This was the maximum number of filters that were used.

5.2. Results and Discussion. Case 1: Actuator Bias. Figures 9 and 10 show the outputs of the diagnostic scheme and the system outputs when a bias of magnitude -4 occurred at $t = 15$ in the actuator, manipulating the cold water flow rate (q_{c}). At $t = 22$, monitoring tool no. 1 indicated an abnormality. Hence the first branching into 7 filters corresponding to different fault scenarios started at $t = 12$ (since $H =$ 10), and additional branchings were done at subsequent time instances. The number of filters ranged between 1 and 127, with 127 filters being used for only 10 time instants. From $t = 55$, only 1 filter was present and the prediction error decreased below threshold value. The monitoring algorithm was started again to detect further abnormalities. In the plot of the actuator 1 fault and its estimates, the solid line (—) denotes the actual fault signal (u_{d1}) caused by the actuator bias. The estimates of different faults obtained with the suboptimal multiple filter (dark dotted line) and that obtained with a single filter and covariance resetting (dashed line) are shown. For the covariance resetting (CR) algorithm, the noise covariance matrix at the time of resetting was taken as

$$\Sigma^e = \mathrm{diag}[5 \quad 5 \quad 1 \quad 1 \quad 25 \quad 25 \quad 25 \quad 25] \tag{43}$$

As can be seen from the plots, the CR algorithm results in a distribution of errors among different faults making the fault isolation infeasible whereas the MF algorithm is able to prevent such bias distribution due to the efficient use of prior knowledge. The bias value is

Table 3. Parameters of the Noise Sequence with Non-Gaussian Probability Density Function

fault	prior probability	covariance for state noise
no fault	0.9	$\Sigma_1^e = \mathrm{diag}[1e^{-6} \times \mathrm{ones}(1, 8)]^a$
process parameter change	0.017	$\Sigma_2^e = \mathrm{diag}[5/M\ 5/M\ 1/M\ 1e^{-6} \times \mathrm{ones}(1, 5)]^a$
disturbance parameter change	0.017	$\Sigma_3^e = \mathrm{diag}[1e^{-6} \times \mathrm{ones}(1, 3)\ 5/M\ 1e^{-6} \times \mathrm{ones}(1, 4)]^a$
actuator 1 fault	0.017	$\Sigma_4^e = \mathrm{diag}[1e^{-6} \times \mathrm{ones}(1, 4)\ 25/M\ 1e^{-6} \times \mathrm{ones}(1, 3)]^a$
actuator 2 fault	0.017	$\Sigma_5^e = \mathrm{diag}[1e^{-6} \times \mathrm{ones}(1, 5)\ 25/M\ 1e^{-6} \times \mathrm{ones}(1, 2)]^a$
sensor 1 fault	0.017	$\Sigma_6^e = \mathrm{diag}[1e^{-6} \times \mathrm{ones}(1, 6)\ 25/M\ 1e^{-6}]^a$
sensor 2 fault	0.017	$\Sigma_7^e = \mathrm{diag}[1e^{-6} \times \mathrm{ones}(1, 7)\ 25/M]^a$

[a] Here ones(1, n) means $\underbrace{(1, 1, ..., 1)}_{n}$

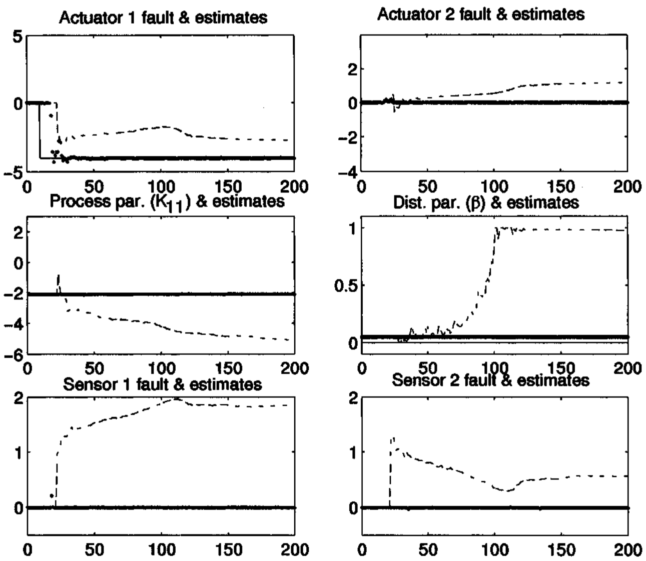

Figure 9. Estimates of faults with multiple filter (dark dotted) and covariance resetting (dashed) algorithms when there was a bias in actuator 1. The true fault signal is represented by a solid line.

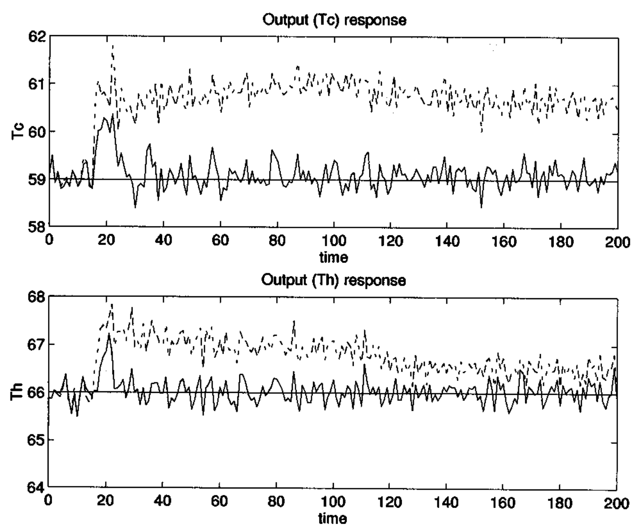

Figure 10. Closed loop output responses, when there was a bias in acutator 1, using multiple filter (solid) and covariance resetting (dashed) algorithms for estimation and diagnosis.

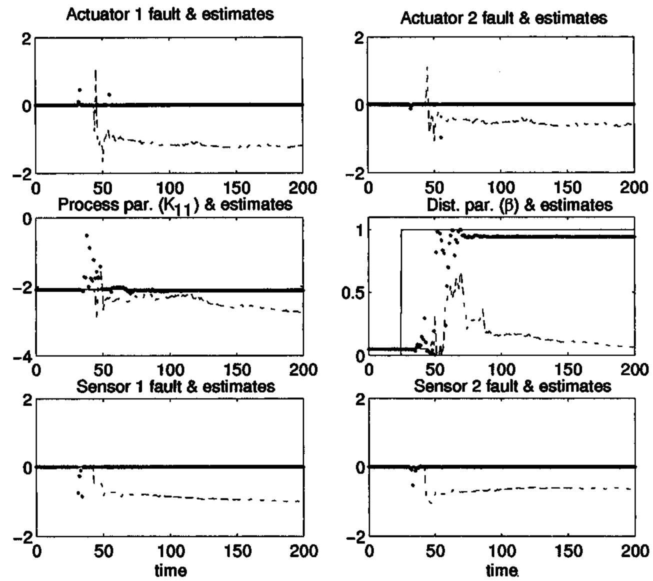

Figure 11. Estimates of faults with multiple filter (dark dotted) and covariance resetting (dashed) algorithms when the disturbance parameter (β) changed. The true fault signal is represented by a solid line.

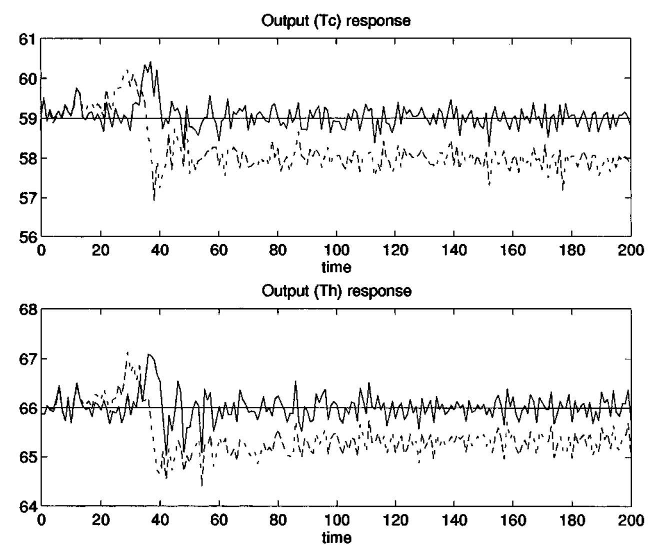

Figure 12. Closed loop output responses when the disturbance parameter changed, using multiple filter (MF) (solid) and covariance resetting (dashed) algorithms for estimation and diagnosis.

passed on to the controller which results in the improved output response after the fault, as shown in Figure 10. The poor output response of the traditional CR algorithm is due to the distribution of the error to different parameters and faults.

Case 2: Disturbance Parameter Change. Figures 11 and 12 show the plots when the disturbance characteristics changed from step to ramp type for $t \geq 15$. Under normal conditions, the disturbance was modeled with an integrated white noise model ($\beta = 0$). Abnormality was detected at $t = 41$, and hence branching was started at $t = 31$. The number of filters that were used varied between 1 and 127, with 127 filters being used only for 15 sample time instances. From $t = 80$ only 1 filter was present and the prediction error decreased below the threshold value. Note that again diagnostic tool No. 2 can be used in altering the prior probabilities of the process/disturbance parameter change, at the time of branching. The estimate of the disturbance parameter β went from 0 to 1, thus adapting the disturbance model to cope with the change in disturbance characteristics of the plant. The CR algorithm again results in the distribution of error. The closed-loop output response in Figure 12 shows the improved performance level that can be obtained with the MF algorithm as compared to that of the CR algorithm.

Case 3: Process Parameter Change. In this case the plant parameter α decreased from 5 to 3 at $t = 15$. The monitoring algorithm detected the abnormality at $t = 22$, and the branching was started at $t = 12$. The number of filters used varied between 1 and 127, with 127 filters being used only for 10 sample time instances. From $t = 60$ only 1 filter was present and the prediction error decreased below the threshold value. The monitoring algorithm was started again to detect further abnormalities. Note that when the number of process parameters are high, diagnostic tool no. 2 can be used to obtain a preliminary idea regarding the presence/absence of process model error. This can in turn be used in altering the prior probabilities at the time of branching. The new parameters (K's) are obtained around the current operating regime using the new input and output values (note that the parameter change causes the model to deviate from the current linear region to another region due to the nonlinearity). The traditional CR algorithm resulted in the distribution of error, affecting fault isolation. Better parameter estimates and output responses were obtained by using the multiple filter approach as in the other cases.

Case 4: Sensor Fault. In this case, sensor bias of magnitude −5 developed in the sensor measuring T_c, at $t = 15$. Monitoring tool no. 1 detected the abnormality at $t = 20$, and the branching was started at $t = 10$. The number of filters used varied between 1 and 49, and from $t = 40$ there was only 1 filter and the prediction error decreased below the threshold value. Even in the multiple filter algorithm, there was a slight change in the process parameter. However, this is due to the nonlinearity of the process and disappeared when a linear model was used for simulating the plant. This could also be tackled by using a nonlinear model for diagnosis and control purposes. The bias information is passed on to the controller to take appropriate action. Better parameter estimates and output responses were obtained by using the multiple filter approach as in the other cases.

Case 5: Consecutive Faults—Sensor Fault Followed by Actuator Fault. Figures 13 and 14 show the diagnostic and output response plots for more than one fault occurring consecutively. Sensor bias of magnitude −5 in the sensor measured T_H at $t = 15$. This was followed by the occurrence of an actuator bias of magnitude −4 that occurred in the actuator 1 at $t = 25$ as in case 1. Abnormality was detected at $t = 22$, and branching was done at $t = 12$. The estimates of the different fault states are shown in Figure 13. The number of filters used varied between 1 and 127, with 127 filters being used only for 15 sample time instants. From $t = 50$ there was only 1 filter and the prediction error decreased. From the figure we see that the suboptimal multiple filter algorithm is able to handle consecutive faults occurring within a short time interval and provide good estimates of the fault. It will be extremely difficult to handle such scenarios with other algorithms such as parallel filter or GLR algorithms. The estimates, when passed on to the controller, result in improved output response as shown in Figure 14. The output response of the CR algorithm is poor.

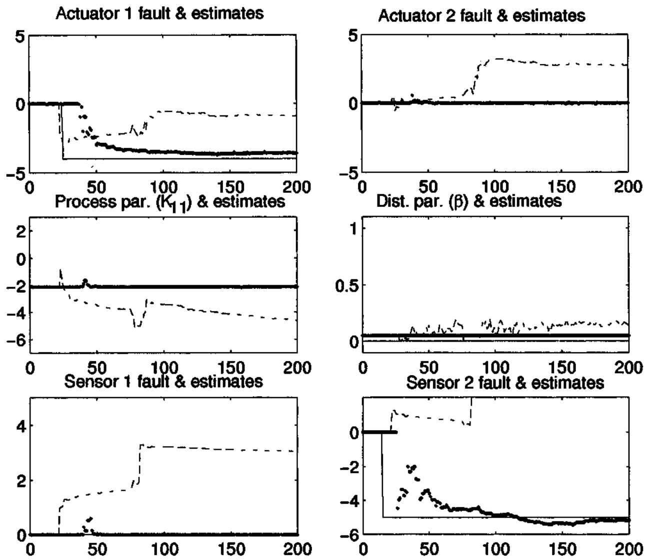

Figure 13. Estimates of faults with multiple filter (dark dotted) and covariance resetting (dashed) algorithms when bias developed in sensor 2 followed by bias in actuator 1. The true fault signal is represented by a solid line.

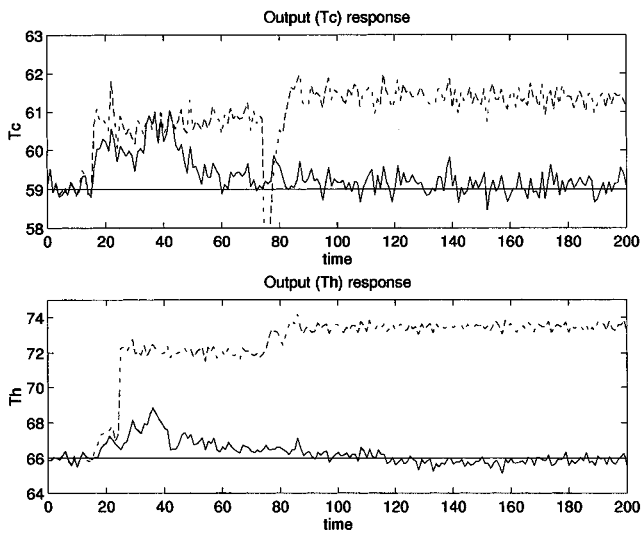

Figure 14. Closed loop output responses when bias developed in sensor 2 followed by bias in actuator 1, using multiple filter (solid) and covariance resetting (dashed) algorithms for estimation and diagnosis.

6. Conclusion

In this paper, we presented several statistical tools for monitoring and diagnosing the performance of multivariable control systems. Various causes such as input saturation, actuator bias, process/disturbance model parameter changes, and sensor bias were considered. Simple diagnostic tools based on prediction error analysis were introduced in order to narrow down the list of potential causes for a closed-loop performance degradation. This can be followed by a detailed fault estimation scheme which estimates the fault parameter vector based on a stochastic fault model. The model requires one to specify a set of potential fault scenarios and their prior probabilities. The results of the diagnosis can be communicated to the operating personnel so that appropriate actions can be taken. Moreover, appropriate modifications to the controller can also be made on an automatic basis, using the analysis. Thus the overall method serves to integrate the statistical process control (SPC) and the automatic process control (APC).

Acknowledgment

The authors gratefully acknowledge the financial support from the NSF NYI Program (CTS #9357827), AspenTech, and Shell Development.

Literature Cited

Ajbar, H.; Kantor J. C. An l_∞ Approach to Robust Control and Fault Detection. AIChE Annual Meeting, November 1993, Paper 146j.

Basseville, M.; Nikiforov, I. V. *Detection of Abrupt Changes*; Prentice Hall: Englewood Cliffs, NJ, 1993.

Box, G. E. P.; Jenkins, G. M. *Time Series Analysis: forecasting and control*; Holden-Day: San Francisco, 1976.

Buxbaum, P. J.; Haddad, R. A. Recursive Optimal Estimation for a Class of Non-Gaussian Process. In *Computer Processing in Communications*; Fox, Jerome, Ed.; Polytechnic Press: New York, 1969.

Desborough, L.; Harris, T. J. Performance Assessment Measures for Univariate Feedback Control. *Can. J. Chem. Eng.* **1992**, *70*, 1186−1197.

Devries, W. R.; Wu, S. M. Evaluation of Process Control Effectiveness and Diagnosis of Variation in Paper Basis Weight via Multivariate Time-Series Analysis. *IEEE Trans. Autom. Control* **1978**, *AC-23*(4), 702−708.

Fathi, Z.; Ramirez, W. F.; Korbicz, J. Analytical and Knowledge Based Redundancy for Fault Diagnosis in Process Plants. *AIChE J.* **1993**, *39*, 42−56.

Harris, T. J. Assessment of Control Loop Performance. *Can. J. Chem. Eng.* **1989**, *67*, 856−861.

Harris, T. J.; Boudreau, F.; MacGregor, J. F. Performance assessment of multivariable Feedback Controllers. AIChE Annual Meeting, Miami Beach, 1995; Paper 182e.

Himmelblau, D. M. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*; Elsevier: Amsterdam, 1978.

Huang, B.; Shah, S. L.; Kwok, E. K. On-Line Control Performance Monitoring of MIMO Processes. *Proceedings of the ACC* **1995**, 1250−1254.

Isermann, R. Process Fault Detection Based on Modeling and Estimation Methods-A Survey. *Automatica* **1984**, *20*, 387−404.

Jacobsen, E. W.; Skogestad, S. Inconsistencies in Dynamic Models for Ill-Conditioned Plants: Applications to Low-Order Models of Distillation Columns. *Ind. Eng. Chem. Res.* **1994**, *33*, 631−640.

Jaffer, A. G.; Gupta, S. C. On Estimation of Discrete Processes Under Multiplicative and Additive Noise Conditions. *Inf. Sci.* **1971**, *3*, 267−276.

Kendra, S. J.; Cinar, A. Frequency domain performance Assessment. AIChE Annual Meeting, Miami Beach, November 1995; Paper 182a.

King, R.; Giles, E. D. Multiple Filter Methods for Detection of Hazardous States in an Industrial Plan. *AIChE J.* **1990**, *36*, 1697−1706.

Kourti, T.; MacGregor, J. F. Multivariate SPC Methods for Monitoring and Diagnosing of Process Performance. *Proceedings of PSE* **1994**, 739−746.

Kozub, D. J. *Monitoring and Diagnosis of Chemical Processes with Automated Process Control*. CPC-V Meeting, Lake Tahoe, 1996.

Ku, W.; Storer, R. H.; Georgakis, C. Use of State Estimation for Statistical Process Control. *Comput. Chem. Eng.* **1994**, 18, S571−S575.

Lee, J. H.; Ricker, N. L. Extended Kalman Filter Based Nonlinear Model Predictive Control. *Ind. Eng. Chem. Res.* **1994**, *33*, 1530−1541.

MacGregor, J. F. On-Line Statistical Process Control. *Chem. Eng. Prog.* **1988**, Oct., 21−31.

Morari, M.; Ricker, N. L. *Model Predictive Control Toolbox* **1994**, The Mathworks Partner Series.

Nett, C. N.; Jacobson, C. A.; Miller, A. T. An integrated approach to Controls and Diagnostics: The 4-Parameter Controller. *Proceedings of the ACC* **1988**, 824−835.

Patton, R.; Frank, P.; Clark, R. *Fault Diagnosis in Dynamic Systems − Theory and Application*; Prentice Hall: Englewood Cliffs, NJ, 1989.

Pryor, C. Auto covariance and Power Spectrum Analysis Derive New information from Process Data. *Control Eng.* **1982**, *Oct.*, 103−107.

Qin, S. J.; Badgwell, T. A. An Overview of Industrial Model Predictive Control Technology. CPC-V Meeting, Lake Tahoe, 1996.

Robertson, D. G.; Lee, J. H.; Rawlings, J. B. A Moving Horizon-Based Approach for Least Squares Estimation. Submitted to *AIChE J.*, **1996**.

Robertson, D. G.; Kesavan, P.; Lee, J. H. Detection and Estimation of Randomly Occurring Deterministic Disturbances. *Proc. ACC* **1995**, 4453−4457.

Stanfelj, N.; Marlin T. E.; MacGregor, J. F. Monitoring and Diagnosis of Process Control Performance: The Single-Loop Case. *Ind. Eng. Chem. Res.* **1993**, *32*, 301−314.

Studebaker, P. Staying on Top of Advanced Controls. *Control Magazine* **1995**, *Dec.*

Tugnait, J. K.; Haddad, A. H. A Detection-Estimation Scheme for State Estimation in Switching Environments. *Automatica* **1979**, *15*, 477−481.

Tyler, M. L.; Morari,M. Performance Monitoring of Control Systems using Likelihood Methods. *Proceedings of the ACC* **1995**, 1245−1249.

Watanabe, K.; Himmelblau, D. M. Application of a Two Level Estimator to a Chemical Reactor. *AIChE J.* **1983**, *29*, 250−260.

Willsky, A. S. A survey of Design Methods for Failure Detection in Dynamical Systems. *Automatica* **1976**, *12*, 601−611.

Willsky, A. S.; Jones, H. L. A Generalized Likelihood Ratio Approach to State Estimation in Linear Systems Subject to Abrupt Changes. *IEEE conf. on Decision and Control* **1974**, 846−853.

Received for review October 17, 1996
Revised manuscript received April 3, 1997
Accepted April 4, 1997[⊗]

IE9606653

[⊗] Abstract published in *Advance ACS Abstracts*, May 15, 1997.