I&EC research
Industrial & Engineering Chemistry Research

Article

Cite This: Ind. Eng. Chem. Res. 2018, 57, 8962−8977

pubs.acs.org/IECR

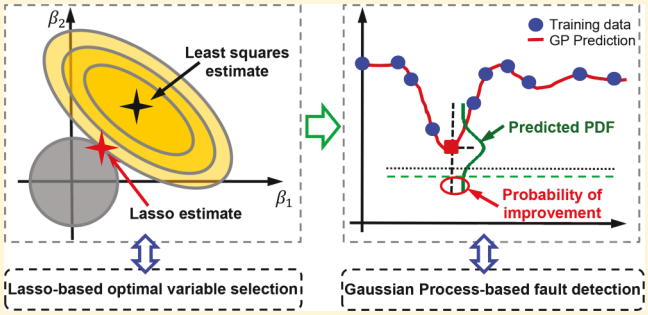
Fault Detection and Classification for Nonlinear Chemical Processes using Lasso and Gaussian Process

Yuncheng Du,*[,†] Hector Budman,[‡] Thomas A. Duever,[‡,§] and Dongping Du[∥]

[†]Department of Chemical and Biomolecular Engineering, Clarkson University, Potsdam, New York 13699, United States
[‡]Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario Canada, N2L 3G1
[§]Department of Chemical Engineering, Ryerson University, Toronto, Ontario Canada, M5B 2K3
[∥]Department of Industrial, Manufacturing, and System Engineering, Texas Tech University, Lubbock, Texas 79409, United States

Supporting Information

ABSTRACT: This paper presents a statistical monitoring methodology to identify and diagnose intermittent stochastic faults occurring in a nonlinear dynamic chemical process. This methodology addresses three important aspects in model-based fault detection and diagnosis (FDD): model simplicity, interpretability, and calibration. The goal is to generate a surrogate model that can be easily interpreted while maintaining model flexibility and efficiency. The key feature is the use of an active set optimization in combination with a Gaussian process (GP) model for fault detection and classification. To optimally select measured variables for inferring faults, an active set optimization with l_1-norm regularization is combined with statistical analysis. This can provide a trade-off between model dimensionality and model prediction error. To ensure sufficient data for the calibration of GP models, an improvement in a probability-based model adjustment algorithm is developed. The performance of the developed FDD scheme is illustrated with two examples: (i) a chemical process consisting of two continuous, stirred tank reactors (CSTRs) and a flash tank separator, and (ii) the Tennessee Eastman benchmark problem. In addition, to deal with multiple-root-cause faults, the GP model based classification was investigated. The summary of the results show that the methodology in this work can cope with both individual and simultaneous occurrences of multiple-root-cause faults in the presence of uncertainty.



1. INTRODUCTION

An important aspect for safe operation and improved product quality of chemical processes is the early detection of abnormal events and malfunctions that are defined as faults.[1] For detectable faults, the fault detection and diagnosis (FDD) algorithm can provide symptomatic fault features, which will be further used by an FDD scheme to identify the root cause of any abnormal behavior. Different methods have been developed in the literature for FDD. These methods can be broadly categorized into three groups:[2,3] (i) analytical methods that are solely based on a first-principles model of process;[4−7] (ii) surrogate (empirical) modeling methods such as multivariate statistical analysis that use the historical data collected from the processes;[8,9] (iii) semiempirical techniques that integrate first-principles models with surrogate (empirical) models.[10−12]

Each of these aforementioned modeling techniques for FDD has its own advantages and disadvantages depending on the process of interest. It is well-recognized that surrogate models are easier to develop, while first-principles models have superior extrapolation ability.[13−15] This work will focus on the development of a surrogate model for fault detection and classification using historical data. Since data in chemical processes generally exhibit high correlation over time and have cross-correlation among variables, multivariate statistical analysis (MVSA) techniques such as partial least squares (PLS) have been used to reduce model complexity, thus leading to improved accuracy.[9] Lower dimensional representations formed with MVSA can often better generalize to new process data, as compared to the representations using the full dimensionality. The main drawback is that models built with MVSA techniques are not interpretable, since they generally rely on subspaces that involve linear combinations of the original physical states.

To build a model in the original physical states, the least absolute selection and shrinkage operator (Lasso) was proposed to select significant variables based on the polyhedral structure of the l_1-norm regularization.[16] However, the use of Lasso may leave a larger number of variables, as compared to

Received: March 13, 2018
Revised: June 14, 2018
Accepted: June 16, 2018
Published: June 17, 2018

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

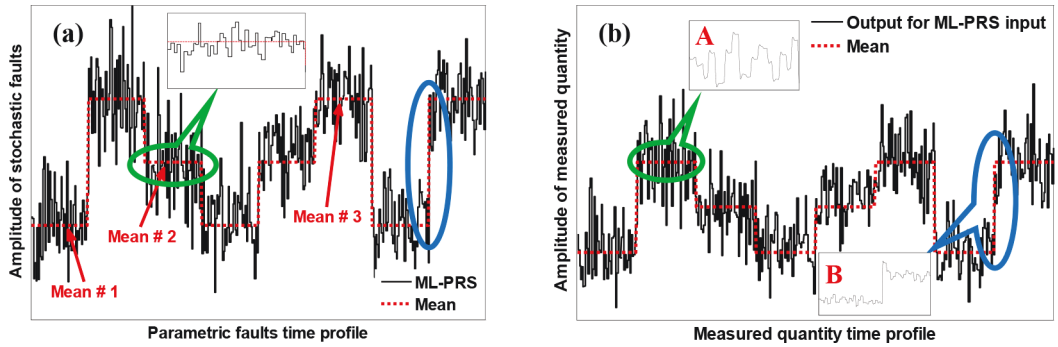


Figure 1. Fault profile denoting an intermittent stochastic input fault and resulting measured variable.

the conventional MVSA methods. To address this, the active set method[17,18] is used to keep the computational cost relatively low and to improve model convergence. Models generated with active set involve a set of physical variables that have significant effects on a chemical process. However, the resulting models are linear and may be less efficient to monitor nonlinear chemical processes.[6]

Uncertainty is one of the major challenges for accurate fault diagnosis and classification, since most of the FDD tools generally rely on models that are not perfect.[3] Such model uncertainty may either result from intrinsic time-varying phenomena that are not considered in the models or originate from inaccurate calibration because of noise in the data used for model calibration.[12] Generally, the effect of uncertainty on model predictions is typically ignored in the reported FDD techniques, leading to a loss of accuracy.[19] To account for the effect of uncertainty on FDD, a Gaussian process (GP) model can be used.[20] GP models present a new, emerging, and complementary approach for system identification[21,22] and design of robust controller.[23,24] A significant feature of the GP models is that they only involve a few tuning parameters, as compared to other surrogate model based methods such as neural networks.[25] In addition, GP models can provide a probabilistic description of uncertainty for hypothesis testing.[26] It is important to note that FDD with GP models comprises regression and classification components.[25] The main difference between the regression and classification components is how the measured quantities are linked to the faults, i.e., continuously or discretely. The regression is concerned with the accurate prediction of the continuous quantities of faults. In contrast, the results of classification are discrete class labels, for which the prediction of possible faults is assigned into one of the predefined classes. For brevity, the classification problem with GP classification models is discussed in the Supporting Information, and we mainly focus on the use of the GP regression model for fault classification here.

Surrogate models to be used for FDD should describe the relationship between faults and measured variables. These models must be calibrated with data, and this calibration step is sensitive to the amount and density of data used for model training. Surrogate models can be inaccurate when dealing with observations that were not used for model calibration.[13] To improve FDD, it is imperative to ensure that sufficient data are available for model training in order to develop a robust model. One possibility is to calibrate models with a large amount of measurements. However, this may require performing many physical experiments, which would be impractical and expensive. Additionally, some measurements used for model calibration may have little effect on improving model accuracy. In this work, we develop a methodology that uses a combination of actual measurements and synthetic data obtained from simulations with the first-principles models for improved model calibration. It should be noted that, if actual data are not available, only simulated data that are selected based on a cumulative probability criterion can be used. Actual data can be used for other than simulations when there are sufficient training data for model calibration.

In summary, a surrogate model is developed using physical process variables while taking into account uncertainty. The proposed method involves three consecutive steps:

(i) Data dimensionality reduction: An active set optimization is combined with statistical analysis to find measured variables that are sensitive to stochastic faults.

(ii) Adaptive GP model calibration: A surrogate model is developed using a GP-based supervised learning method, which is calibrated with synthetic data based on a minimal model adjustment algorithm.

(iii) Stochastic fault detection and diagnosis: The GP model is used to infer intermittent faults consisting of stochastic perturbations superimposed on intermittently changing mean values of a particular input. In addition, a GP classification model is developed in order to deal with multiple-root-cause faults (see the Supporting Information).

The paper is organized as follows. In section 2, the formulation of a fault detection and classification problem is presented, which is followed by the theoretical background of the active set optimization and the GP theory. The fault detection and diagnosis (FDD) algorithm developed in this work is explained in section 3. A nonlinear chemical process, consisting of two continuously stirred tank reactors (CSTRs) and a flash tank separator, and the Tennessee Eastman benchmark process are used as case studies in section 4. Simulation results and brief discussion of the results are given in section 5 followed by the conclusion in section 6. For multiple-root-cause faults, the Lasso and GP model based stochastic fault classification problem is discussed in the Supporting Information for brevity.

2. PROBLEM FORMULATION AND MATHEMATICAL BACKGROUND

2.1. Formulation of Intermittent Stochastic Faults. A nonlinear chemical plant, subjected to uncertain parametric input faults, is described by a dynamic model as

$$\dot{\boldsymbol{x}} = \mathcal{B}(t, \boldsymbol{x}, \boldsymbol{u}; \mathbf{G}) \qquad 0 \le t \le t_f, \boldsymbol{x}(0) = \boldsymbol{x}_0 \tag{1}$$

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

where the vector $\boldsymbol{x} \in \boldsymbol{R}^n$ represents the states of the system including measured variables with initial conditions $\boldsymbol{x}_0 \in \boldsymbol{R}^n$ over time domain $[0, t_f]$; $\boldsymbol{u}$ denotes the known (measurable) inputs of the system. The matrix $\mathbf{G} \in \boldsymbol{R}^{p \times n_g}$ represents the unknown (unmeasured) stochastic time-varying input faults, which will be inferred with FDD algorithms. The function $\mathcal{B}$ builds the relationship between inputs ($\boldsymbol{u}$ and $\mathbf{G}$) and system states ($\boldsymbol{x}$). It should be noted that $\mathbf{G}$ in this work denotes the multiple-root-cause stochastic faults, i.e., $\mathbf{G} = [\mathbf{g}_1, ..., \mathbf{g}_j, ..., \mathbf{g}_p]^T$, where $\mathbf{g}_j$ ($j = 1, ..., p$) is a row vector involving a finite set of mean values, i.e., n_g. The definition of $\mathbf{g}_j$ and the formulation of faults will be further explained below using a single-root-cause fault example, which can be subsequently further extended to the multiple-root-cause fault problem in the Supporting Information. Also, it is assumed that measured quantities used for FDD in this work are corrupted by additive measurement noise.

For brevity, only the formulation of fault detection algorithms for a single-root-cause fault is discussed here to show the definition of stochastic faults and to demonstrate how the fault detection algorithm operates in this current work. For clarify, let assume $p = 1$; thus $\mathbf{G}$ can be simplified as a row vector, i.e., $\mathbf{G} = \mathbf{g} = \{g_i\}$, where $i = 1, ..., n_g$. The input fault of $\mathbf{g}$, consisting of stochastic perturbations superimposed on n_g sets of mean values, can be described in Figure 1a, which can be further mathematically defined as

$$g_i = \bar{g}_i + \Delta g_i \qquad (i = 1, ..., n_g) \tag{2}$$

where $\{\bar{g}_i\}$ are a set of constant mean values (operating modes); $\{\Delta g_i\}$ are stochastic variations around each mean value. The statistical distributions of the changes Δg_i are assumed to be time invariant and estimated from a model calibration algorithm. The occurrence of a new steady state, i.e., the constancy of the mean values of $\{\bar{g}_i\}$, can be experimentally inferred from the constancy of measured variables in Figure, 1b, such as the manipulated variables.

As shown in Figure 1, the changes in the mean values of $\{\bar{g}_i\}$ (faults) follow a multilevel pseudorandom signal (ML-PRS).[27] The faults defined in eq 2 are typical in chemical plants that can experience changes in means of operating variables and additional continuous random perturbations superimposed on each of the means of faults.[12] To illustrate the how the fault classification algorithm operates, it is assumed that the system in this work has been operated for long periods around a specific mean value and the objective is to identify and diagnose the mean value of the fault, i.e., $\{\bar{g}_i\}$, in the presence of perturbations $\{\Delta g_i\}$. In summary, the FDD algorithm in this work has two following objectives.

Objective 1 (fault detection): The first objective is to identify any possible stepwise changes between mean values of $\{\bar{g}_i\}$, where each $\bar{g}_i$ will be alternatively referred hereinafter as to an operating mode in this current work.

Objective 2 (fault classification): The second objective is to diagnose (or classify) a specific operating mode $\bar{g}_i$ at a given time instant t while taking into account the perturbation $\{\Delta g_i\}$.

The FDD problem can be further extended to multiple-root-cause faults. In this case, the objective of the fault detection is to identify the possible step changes between mean values in any row of $\mathbf{G}$, whereas the objective of the fault classification is to find a particular entry in $\mathbf{G}$, which represents the operating mode corresponding to a pair of specific mean values in $\mathbf{G}$. This will yield a classification problem that is discussed in the Supporting Information. For brevity, only the FDD of a single-root-cause fault is discussed here.

2.2. Active set Optimization with Regularization. To select the measured variables that are most sensitive to faults, let assume that, for the purpose of model training, measurements of a single-root-cause fault $\mathbf{G} = \mathbf{g} = \{g_i\}$ and measured variables $\boldsymbol{x}$ are known and can be rewritten as

$$\mathbf{g} = [g_1, ..., g_i, ..., g_m]^T \tag{3}$$

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & & \vdots \\ x_{i,1} & \ddots & x_{i,n} \\ \vdots & & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{pmatrix} \tag{4}$$

where m denotes the total number of measurements, and n is the number of measured variables in eq 1. Each value in eq 3 can be compared to the set of mean values $\{\bar{g}_i\}$ in eq 2 based on a minimum distance criterion to identify the operating mode, i.e., normal vs faulty.

Linear squares regression can find a linear combination of $\{\boldsymbol{x}_j\}$ ($j = 1, ..., n$) to approximate the relationship between $\mathbf{g}$ and $\mathbf{X}$. However, the variances of the regression coefficients can be unacceptably high when n is large or the measured variables are highly correlated as previously reported.[16] Lasso can be used to minimize the residual of regression with constraints on the l_1 regularization of regression coefficients,[16,17] according to the following optimization problem:

$$\min_{\beta_1,...,\beta_n} J = \frac{1}{2} \sum_{j=1}^{m} \left(g_j - \sum_{i=1}^{n} x_{i,j} \beta_i \right)^2 \tag{5a}$$

subject to

$$\sum_{i=1}^{n} |\beta_i| \le \alpha \tag{5b}$$

where $\alpha > 0$. For smaller values of α, Lasso will ultimately drive the regression coefficients toward zero, thus making the algorithm useful for selecting variables while ruling out others. However, it is not trivial to solve eq 5a, when n is large and the number of measurements is limited.[17] The active set method is one of the most influential work since the original Lasso for solving the optimization efficiently.

The optimization problem in eq 5a can be reformulated as follows:

$$\min_{\beta=\{\beta_1,...,\beta_n\}} J = \frac{1}{2} f(\boldsymbol{\beta}) = \frac{1}{2} (\mathbf{g} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{g} - \mathbf{X}\boldsymbol{\beta}) = \frac{1}{2} \boldsymbol{r}^T \boldsymbol{r} \tag{6a}$$

subject to

$$\psi(\boldsymbol{\beta}) \ge 0 \tag{6b}$$

where f is a continuous function to represent the cost function of the regression problem, $\psi(\boldsymbol{\beta}) = \alpha - \sum_{i=1}^{n} |\beta_i|$, $\boldsymbol{r} = \boldsymbol{r}(\boldsymbol{\beta})$ is a vector of residuals related to regression coefficients $\boldsymbol{\beta}$, and $\psi(\boldsymbol{\beta})$ is implicitly a function of α given by eq 5b. The optimization in eq 6a can be solved with an iterative algorithm.[17] A key feature of the active set method is the use of a local linearization about the current value of $\boldsymbol{\beta}$ whereby the basic procedure involves the

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

computation of a correction $\boldsymbol{h}$ with respect to the local linearization leading to the following optimization:

$$\min_{\boldsymbol{h}} J = f(\boldsymbol{\beta} + \boldsymbol{h}) \tag{7a}$$

subject to

$$\boldsymbol{\theta}_\sigma^{\mathrm{T}}(\boldsymbol{\beta}_\sigma + \boldsymbol{h}_\sigma) \le \alpha \tag{7b}$$

$$\boldsymbol{h} = P^{\mathrm{T}}\begin{pmatrix}\boldsymbol{h}_\sigma \\ 0\end{pmatrix} \tag{7c}$$

where P is a permutation matrix that collects the nonzero components of $\boldsymbol{\beta}$ associated with the first $|\sigma|$ components. At each iterative step, the ith component of $\boldsymbol{\beta}$ is nonzero only if $i \in \sigma$, where σ is referred to as the index set (members of the active set) and is updated at each step of the optimization. In addition, $\boldsymbol{\theta}_\sigma = \mathrm{sign}(\boldsymbol{\beta}_\sigma)$ has entry 1 if the corresponding entry in $\boldsymbol{\beta}_\sigma$ is positive and -1 otherwise. For each step, $\boldsymbol{\beta}$ has to be feasible with respect to eq 5b, i.e., $\boldsymbol{\theta}_\sigma^{\mathrm{T}}\boldsymbol{\beta}_\sigma \le \alpha$. If the constraint is active, the optimization results, satisfying the KKT (Karush−Kuhn−Tucker) conditions of the optimization given by eq 7a, are described as follows:

$$\mu = \max\left(0, \frac{\boldsymbol{\theta}_\sigma^{\mathrm{T}}(\boldsymbol{X}_\sigma^{\mathrm{T}}\boldsymbol{X}_\sigma)^{-1}\boldsymbol{X}_\sigma^{\mathrm{T}}\mathbf{g} - \alpha}{\boldsymbol{\theta}_\sigma^{\mathrm{T}}(\boldsymbol{X}_\sigma^{\mathrm{T}}\boldsymbol{X}_\sigma)^{-1}\theta_\sigma}\right) \tag{8a}$$

$$\boldsymbol{h}_\sigma = (\boldsymbol{X}_\sigma^{\mathrm{T}}\boldsymbol{X}_\sigma)^{-1}(\boldsymbol{X}_\sigma^{\mathrm{T}}(\mathbf{g} - \boldsymbol{X}_\sigma\boldsymbol{\beta}_\sigma) - \mu\boldsymbol{\theta}_\sigma) \tag{8b}$$

where μ is a positive scalar, $\boldsymbol{X}_\sigma$ is a finite subset of measurements defined with respect to the active set σ, which is initially empty, and a zero-value element will be added to this set at the end of each iteration. Specifically, elements that are not included in σ and exhibit the largest violation will be added to the active set. Let suppose $\boldsymbol{\beta}^+ = \boldsymbol{\beta} + \boldsymbol{h}$ and define a violation as follows:

$$\boldsymbol{v}^+ = \frac{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{r}^+}{\|\boldsymbol{X}_\sigma^{\mathrm{T}}\boldsymbol{r}^+\|_\infty} \tag{9}$$

where $\boldsymbol{r}^+ = \mathbf{g} - \boldsymbol{X}\boldsymbol{\beta}^+$. Note that variables outside the active set will be 0 as will corresponding variables of $\boldsymbol{h}$. Since sign(0) is not well-defined, the θ_i $(1 \le i \le \sigma)$ value for the variable to be introduced into the active set is set to the sign of the corresponding violation. If the magnitude of the violation for all variables outside the active set is less than 1, then optimality is considered to be achieved.[17] Based on this, the variable that results in the largest violation will be added to σ, and then solve for μ and subsequently $\boldsymbol{h}_\sigma$. The optimization becomes more complicated when a variable in the active set may change sign during an iteration, which requires additional treatments,[17] but it is not discussed for brevity.

The active set method can handle data in which measured variables are highly correlated. It is particularly useful when the number of available measurements used for model training is limited. Measured variables that have equal impact on the faults will be given equal regression coefficients, and zero regression weights will be given to variables that have negligible effect on the faults. These properties of Lasso make it useful for selecting a reduced set of measured variables to be used for FDD. Further details on the use of the active set method for measurement selection are given in section 3.

2.3. Gaussian Process Model. The modeling with Gaussian process (GP) involves multivariate Gaussian distributions of infinite dimensionality.[25] For algorithm clarification, we present the formulation of a GP model to estimate values of $\mathbf{g}$ with measurements $\boldsymbol{x}$. Based on a training set $\mathcal{D} = \{(x_i, \mathrm{g}_i)\}$ $(i = 1, ..., N)$ with N pairs of measurements, a GP regression model is defined with respect to *z-scored* measured variables, i.e., mean-centered around zero and scaled by their standard deviations:

$$\mathbf{g}_i = \mathcal{G}(\boldsymbol{x}_i) + \varepsilon_i \tag{10}$$

$$\varepsilon_i \sim \mathcal{N}(0, {\sigma_g}^2) \tag{11}$$

where $\mathcal{G}$ is the GP surrogate model and ε_i is the residual, which can be approximated with a Gaussian noise model $\mathcal{N}$ with a mean of zero and a standard deviation of σ_g. Accordingly, $\mathbf{g}_i$ is nonlinearly related to $\boldsymbol{x}_i$ via an unknown function $\mathcal{G}$ that is approximated with a GP model. Furthermore, each measurement within the training set $\boldsymbol{X} = \{x_i\}$ is related to another measurement through a covariance function $K = \{k_{ij}\} = \{k(x_i, x_j)\}$. The covariance function K is a squared exponential kernel function in this work,[28] which is defined as

$$k_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j) = {\sigma_G}^2 \exp\left(-\frac{1}{2l^2}(\boldsymbol{x}_i - \boldsymbol{x}_j)^2\right) + {\sigma_g}^2\delta(\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{12}$$

where δ_{ij} denotes the Kronecker delta function. Unknown parameters $\boldsymbol{\theta} = (\sigma_G, l, \sigma_g)$ are referred to as *hyperparameters*. σ_G is the maximum allowable covariance. For example, $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ will approach the maximum when $\boldsymbol{x}_i \approx \boldsymbol{x}_j$, meaning that $\mathcal{G}(\boldsymbol{x}_i)$ is nearly perfectly correlated with $\mathcal{G}(\boldsymbol{x}_j)$. When $\boldsymbol{x}_i$ is very distant from $\boldsymbol{x}_j$, $k(\boldsymbol{x}_i, \boldsymbol{x}_j) \approx 0$, implying that distant measurements may have negligible effects to interpolate new measurements.

For measurements $\boldsymbol{x}$ in the training set and *hyperparameters*, the covariance for all possible combinations of this N set of data points can be calculated with eq 12. Suppose that K is the covariance matrix of N training measurements; i.e., $K = \{k_{ij}\}$ and $1 \le i, j \le N$. The covariance matrix K can be written as follows:

$$K = \begin{bmatrix} k(\boldsymbol{x}_1, \boldsymbol{x}_1) & k(\boldsymbol{x}_1, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_1, \boldsymbol{x}_N) \\ k(\boldsymbol{x}_2, \boldsymbol{x}_1) & k(\boldsymbol{x}_2, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_2, \boldsymbol{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_N, \boldsymbol{x}_1) & k(\boldsymbol{x}_N, \boldsymbol{x}_2) & \cdots & k(\boldsymbol{x}_N, \boldsymbol{x}_N) \end{bmatrix} \tag{13}$$

where the diagonal element of K is ${\sigma_G}^2 + {\sigma_g}^2$. Using the available measurements of a training set $\mathcal{D} = \{(x_i, \mathrm{g}_i)\}$, the objective is to predict g* for a set of new measurements of x*. A key assumption in GP modeling is that measurements $\boldsymbol{x}$ can be represented as samples from a multivariate Gaussian distribution as

$$\begin{bmatrix}\mathbf{g} \\ \mathrm{g}*\end{bmatrix} = \mathcal{N}\left(0, \begin{bmatrix} K & K_*^{\mathrm{T}} \\ K_* & K_{**}\end{bmatrix}\right) \tag{14}$$

where

$$K_* = [k(\boldsymbol{x}_*, \boldsymbol{x}_1) \quad k(\boldsymbol{x}_*, \boldsymbol{x}_2) \quad \cdots \quad k(\boldsymbol{x}_*, \boldsymbol{x}_n)]$$

$K_{**} = k(\boldsymbol{x}_*, \boldsymbol{x}_*)$, and "T" represents matrix transpose.[25,28] Given the measurements of $\mathbf{g}$, the conditional probability $p(\mathrm{g}_*|\mathbf{g})$ follows a Gaussian distribution as

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

$$p(g_*|\mathbf{g}) = \mathcal{N}(K_*K^{-1}\mathbf{g},\ K_{**} - K_*K^{-1}K_*^{T}) \quad (15)$$

The best estimate of g_* is the mean of the distribution in eq 15, which is defined as

$$\bar{g}_* = K_*K^{-1}\mathbf{g} \quad (16)$$

The uncertainty in the estimate can be calculated with its variance as

$$\mathrm{var}(g_*) = K_{**} - K_*K^{-1}K_*^{T} \quad (17)$$

These estimates are the key since the estimated mean is used for fault classification and the variance or uncertainty of the estimated mean is used for a model adjustment operation explained in section 3.

As seen in eqs 16 and eq 17, the reliability of the prediction of mean and variance is dependent on the covariance function K, which is related to the hyperparameters $\boldsymbol{\theta}$. The calibration of the GP requires determining the unknown *hyperparameters* in eq 12, i.e., $\boldsymbol{\theta} = \{\sigma_G, l, \sigma_g\}$, based on a given training set $\mathcal{D}$. The parameters $\boldsymbol{\theta}$ can be obtained with an empirical Bayes estimation technique by maximizing a likelihood function as[28]

$$\arg\max \log p(\mathbf{g}|\boldsymbol{x}, \boldsymbol{\theta}) = -\frac{1}{2}N\log(2\pi) - \frac{1}{2}\log|K| - \frac{1}{2}\mathbf{g}^{T}(K)^{-1}\mathbf{g} \quad (18)$$

This optimization can be simply solved with multivariate optimization algorithms such as conjugate gradients,[25,28] which are used in this work. Figure 2 schematically shows the GP

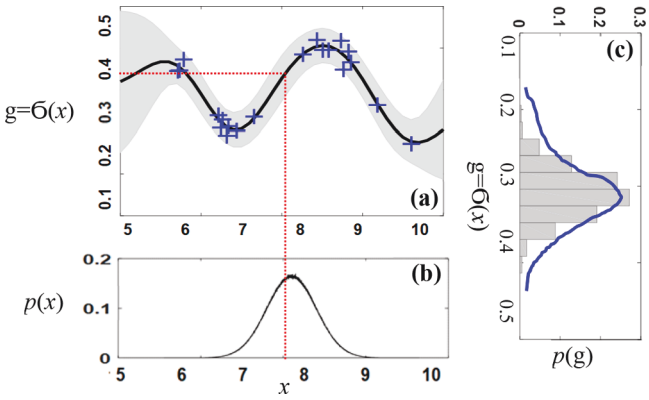


Figure 2. Illustration of GP model and its prediction.

training and prediction results, where Figure 2a shows a GP posterior distribution, Figure 2b shows the distribution of the measured variable around a mean, and Figure 2c shows the corresponding model prediction with a GP model. The bins in Figure 2c representing the distribution of the model prediction such as faults are obtained from Monte Carlo sampling of the resulting posterior distribution.

It is important to note that one major challenge using the GP surrogate model is the computational burden for a high-dimensional parameter space. For example, for a training set with n-measured variables and m measurements of each variable, the predictions require $O(n^2m)$ operations in addition to the $O(n^3)$ operations involved in inverting the covariance matrix,[29] thus making GP computationally intensive for high-dimensional application. To overcome this computational challenge, the active set optimization based variables selection algorithm explained in section 2.2 is first applied to identify variables that are sensitive to faults. A GP model is then generated with measurements of these sensitive variables for fault classification.

3. FAULT DETECTION AND DIAGNOSIS ALGORITHMS BASED ON GP MODELS

3.1. Selection of Measured Variables for FDD with Active Set Optimization. Since some of the available data may provide limited information on faults thus increasing computational burden with no additional gain, the appropriate selection of measured quantities that can be used for FDD is useful. A multistep algorithm is developed for this purpose to identify variables that are sensitive to faults. Specifically, the active set optimization is combined with a Latin hypercube sampling technique to reduce model dimensionality.

A linear regression model, describing the relationship between faults and measured variables, can be formed to find an optimal set of measured variables and to minimize the residual sum of squared errors as

$$g_i = \beta_0 + \sum_{k=1}^{p} x_{i,k}\beta_k + \epsilon_i \quad (19)$$

where ϵ_i is the error between the ith fault value and the model prediction, $i = 1, ..., m$. To simplify the implementation, data of faults $\{g_i\}$ is normalized by setting β_0 to the mean value as $\bar{g} = \sum_{i=1}^{m} g_i/m$. Since each measured variable may have different units and orders of magnitude, each of them is normalized with respect to its mean value as done for the faults, thus resulting in the following problem:

$$g_i' = \sum_{k=1}^{p} x_{i,k}'\beta_k + \epsilon_i \quad (20)$$

where $x_{i,k}'$ and g_i' are the normalized measurements (or synthetic data) obtained from simulations with a first-principles model, $x' \in R^p$, and $\boldsymbol{\beta} \in R^p$. The value of p that represents the number of measured variables that are sensitive to the faults while minimizing the error ϵ_i is determined with the active set optimization method. It should be noted that the linear regression model is used to determine measured variables that are sensitive to the faults. If there is sufficient evidence to support a more complicated model such as nonlinear models, then nonlinear algebraic transformation can be applied to the measurements of $\boldsymbol{X}$ and $\mathbf{g}$ to obtain better models.

The number of available measurements required for training according to eq 20 may affect the accuracy of regression coefficients. To ensure sufficient data, a large amount of measurements may be required, since some measurements may contain little useful information. To overcome this challenge while accounting for different faults, we propose in this work the use of synthetic data generated by first-principles models, instead of using actual measurements. A Latin hypercube sampling (LHS) approach is used to generate synthetic data to improve computational efficiency. The active set optimization is then applied to identify measured variables that are sensitive to faults. Once again, actual data can be used in lieu of synthetic data, when there are sufficient data and/or first-principles models are not available. The optimal selection of

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

sensitive measured variables used for generating an FDD model (GP model) is summarized as per the following steps.

Step 1. To effectively identify sensitive measured variables, eq 20 is reorganized as

$$\zeta = \sum_{i=1}^{m} \epsilon_i^{\,2} = \sum_{i=1}^{m} \left(g_i' - \sum_{k=1}^{p} x_{i,k}' \beta_k \right)^2 \tag{21}$$

where m is the total number of data used for sensitive variables selection. The goal is to find a vector of significant measured variables p and their corresponding regression parameters β_k that can minimize the sum of squared errors ζ between the model predictions and the training data.

Step 2. To generate training data, a set of constant mean values and stochastic variations around each mean of faults must be first approximated from offline model calibration algorithms.[4] A Latin hypercube sampling (LHS) is used to generate a training set with the estimated mean and variance. For each mean value of the faults, n_s samples are simulated which results in M samples in total, i.e., $M = n_g \times n_s$. Each sample is then used to simulate measurements of $\boldsymbol{x}$. This will result in M measurements for each measured quantity. Gaussian noise is added to each measurement. This step generates a training set $\mathcal{D} = \{(\boldsymbol{x}_i', g_i')\}$ involving M pairs of measurements $(i = 1, ..., M)$.

Step 3. An optimal combination of the measured quantities p in eq 21 will be identified with the training set $\mathcal{D}$ and a *cross-validation procedure*. For a specified value of α in eq 7b, a $(1/b)$ portion of the training set $\mathcal{D}$ is randomly selected and used to find the optimal solution of $\{\beta_k\}$ in eq 21, using the active-set optimization with regularization. An initial subset of measured variables, i.e., p', is first considered to be the best selection. Based on the optimization results of $\{\beta_k\}$ with the initial subset, measured quantities will be removed from the initial subset p', when a regression coefficient β_k is found to be zero or smaller than a given threshold ζ. The total number of the removed quantities is defined as r. For an updated subset with $(p' - r)$ measured quantities, the optimization with eq 7a will be repeated until the identified measured quantities p converge to a constant value, and the corresponding optimization results $\{\beta_k\}$ for the identified measured variables (p) are stored.

Step 4. Using the results obtained in step 3, a regression model as eq 20 can be generated. The remaining portion of the training set $\mathcal{D}$ that was not used in step 3 is used for model validation. The residual sum of squared errors $\{\epsilon_i\}$ between the model predictions and the measurements of faults are computed and stored.

Step 5. To avoid model overfitting, steps 3 and 5 can be repeated several times. Each time, a $(1/b)$ portion of the training set $\mathcal{D}$ is randomly selected. The frequency of the measured quantities and the residual sum of squared errors are recorded. The model is then generated with the measured quantities that have the highest frequency and the lowest residual sum of squared errors, thus leading to the successful identification of variables that are most sensitive to any possible faults.

It should be noted that the generation of a training set in step 3 can also be repeated several times when the variance of the faults is found to be relatively large. For different training sets, steps 3 and 4 can be applied by following the same procedure explained above. In addition, the use of the Latin hypercube sampling (LHS) eliminates the possibility that sampling points will come from the same local domain as compared to the Monte Carlo simulations. This will be further discussed in section 5 with an example.

3.2. GP Model Calibration. The active set optimization based variables selection in a previous section finds a linear relationship between faults and measured variables. However, the FDD performance with the resulting linear model may be low as found in the case study presented later. In contrast, the GP model is a nonlinear model that can provide improved fault classification.

For each set of available measurements $\boldsymbol{x}^*$, the GP model can predict the dynamic value of g* by calculating its mean value and its variance with eqs 16 and 17. However, the prediction accuracy of the GP models is sensitive to the density and amount of available data points that can be used for the model calibration. To ensure that sufficient data is available for the model calibration, a model adjustment algorithm is developed, which will add new data of sensitive variables into a training set for the improved model calibration. This method will combine a cumulative distribution function (CDF) using probability improvement with an adaptive selection criterion of new training data as explained later.[30]

To quantify the amount of additional data required for GP model calibration, a measure of the model discrepancy between the GP model prediction and the actual value of $\{\mathbf{g}_i\}$ used for model calibration can be defined as

$$\varepsilon = \mathbf{g}_i - \partial\mathcal{G}(\bar{\mathbf{g}}_i|\boldsymbol{x}_i, \boldsymbol{\theta}) \tag{22}$$

where $\boldsymbol{x}_i$ and $\mathbf{g}_i$ are the ith set of available data points in a given initial training set, i.e., $\mathcal{D}_0 = \{(\boldsymbol{x}_i, \mathbf{g}_i)\}$ and $i = 1, ..., N$, and N is the total number of measurement sets. In eq 22, $\mathcal{G}(\cdot)$ represents the GP model that can predict the mean value of faults $\bar{\mathbf{g}}_i$ based on $\{\boldsymbol{x}_i\}$ and hyperparameters $\boldsymbol{\theta}$ that is obtained with the initial training set $\mathcal{D}_0$. It is worth mentioning that $\{\boldsymbol{x}_i\}$ represents measurements of variables that are sensitive to faults, which are identified with the active set optimization explained in section 2. Figure 3 shows a schematic of the model adjustment for clarification.

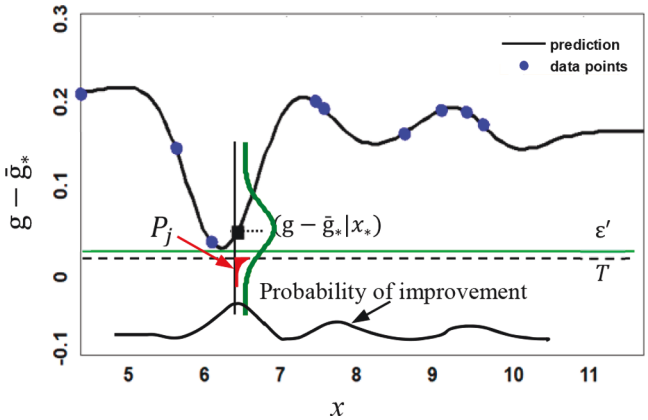


Figure 3. Schematic of improvement in the probability for a target value T.

Based on the information on model discrepancy, a cumulative distribution function (CDF) is used in this work to identify an optimal amount of data that are necessary for maximizing the probability improvement while minimizing the model discrepancy beyond a predefined target T.[30,31] The

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

improvement in probability P_j used to search additional data points for the GP model calibration can be defined as follows:

$$P_j = \psi[(T - E(\mathbf{g}_j'|\boldsymbol{x}_j))/s(\mathbf{g}_j'|\boldsymbol{x}_j)] \quad (23)$$

where ψ is a normal cumulative distribution function, $E(\cdot)$ is the mean value, and $s(\cdot)$ is the standard deviation of model discrepancy, respectively. For any given measurements $\boldsymbol{x}_j$, the model prediction $\overline{\mathbf{g}}_{j*}$ can be defined with a probability density function (PDF) since the GP model is used. The PDF is compared with fault value $\mathbf{g}_j$, thus producing a PDF of the model discrepancy, i.e., $\mathbf{g}_j' = \mathbf{g}_j - \overline{\mathbf{g}}_*$, which can be used to calculate the mean and standard deviation in eq 23. In addition, T is a predefined target value used to tune the model and to evaluate the improvement in probability. For instance, T can be defined as $T = \varepsilon' - 0.2|\varepsilon'|$ to produce at least a 20% improvement (see Figure 3), where the red area, i.e., P_j, is the maximum improvement of probability for a set of given measurements x_*.

The optimal selection of synthetic data used for model adjustment can be summarized as follows.

(i) Build a GP model to obtain the initial *hyperparameters* $\boldsymbol{\theta}_0$ with an initial training set $\mathcal{D}_0$.

(ii) Specify the model discrepancy criterion ε' and calculate the target value T for a predefined improvement in probability.

(iii) Generate new synthetic data $\mathcal{D}_1$ with simulations, which consist of N_1 values for n variables.

(iv) For each set of measurements $\boldsymbol{x}_j$ in $\mathcal{D}_1$ (j = 1, ..., N_1), estimate the predicted mean value and the predicted standard deviation of g_j using the GP model with initial hyperparameters $\boldsymbol{\theta}_0$, and then calculate the difference with respect to training data $\mathbf{g}_j$ used to generate $\boldsymbol{x}_j$.

(v) Compute the probability improvement P_j with eq 23 with the estimated mean $E(\mathbf{g}_j'|\boldsymbol{x}_j)$ and the standard deviation $s(\mathbf{g}_j'|\boldsymbol{x}_j)$.

(vi) Synthetic data in $\mathcal{D}_1$ with the maximum probability improvement P_j will be used and added to the initial training set $\mathcal{D}_0$. This will yield a new training set, which now consists of $(N + 1)$ set of data points for n variables.

(vii) Calibrate the GP model using the new training set obtained in step vi, which involves $(N + 1)$ set of samples. This will provide a new set of *hyperparameters* $\boldsymbol{\theta}_0'$.

(viii) Calculate the model discrepancy ε and $\boldsymbol{\theta}_0'$.

(ix) Replace the initial hyperparameter $\boldsymbol{\theta}_0$ with $\boldsymbol{\theta}_0'$ calculated before; repeat steps iii−viii and keep appending new synthetic data into the initial training set $\mathcal{D}_0$ until $\varepsilon < \varepsilon'$.

It is worth mentioning that for any two given samples $\boldsymbol{x}$ and $\boldsymbol{x}'$, when $\boldsymbol{x}$ is distant from $\boldsymbol{x}'$, the covariance function defined in eq 13 is negligible, i.e., $k(\boldsymbol{x}_i, \boldsymbol{x}_j) \approx 0$. In such a case, this pair of samples may have aninsignificant effect on the interpolation of a GP model. To take this information into account, a second GP model adjustment criterion η can be developed, for which each of the new samples that will be considered as an addition to the original training set is further examined based on the corresponding covariance value to ensure that the latter is larger than a criterion η.

3.3. FDD Algorithm with GP Model. The main idea of the GP model based FDD is to estimate the dynamic values of faults g* using measurements $\boldsymbol{x}^*$ of sensitive variables, and then discretize the results using a minimum distance criterion for fault classification. The FDD proceeds as per the following procedures.

(a) Decide the total number of possible mean values $\{\overline{g}_i\}$ of a process by examining the constancy of measured quantities such as manipulated and/or controlled variable.

(b) Estimate these mean values $\{\overline{g}_i\}$ and their corresponding variances using collected measurements through an offline calibration step.[4]

(c) Identify measured quantities that are sensitive to faults using the active set based optimization in combination with the Latin hypercube sampling (LHS) technique as explained in section 3.1.

(d) Generate a GP model for measured variables identified from step c using the improvement in probability algorithm developed in section 3.2.

(e) New collected measurements $\boldsymbol{x}^*$ are normalized and substituted into the GP model.

(f) The mean and the variance of g* are then approximated with measurements of $\boldsymbol{x}^*$ from eqs 16 and 17, respectively.

(g) The mean of faults obtained in step f is compared to a finite set of mean values $\{\overline{g}_i\}$ estimated in step b based on a minimum distance criterion, from which the corresponding operating mode (mean value) can be identified.

The minimum distance between the estimate of g* and each of the mean values $\{\overline{g}_i\}$ can be calculated as

$$\min J_i = (g^* - \overline{g}_i)^2 \quad (24)$$

$$\text{operating mode}(\overline{g}_i) = \arg\min\{J_i\} \quad (25)$$

This criterion is performed for each estimate of faults {g*}, and then the minimum distance J_i given in eq 25 is used to identify the mean value of the fault, i.e., classification of operating mode.

4. CASE STUDY EXAMPLES

To demonstrate the efficiency of the FDD algorithm, two examples are used in this current work. In the first example, a nonlinear process involving two continuously stirred tank reactors (CSTRs) and a separator with recycle unit is used.[32,33] In the second example (section 5.4), the efficiency of FDD is studied for the Tennessee Eastman process[34,35] that has been widely used as a standard benchmark problem. The objective is to (i) identify significant measured variables that are sensitive to faults affecting a system in a stochastic fashion, and (ii) build a GP model with these identified variables for fault classification while minimizing the model dimensionality.

4.1. Reactor-Separator Chemical Process (Example 1). Figure 4 shows a schematic of the first example with three temperature control loops in this work. As seen, a stream of reactant A is added to each CSTR and converted to the final product B, where C is the side product in this chemical process. We deliberately chose this process, since it is considered sufficiently large to illustrate the computational

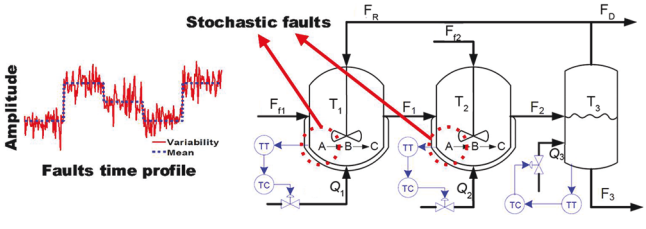


Figure 4. Schematic diagram of the first example (reactor−separator process).

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

efficiency of the methodology and it involves nonlinear behavior and uncertainty.

The feed mass fraction of reactant A (x_{A0}) is assumed to be the unknown (unmeasured) stochastic fault ($\mathbf{g}$) in this work. The changes in x_{A0} are schematically shown in Figure 4. It is assumed that perturbations around three mean values (operating modes) as described in eq 2 are normally distributed. The first-principles model of this process is used in this study in order to generate the synthetic data that can be used to identify significant measured variables and for GP model calibration. This process is controlled with three PI controllers and can be described in their velocity formulations by a set of ordinary differential equations (ODEs) as below.

$$\dot{H}_1 = (1/\rho A_1)(F_{f1} + F_R - F_1) \tag{26}$$

$$\dot{x}_{A1} = (1/\rho A_1 H_1)(F_{f1}x_{A0} + F_R x_{AR} - F_1 x_{A1}) - k_{A1}x_{A1} \tag{27}$$

$$\dot{x}_{B1} = (1/\rho A_1 H_1)(F_R x_{BR} - F_1 x_{B1}) + k_{A1}x_{A1} - k_{B1}x_{B1} \tag{28}$$

$$\dot{T}_1 = (1/\rho A_1 H_1)(F_{f1}T_0 + F_R T_R - F_1 T_1) - (1/C_p)(k_{A1}x_{A1}\Delta H_A + k_{B1}x_{B1}\Delta H_B) + (Q_1/\rho A_1 C_p H_1) \tag{29}$$

$$\dot{H}_2 = (1/\rho A_2)(F_{f2} + F_1 - F_2) \tag{30}$$

$$\dot{x}_{A2} = (1/\rho A_2 H_2)(F_{f2}x_{A0} + F_1 x_{A1} - F_2 x_{A2}) - k_{A2}x_{A2} \tag{31}$$

$$\dot{x}_{B2} = (1/\rho A_2 H_2)(F_1 x_{B1} - F_2 x_{B2}) + k_{A2}x_{A2} - k_{B2}x_{B2} \tag{32}$$

$$\dot{T}_2 = (1/\rho A_2 H_2)(F_{f2}T_0 + F_1 T_1 - F_2 T_2) - (1/C_p)(k_{A2}x_{A2}\Delta H_A + k_{B2}x_{B2}\Delta H_B) + (Q_2/\rho A_2 C_p H_2) \tag{33}$$

$$\dot{H}_3 = (1/\rho A_3)(F_2 - F_D - F_R - F_3) \tag{34}$$

$$\dot{x}_{A3} = (1/\rho A_3 H_3)(F_2 x_{A2} - (F_R + F_D)x_{AR} - F_3 x_{A3}) \tag{35}$$

$$\dot{x}_{B3} = (1/\rho A_3 H_3)(F_2 x_{B2} - (F_R + F_D)x_{BR} - F_3 x_{B3}) \tag{36}$$

$$\dot{T}_3 = (1/\rho A_3 H_3)(F_2 T_2 - (F_R + F_D)T_R - F_3 T_3) + (Q_3/\rho A_3 C_p H_3) \tag{37}$$

where the subscripts i in each equation (i.e., 1, 2, 3) refer to the vessels, x_i denotes the mass fraction of reactant A and chemical product B, respectively, T_i is the temperature in each tank, H_i is the level in each tank, F_i represents the flow rate, and the reaction terms are defined as

$$F_i = k_{vi}H_i \tag{38}$$

$$k_{Ai} = k_A \exp(-E_A/RT_i) \tag{39}$$

$$k_{Bi} = k_B \exp(-E_B/RT_i) \tag{40}$$

The recycle flow and the weight percent factors satisfy

$$F_D = 0.01F_R \tag{41}$$

$$x_{AR} = \alpha_A x_{A3}/\bar{x}_3 \tag{42}$$

$$x_{BR} = \alpha_B x_{B3}/\bar{x}_3 \tag{43}$$

$$\bar{x}_3 = \alpha_A x_{A3} + \alpha_B x_{B3} + \alpha_C x_{C3} \tag{44}$$

$$x_{C3} = 1 - x_{A3} - x_{B3} \tag{45}$$

Each of the tank in this example has an external heat input Q_i that is controlled by a PI controller:

$$Q_i(t) = Q_{(ss),i}(t) + K_{p,i}(T_{(set),i} - T_i(t)) + K_{p,i}/\tau_i \int_0^t (T_{(set),i} - T_i(t^*))\, dt^* \tag{46}$$

The descriptions of these parameters, parameter values, and controller parameters used for the computer experiments are given in Tables 1, 2, and 3, respectively.

Table 1. Process Variables

symbol	description
x_{A1}, x_{A2}, x_{A3}	mass fraction of A in vessels 1, 2, 3
x_{B1}, x_{B2}, x_{B3}	mass fraction of B in vessels 1, 2, 3
x_{C3}	mass fraction of C in vessel 3
x_{AR}, x_{BR}	mass fraction of A, B in the recycle
T_1, T_2, T_3	temperature in vessels 1, 2, 3
$T_{(set),1}$, $T_{(set),2}$, $T_{(set),3}$	temperature set point in vessels 1, 2, 3
T_0	feed stream temperature
F_{f1}, F_{f2}	feed stream flow rates to vessels 1, 2
F_1, F_2, F_3	effluent stream flow rates to vessels 1, 2, 3
F_R, F_D	flow rate of recycle and purge
H_1, H_2, H_3	level of vessels 1, 2, 3
Q_1, Q_2, Q_3	manipulated input in vessels 1, 2, 3
ΔH_A, ΔH_B	heats of reaction
k_{Ai}, k_{Bi}	pre-exponential values of reactions 1, 2
α_A, α_B, α_C	relative volatilities of A, B, C
E_A/R, E_B/R	ratio of activation energy and gas constant for reactions 1, 2
A_1, A_2, A_3	cross-section area of vessels 1, 2, 3
C_p, ρ	heat capacity, solution density

4.2. Tennessee Eastman Process (Example 2). The Tennessee Eastman (TE) benchmark process has five major units as shown in Figure 5, i.e., a product condenser, a recycle compressor, a product stripper, a vapor−liquid separator, and a reactor.[34] This process in total has 41 measured variables, 12 manipulated variables, and 20 disturbances that can be considered as faults. The decentralized multiloop control strategy[36] is used in this example. In this current work, the A/C feed ratio and B composition in stream 4, i.e., load disturbance IDV(1), is defined as the stochastic faults ($\mathbf{g}$) to demonstrate the efficiency of the proposed FDD algorithm. Since large perturbations in the feed can be harmful, three different mean values in IDV(1) are considered in this work. The smallest mean value is treated as a normal operation, while the rest denote the faulty operating conditions. The detailed description of the normal and faulty operations will be discussed in section 5.4. Note that the objective in this example is to identify the mean value of the feed IDV(1) in the presence of feed perturbations. The multiple-root-cause fault classification is discussed in the Supporting Information.

5. RESULTS AND DISCUSSION

5.1. Faults Distribution and LHS Sampling. The goal in the first example is to identify and diagnose (classify) the mean value (operating mode) of the unmeasured (unknown) feed mass fraction x_{A0} with available data of easily measured quantities. For clarification, three mean values of the feed mass fraction (x_{A0}) are considered in this example, i.e., x_{A0} = 0.65,

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

Table 2. Parameter Values of Process Variables

symbol	value	units	symbol	value	units	symbol	value	units
F_{f1}	10	kg/s	k_{v1}	2.5	kg/m s	ρ	0.15	kg/m^3
F_{f2}	1	kg/s	k_{v2}	2.5	kg/m s	A_1	3	m^2
F_R	60	kg/s	k_{v3}	2.5	kg/m s	A_2	3	m^2
$T_{(set),1}$	315	K	k_A	0.02	1/s	A_3	1	m^2
$T_{(set),2}$	315	K	K_B	0.018	1/s	α_A	3.5	
$T_{(set),3}$	400	K	E_A/R	−1000	K	α_B	1.1	
T_0	310	K	E_B/R	−500	K	α_c	0.5	
T_R	310	K	ΔH_A	−40	kJ/kg			
C_p	2.5	kJ/kg K	ΔH_B	−50	kJ/kg			

Table 3. Controller Parameters

	vessel					
	1		2		3	
	$K_{p,1}$	τ_1	$K_{p,2}$	τ_2	$K_{p,3}$	τ_3
value	0.25	0.0025	0.25	0.0025	0.25	0.0025

0.75, and 0.85 (n_g = 3 in eq 2). It is assumed that stochastic perturbations in x_{A0} occur around each aforementioned mean value, which follows a normal distribution with zero mean and a standard deviation of 0.1.

To reduce the number of simulation runs for generating the synthetic training set, the Latin hypercube sampling (LHS) technique is used to encompass the entire domain of faults and to reduce computational burden. The LHS can account for the previously generated sample points, as compared to Monte Carlo (MC) sampling techniques. This ensures that samples are evenly distributed in the domain. A comparison between LHS and random sampling with MC is shown in Figure 6.

For clarification, two mean values of faults are used in Figure 6, and five samples are generated around each mean with both

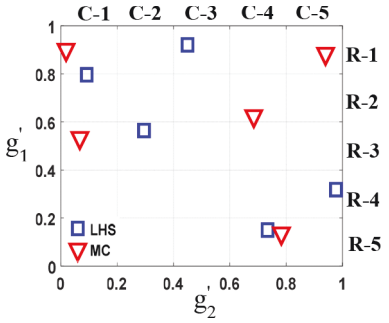


Figure 6. Samples generated with LHS and MC (R, row; C, column).

LHS and MC methods. For LHS and MC techniques, each sample is normalized with respect to the maximum value of the five samples for illustration. The two-dimensional fault domain is divided into 5 × 5 subdomains with the LHS. As shown, the fault domain is evenly covered. One sample is found in each row and each column for LHS, whereas no samples fall into the fourth row, the second and the third columns, with MC. Thus, more samples are required for MC. This shows that the

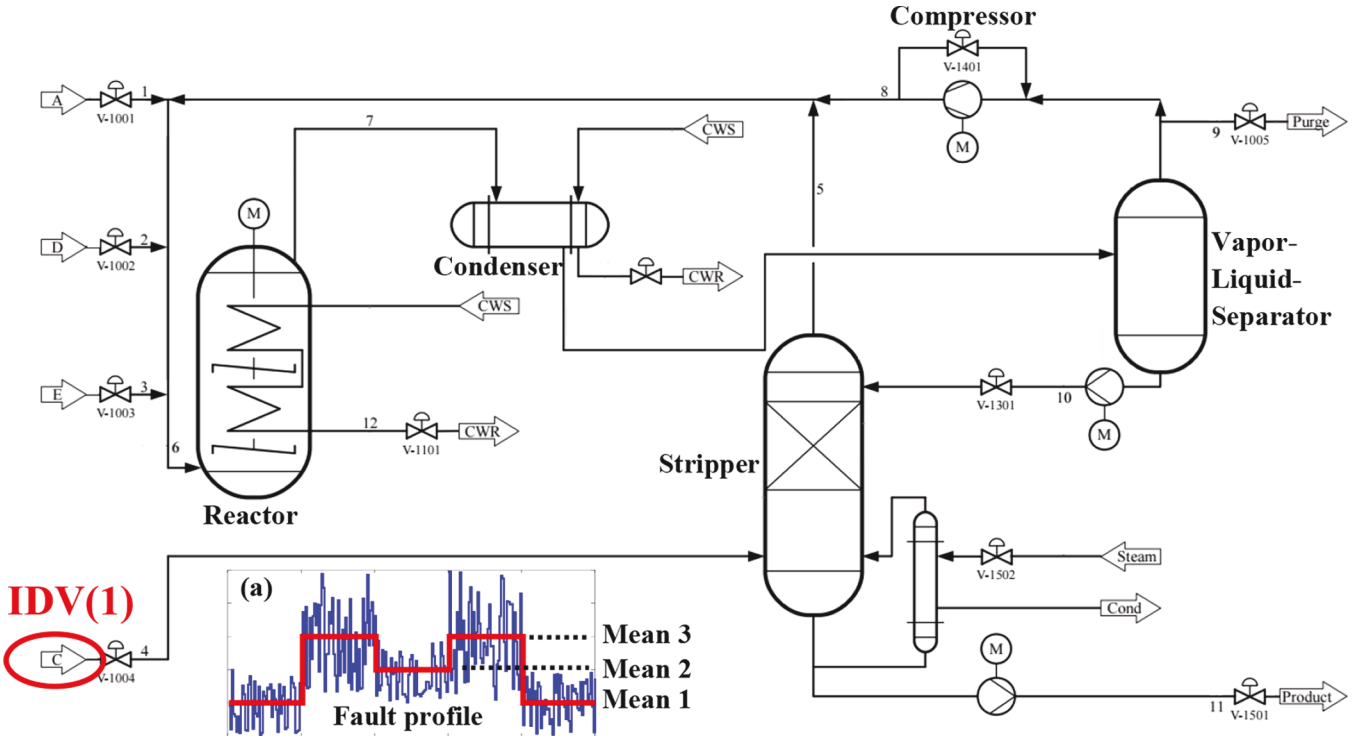


Figure 5. Tennessee Eastman (TE) process described by Bathelt et al.[34]

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

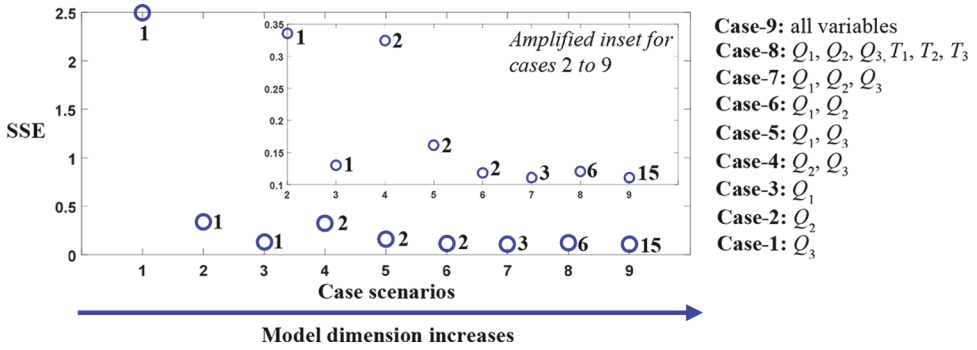


Figure 7. Illustration of trade-off between model dimensionality and model prediction accuracy. A number beside a symbol denotes the total number of measured variables used for calculating SSE.

computational cost for generating a training set used for sensitivity analysis can be reduced with LHS, which is particularly critical for problems involving large number of variables.

5.2. Case Study 1. Single Stochastic Fault (Example 1). *5.2.1. Optimal Selection of Measured Variables for FDD.* To optimally choose training data for FDD, the active set optimization is used as explained in section 3.1. To generate a training set, 100 samples with the LHS technique are generated around each mean value of x_{A0}. This results in a training set $\mathcal{D}$ involving 300 samples of three faults, i.e., $M = n_g \times n_s = 3 \times 100 = 300$. All variables including the faults are normalized with respect to the mean value of 300 samples.

For the *cross-validation procedure*, one-third of the samples in the training set $\mathcal{D}$, i.e., 100 pairs of measurements of faults and measured variables, are used to solve $\{\beta_k\}$ in eq 21. The remaining two-thirds of the samples are used as validation data to evaluate the model accuracy that is assessed by the residual sum of squared error (SSE) between the predictions and simulated measurements of faults with the validation data. The SSE is also chosen as a criterion to find the trade-off between the model dimensionality and model accuracy. The *cross-validation procedure* is repeated 100 times in this case study, and the average of SSE is recorded. It should be noted that, for each *cross-validation procedure*, one-third of the data is randomly chosen from the training set $\mathcal{D}$ and there are no identical training sets.

To find the most sensitive variables, this method is evaluated with different initial subsets of measured variables, i.e., p' in section 3.1. First, it is assumed in this case study that all variables can be measured. Thus, there are 15 unknown coefficients $\{\beta_k\}$ in eq 20, i.e., $p = p' = 15$ and $k = 1, ..., p$. Each coefficient β_k determines the contribution of the kth parameter to the total SSE. It is found that the regression coefficients $\{\beta_k\}$ of the level $\{H_i\}$ in three vessels are 0 as expected, since the perturbations in mass fraction x_{A0} have negligible effect on the level. In addition, the regression coefficients of the mass fractions of A and B (i.e., $\boldsymbol{x}_i$) are smaller in each tank in this case study, as compared to the coefficients of temperature and external heat. To reduce the model dimensionality, a threshold $\zeta_1 = 0.001$ is used to remove the measured variables that have smaller coefficients, i.e., the mass fractions of reactant A and product B. When all the variables are used in eq 21, the average of SSE for the validation data for 100 repeated *cross-validation procedures* is found to be ~0.1108.

Based on the results above, eight additional cases are investigated. Figure 7 shows the initial subset of variables used in each case scenario and the sensitivity analysis results, where the number beside a symbol represents the total number of measured variables used in eq 21 in section 3.1. The sensitivity analysis is evaluated with the SSE for each case and used to determine the variables that can be used for FDD. For example, it is found that the temperatures $\{T_i\}$ in case 8 have negligible effect on the model predictions, since the regression coefficients $\{\beta_k\}$ of $\{T_i\}$ are smaller as compared to the manipulated variable $\{Q_i\}$. This result is expected since the temperatures are controlled variables and less sensitive to the variations in the feed x_{A0}.

As seen in Figure 7, the SSE for case scenarios 3, 6, 7, 8, and 9 are very similar. For example, the SSE is ~0.1201 for case 8, when all measurements of $\{T_i\}$ and $\{Q_i\}$ are used. In contrast, the SSE is ~0.13 when only the external heat of Q_1 is used to estimate the faults in case 3. In addition, as compared to the above-mentioned study where all variables are used for model predictions (case 9 in Figure 7), the difference in SSE is negligible. As mentioned, the SSE is found to be ~0.1108 when 15 variables are used in eq 20. As compared to the case 3 here, the SSE is only slightly increased from ~0.1108 to ~0.13.

The sensitivity analysis with active set optimization is also investigated for two additional case studies; i.e., Q_2 and Q_3 are used alone in eq 20. As shown in Figure 7, the SSE values of cases 1 and 2 are clearly larger as compared to the others. Due to the small difference in the prediction errors, only Q_1 is used for model calibration in this case study, since it is found to be the most sensitive variable.

5.2.2. Calibration of GP Model. Single Fault. Using the measurements of Q_1, an initial training set $\mathcal{D}_0 = \{(\mathbf{Q}_{1,i}, \mathbf{g}_i)\}$ $(i = 1, ..., 30)$ with 30 pairs of synthetic samples are generated. The model calibration results using the model adjustment algorithm explained in section 3.2 are given in Table 4. For the improvement in probability, the model discrepancy is defined as $\varepsilon' = 1 \times 10^{-2}$ in this case study, and a 15 percent point probability improvement is used. Also, the second model adjustment criterion η is set to 2×10^{-2}.

Table 4. Hyperparameters of GP Model

model	σ_G	l	σ_g
no adjustment	1.0473	4.2163	0.1001
with adjustment	2.0086	4.5724	0.1010

As can be seen in Table 4, the GP model parameters obtained from the minimal model adjustment algorithm in section 3.2 are different from those computed without the GP model adjustment procedure. The efficiency of the model

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

adjustment technique in this work will be further discussed in terms of fault detection rate below. For the improvement in the probability-based model adjustment, 206 sets of actual measurements are simulated and 178 pairs of them (Q_1 and g) are added to the initial training set $\mathcal{D}_0$ following the steps in section 3.2.

5.2.3. Evaluation of Performance. To evaluate the efficiency of FDD, two indices are used,[37] i.e., fault classification rate r_d and false alarm rate r_f:

$$r_d = d_i/D_{To} \tag{47}$$

$$r_f = m_i/D_{Ff} \tag{48}$$

where d_i is the number of testing samples (e.g., x_{A0}) that have been correctly identified, m_i is the number of samples that indicates the occurrence of faults, but a fault has not occurred actually, D_{Ff} is the total number of fault free testing samples, and D_{To} is the number of testing samples including both faulty and normal samples that are used for FDD.

It should be noted that the focus of this work is to identify the switch between a normal operating mode and a faulty operating mode in a process of an intermittent manner; thus samples representing the normal operating mode are also included in d_i for the evaluation of FDD performance. In other words, the fault classification rate is used to evaluate the capability of correctly identifying either a normal or a faulty operating mode with available measurements. To evaluate the FDD performance, a misdetection rate can be used. In this work, the result of the misdetection rate is not given for brevity. However, it can be easily calculated with the fault classification rate. For example, the classification rate of the normal and faulty operating modes with an FDD algorithm in principle would be 1; i.e., all the normal and faulty operating modes can be accurately identified. However, there is misdetection due to uncertainty such as measurement noise, and the classification rate is often smaller than 1. Thus, the misdetection rate can be estimated by calculating the difference between 1 and the fault classification rate r_d obtained in eq 47.

The classification rate and the false alarm rate in this current work are evaluated for two different case scenarios: (i) GP model that is calibrated without the adjustment technique in section 3.2 and (ii) GP model that was calibrated with the probability improvement algorithm. The classification rates r_d are summarized in Table 5, while the false alarm rates r_f are given in Table 6.

Table 5. Summary of the Fault Classification Rate

	noise level		
method	1%	2%	3%
GP	0.80	0.77	0.70
GP adjustment	0.88	0.86	0.83

Table 6. Summary of the False Alarm Rate

	noise level		
method	1%	2%	3%
GP	0.10	0.14	0.16
GP adjustment	0.06	0.09	0.12

As seen in Tables 5 and 6, the fault classification rates decrease as the noise level increases, while the false alarm rates increase. For example, as shown in Table 5, the GP model calibrated with the probability improvement algorithm can consistently provide better performance, as compared to a model without using the model adjustment algorithm. This confirms that the design of the training set is instrumental for FDD performance.

5.3. Case Study 2. Single Fault and Model−Plant Mismatch (Example 1). In the previous case studies, it is assumed that the process is only affected by the perturbations in the feed mass fraction x_{A0}, and the other components are perfectly known and accurate to a modeler. In contrast, this case study investigates the FDD in the presence of one parametric stochastic fault and unknown model−plant mismatch. As done in a previous section, perturbations are assumed to be superimposed on step changes of mean values in the feed x_{A0} as shown in Figure 8. To introduce model−plant mismatch, it is assumed that the ratio between the activation energy and the universal gas constant in eq 39, i.e., E_A/R, is not a fixed constant value. Instead, the ratio E_A/R follows a uniform distribution and varies randomly between −980 and −1020 K with respect to time. However, this uncertainty is assumed to be a priori unknown to the modeler.

The objectives are to (i) identify measured variables that are sensitive to these variabilities in the feed and the ratio of E_A/R, (ii) build a GP model to improve FDD performance, and (iii) diagnose the mean of $\overline{g}_i$ of the feed mass fraction x_{A0}. To illustrate the variations in both the ratio E_A/R and the feed x_{A0}, three mean values of x_{A0}, i.e., 0.65, 0.75, and 0.85, are used as investigated in the previous case study. Figure 8 shows the perturbations in E_A/R and x_{A0} for clarification.

5.3.1. Selection of Measured Variables for FDD. Similar to the single fault study, the active set optimization in combination with statistical analysis are used to identify sensitive variables. To generate a synthetic training set, 100 samples are generated with the LHS technique for each mean value of x_{A0}. This produces a training set $\mathcal{D}$ with 300 samples of x_{A0}. For each sample of x_{A0}, a random value of the ratio E_A/R is sampled from the range of −980 to −1020 K to simulate data to be used for sensitivity analysis. All variables including the faults are normalized with respect to the mean value of 300 samples.

For the *cross-validation procedure*, a (one-third) portion of the training set $\mathcal{D}$, i.e., 100 pairs of measurements of faults and measured variables, is used to solve $\{\beta_k\}$ in eq 20. The remaining 200 samples are used as validation data to calculate the SSE. The *cross-validation procedure* is repeated 100 times and the average of SSE for different initial sets of variables is recorded. This is used as a criterion to find the trade-off between the model dimensionality and prediction accuracy as done in the first case study. Figure 9 summarizes the variables used in each initial subset and the SSE results. The number beside a symbol means the total number of measured variables used for sensitivity analysis.

As shown in Figure 9, the SSE decreases as the model dimensional increases. For example, the SSE for case 10 is found to be ∼0.1441, where 15 variables are used in the initial set in eq 20. Similar to the first case study, the regression coefficients $\{\beta_k\}$ of the level $\{H_i\}$, mass fraction of reactant A and product B (i.e., $\boldsymbol{x}_i$) are zero or have negligible effect on the prediction of fault x_{A0}. Thus, additional studies (cases 1−7), focusing only on the external heat $\{Q_i\}$, are investigated to find the trade-off between model dimensionality and prediction accuracy.

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

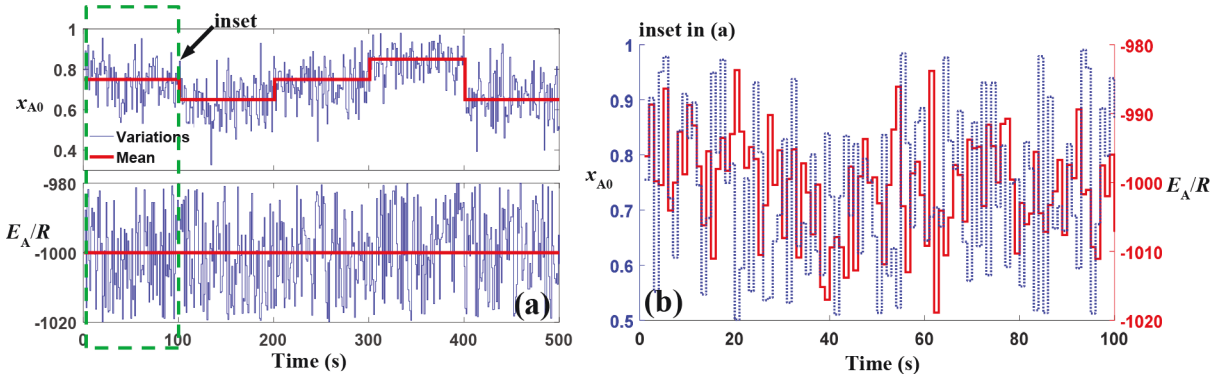


Figure 8. Profiles of input faults x_{A0} and variations in Arrhenius equation parameter E_A/R.

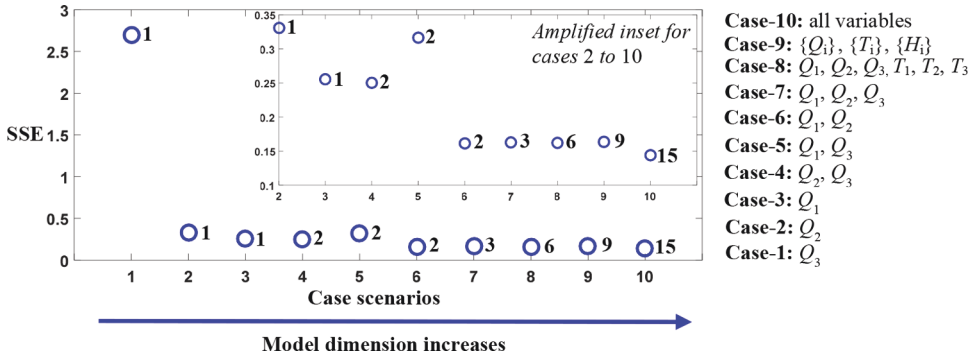


Figure 9. Illustration of trade-off between model dimensionality and model prediction accuracy. A number beside a symbol denotes the total number of measured variables used for calculating SSE.

For the single fault study in the previous section, we found that one measured variable Q_1 can provide sufficiently accurate results, as compared to the case where all variables are used for predictions. On the other hand, it is found that Q_1 alone, i.e., case 3 in Figure 9, fails to provide accurate predictions for the current case study involving additional uncertainty. For example, the SSE is ∼0.2557 for case 3, which is significantly higher than case 10. It is also found that the SSE values for cases 6−9 are quite similar, varying between ∼0.1612 and ∼0.1633. In summary, Q_1 and Q_2 as chosen in case 6 are identified as the sensitive variables in this case study.

5.3.2. GP Model Calibration. An initial training set $\mathcal{D}_0 = \{(\mathbf{Q}_{1,i}, \mathbf{Q}_{2,i}, \mathbf{g}_i)\}$ $(i = 1, ..., 50)$ with 50 sets of measurements is generated. Following the GP model adjustment method in section 3.2, the GP model parameters are given in Table 7. The model discrepancy criterion is defined as

Table 7. Summary of Hyperparameters for GP Model

model	σ_G	l	σ_g
no adjustment	1.8963	3.7231	0.1047
with adjustment	2.0199	4.7120	0.1103

$\varepsilon' = 1 \times 10^{-2}$, a 15% improvement in probability is used, and the second model adjustment criterion η given in section 3.2 is set to 2×10^{-2}.

For the improvement in the probability-based GP model adjustment, 343 set of synthetic samples are required and 295 pairs of them (Q_1, Q_2, and g) are appended to the initial training data set $\mathcal{D}_0$. As expected, more training data are required for the model calibration as two sources of uncertainty are investigated in this case study, thus necessitating more data points to satisfy the accuracy criterion defined in eq 22.

5.3.3. Evaluation of Performance. To compare the fault classification results, two different case scenarios are studied in this work. For the first case scenario, the GP model is only calibrated with an initial training set $\mathcal{D}_0$. The hyperparameters shown in Table 7 are used to generate a two-dimensional GP regression model of Q_1 and Q_2. For comparison, the hyperparameters with the improvement in the probability method are used to build a second GP model. For each case studies, 1000 testing samples for each mean value of the feed mass fraction x_{A0} are used to study the detection rate. To introduce variations in the Arrhenius equation (i.e., E_A/R), different E_A/R values are used for each sample of x_{A0}. These values are randomly selected from the range of −980 to −1020 K. For the simulations, all measurements are generated from the model predictions corrupted with Gaussian noise.

For the first case scenario, the fault classification rate is found to be ∼0.81 for the GP model with a model adjustment step, whereas the fault classification rate is ∼0.70 for the GP model calibrated only with an initial training set. The fault classification rate is increased by ∼11 percent points with the model adjustment steps outlined in section 3.2. This confirms that the GP model is very sensitive to the density and amount of data used for calibration. The false alarm rates are ∼0.11 and ∼0.14 for the GP model calibrated with a model adjustment procedure and the GP model without adjustment, respectively.

5.4. Case Study 3. FDD of the Tennessee Eastman Benchmark Process (Example 2). *5.4.1. Selection of Measured Variables for FDD.* To illustrate the efficiency, the Tennessee Eastman process is used as a standard benchmark problem in this case study. The FDD algorithm

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

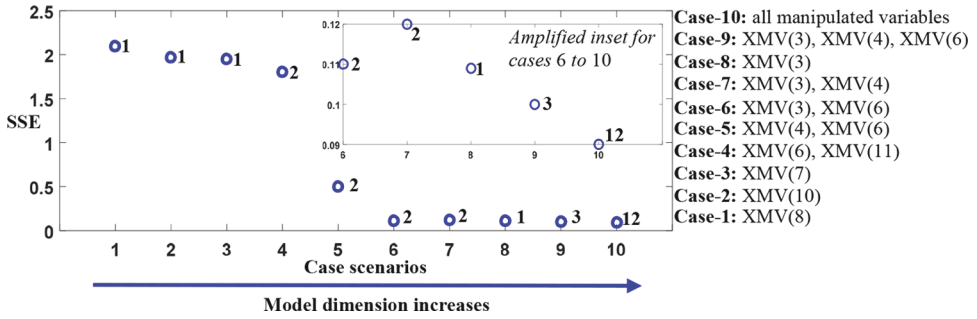


Figure 10. Illustration of trade-off between model dimensionality and model prediction accuracy of the TE process. A number beside a symbol denotes the total number of manipulated variables used for calculating SSE.

is used to identify and diagnose the operating modes (mean values) when the system is operated at steady states in the presence of feed perturbations. The load disturbance IDV(1), i.e., A/C feed ratio, B composition constant in stream 4, is assumed to be the unknown stochastic fault ($\mathbf{g}$) in this case study. Three different operating modes, as shown in Figure 5, are defined in terms of the percent point of perturbations, since significant variations in the feed IDV(1) may affect the product quality, thus resulting in economic loss. Three mean values of IDV(1) are defined as 5, 7, and 9%, respectively. It is assumed here that the perturbations around each mean value follow a normal distribution with a standard deviation of 1 percent point. The first mean value is used to represent a normal operation, while the other two mean values are two different faulty operating modes.

Similar to previous case studies, the active set optimization is first used to identify sensitive variables that can be used for FDD. To build the training set, 100 samples are generated with the LHS technique for each mean value of IDV(1), resulting in a data set $\mathcal{D}$ involving 300 samples. All variables including the fault variables are normalized with respect to the mean value of 300 samples. For the *cross-validation procedure*, a (one-third) portion of the training set $\mathcal{D}$, i.e., 100 pairs of measurements of faults and measured variables, is used to solve $\{\beta_k\}$ in eq 20. The remaining 200 samples are used as validation data to calculate the SSE. The *cross-validation procedure* is repeated 100 times, and the average of SSE for different initial sets of variables is recorded. For clarification, Figure 10 summarizes different initial sets of manipulated variables that were used for variable selection and their corresponding SSE results. The number beside a symbol represents the total number of manipulated variables used for sensitive variable selection in each case study.

As shown in Figure 10, the SSE generally decreases as the model dimensions increase. For example, the SSE for case 9 is found to be ∼0.1001, where three variables are used in the initial set in eq 20. In addition, it was found that the regression coefficients $\{\beta_k\}$ of the compressor recycle valve XMV(5), stripper steam valve XMV(9), and agitator speed XMV(12) are zero and they have negligible effect on the prediction of faults IDV(1). The regression coefficients of the separator pot liquid flow XMV(7), stripper liquid product flow XMV(8), and reactor cooling water flow XMV(10) are smaller, compared to the coefficients of A feed flow XMV(3), A and C feed flow XMV(4), and purge valve XMV(6). Thus, additional case studies, i.e., cases 5−9 focusing only on XMV(3), XMV(4), and XMV(6), are investigated to find the trade-off between model dimensionality and the prediction accuracy. The SSE results are compared with a case study where all manipulated variables are used in eq 20. Note that the SSE values in these case studies are quite similar, varying between ∼0.09 and ∼0.12. To maintain model simplicity, the A feed flow XMV(3) is chosen as the sensitive variable in this case study that can be used for FDD. The selection of this variable is in good agreement with prior knowledge about the process, since the manipulated variable, i.e., the feed flow of A, should be changed to eliminate any perturbations in the A/C feed ratio IDV(1).[35,38]

5.4.2. Calibration of GP Model and Evaluation of FDD Performance. For the model calibration, an initial set of measurements $\mathcal{D} = \{(\boldsymbol{x}_{s,i}, \boldsymbol{g}_i)\}$ (i = 1, ..., 50) with 50 pairs of measurements around each mean value are formulated. To evaluate and compare the FDD performance, two different case scenarios were investigated. For the first case study (GP-1), only the initial training set $\mathcal{D}$ of measurements was used to identify the *hyperparameters* $\boldsymbol{\theta}$, whereas the *hyperparameters* for the second case study (GP-2) were determined with the model adjustment algorithm as explained in section 3.2. For the first case study here, the *hyperparameters* of GP-1 are σ_G = 0.6879, l = 1.3743, and σ_g = 0.1052, respectively. For the second case study, 201 additional sets of measurements were simulated, and 157 sets of them were ultimately appended to the initial training set $\mathcal{D}$ based on covariance values. The *hyperparameters* of the second case study (GP-2) are σ_G = 1.1036, l = 1.8147, and σ_g = 0.0901, respectively.

The efficiency of the model adjustment is further studied in terms of the fault classification rate and the false alarm rate with the TE process. It was found that the fault classification rate is ∼0.89 using the GP model calibrated with the model adjustment steps, while the fault classification rate is ∼0.81 for the GP model which was only calibrated with the initial training set $\mathcal{D}$. This indicates that the empirical model-based FDD such as the GP model in this work is sensitive to the training data set, and it is important to optimally select model parameters with appropriate training data. The fault detection results also confirm that the model adjustment procedures in this current work can improve the FDD performance for large-scale nonlinear chemical processes. Similar to the case studies above, the model of the TE process (GP-2), calibrated with the model adjustment steps in section 3.2, has lower false alarm rates, i.e., ∼0.08, as compared to GP-1, which has a false alarm rate of ∼0.15 without the adjustment procedures.

5.5. Comparison and Discussion. In this section, the GP model-based FDD is compared with the other surrogate model based methods, and the fault classification rate was used to evaluate the FDD performance. For the fault (x_{A0}) in the first example, the accuracy of a GP model developed with an

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

Table 8. Summary of Fault Classification Rate with Different Models (Example 1)

	PLS-1	PLS-2	PLS-3	Lasso-1	Lasso-2	ANN	GP-1	GP-2
r_{rate}	0.82	0.78	0.83	0.67	0.61	0.83	0.81	0.88

improvement in the probability method is compared to the partial least squares (PLS) regression, a linear Lasso regression model generated with the active set optimization approach, and the neural network.

The PLS model generates a linear regression model by projecting the predicted variables and the measured variables into a reduced subspace. Variables generated from the subspace generally do not have direct physical explanation and are not easy to interpret. The active set optimization can find a linear relationship between faults and measured variables as explained in section 3.1.

Different case scenarios are investigated. Three models are developed with the PLS. All variables are used for model training, and all the components after the linear projection are remained to find the relation between faults and data, i.e., PLS-1. For the second model, PLS-2, although all variables are used for model training, only the first principal component is used to identify the relation between faults and measurements. A third model, PLS-3, is built to find the relation between the faults and the most sensitive variable (Q_1) identified from the active set optimization. For the Lasso regression model, two models are generated with the active set optimization. In the first model to be referred as Lasso-1, all the variables are used in eq 20 to build the relationship between faults and measured quantities. For model Lasso-2, to reduce the model dimensionality, only the most sensitive variable (Q_1) is used to identify the relationship between the faults and external heat. Note that 300 sets of simulated measurements were used for the calibration of the PLS and Lasso models.

Two GP models were developed to compare the FDD results. With an initial training data set of 30 pairs of measured quantities, all variables are used for model calibration for the first GP model (GP-1). However, the improvement in probability of model adjustment method is not used. Using the sensitivity analysis results obtained with active set optimization, a second GP model (GP-2) is generated with the most sensitive variable Q_1 and the improvement in probability of model adjustment procedure is used. To further demonstrate the capability of the GP model, it was compared with a neural network that is another type of nonlinear surrogate model. The artificial neural network (ANN) toolbox of Matlab was used. For ANN training, 300 sets of data of sensitive variables were used. The ANN involved two hidden layers, and the *tan-sigmoid* function ("*tansig*") was used as the activation function. The rest of the parameters in the network were set up by using the default setting in the ANN toolbox. For example, the "*trainlm*" was used as the training function. Table 8 summarizes the comparison results of these studies in terms of fault detection rate.

In Table 8, the variation in x_{A0} follows the same assumptions as explained in section 5.2, and 1% measurement noise is used for simulations. As can be seen, the linear models generated with Lasso are found to be inaccurate and lead to a higher misdetection rate. The PLS models outperform Lasso models because PLS can address correlations better than Lasso models.

As compared to PLS and Lasso models, GP models provide more accurate fault detection results. The results obtained with ANN are ∼0.83, which are similar to the fault detection results obtained with the GP models. However, it is worth mentioning that 300 measurements were used for the calibration of PLS-1 and ANN models, which is larger than GP models. As compared to the ANN model, the GP model in this work has fewer training parameters and can provide a measure of confidence interval in prediction.[25,39] Also, for the single fault case study as investigated in section 5.2, the fault classification rate was found to be ∼0.80, when only the most sensitive variable is used to generate GP model without an adjustment procedure. The fault classification rate only increases by ∼1 percent point in Table 8 with GP-1, when all the variables were used to calibrate the models. This clearly shows the importance of identifying measured variables that are more sensitive to variations in faults, since the classification rate can only be improved slightly with a larger number of measurements. The fault classification rate for the second GP model (GP-2), which was generated with the most sensitive variable and a model adjustment step, is ∼0.88. This clearly indicates that the GP model in combination with the Lasso-based sensitivity analysis can provide more accurate fault detection results, as compared to other empirical models in this work.

Additionally, although simultaneous faults are less common than single faults,[40] the FDD method is validated with multiple changes in a process. Especially, our objective is to investigate whether changes such as mean value shifts in other disturbances would be mistaken as mean value changes in the chosen fault (process variable) of interest. For the reactor−separator process, the feed mass fraction of reactant A is assumed to be the fault of interest, but the gas constant (i.e., E_A/R) in eq 39 can exceed the range defined in case study 2 in section 5.3. That means random values of E_A/R were sampled from the range −960 to −1040 K other than the range between −980 and −1060 K used in case study 2, thus producing a different mean value. However, such a change was assumed to be unknown to the modeler and measurements were not available for the GP model calibration. In this case study, it was found that the fault classification rate was ∼0.74, which is approximately 5 percent points lower than the results obtained in section 5.3. We further validated the algorithm, by assuming that measurements when E_A/R exceeds its assumed range are available, and these measurements were then used for recalibrating the GP model. It was found that a similar fault classification rate, i.e., ∼0.79, can be obtained. This confirms that the performance of GP models is sensitive to the data used for model calibration, and therefore careful design of data to be used for calibration is essential for detection accuracy. It should be noted that the focus in this case study is to identify and diagnose the mean value changes in the feed mass fraction other than the mean value changes in the gas constant. When the changes between mean values and fault magnitude for multiple-root-cause faults have to be identified and estimated, the GP approach can be extended to a fault classification problem. An additional case study is presented in the Supporting Information on how to extend the approach to a classification problem, which is not discussed here for brevity.

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977

6. CONCLUSION

In this work, the Gaussian process (GP) theory is combined with the active set optimization to build a surrogate model to identify and diagnose intermittent stochastic faults. Three important aspects of surrogate modeling based fault detection and diagnosis (FDD) have been discussed, i.e., model simplicity, interpretability, and calibration. The proposed method can find a trade-off between model dimensionality and fault classification accuracy. Different from the dimension reduction techniques in the literature, the sensitivity analysis approach in this work does not rely on a transformed subspace of measured variables, and it is performed with the physical variables of the problem, thus facilitating interpretability of surrogate models. To ensure sufficient data for GP model calibration, an improvement in probability-based model adjustment algorithm is developed to improve the model accuracy. It is demonstrated that the GP model in combination with the active set optimization based sensitivity analysis can provide more accurate fault classification results, as compared to other techniques such as multivariate analysis, Lasso-based linear surrogate models, and artificial neural network.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.iecr.8b01110.

Gaussian process (GP) model based multiple fault identification and classification (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: 315-268-2284 E-mail: ydu@clarkson.edu.

ORCID

Yuncheng Du: 0000-0003-3652-7878

Hector Budman: 0000-0002-0773-7457

Funding

H. Budman and T. Duever acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support in this work. D. Du acknowledges the Natural Science Foundation (NSF-CMMI-1646664 and NSF-CMMI-1728338) for financial support. Y. Du acknowledges the Natural Science Foundation (NSF-CMMI-1727487) for financial support.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Gerlter, J. *Fault Detection and Diagnosis in Engineering Systems*; Taylor & Francis: 1998.

(2) Isermann, R. Model based fault detection and diagnosis - status and applications. *Annu. Rev. Control* **2005**, *29*, 71−85.

(3) Venkatasubramanian, V.; Rengaswamy, R.; Yin, K.; Kavuri, S. A review of process fault detecion and diagnosis Part I: Quantitative model-based methods. *Comput. Chem. Eng.* **2003**, *27*, 293−311.

(4) Du, Y.; Duever, T.; Budman, H. Fault detection and diagnosis with parametric uncertainty using generalized polynomial chaos. *Comput. Chem. Eng.* **2015**, *76*, 63−75.

(5) Wong, W. C.; Lee, J. H. Fault detection and diagnosis using hidden markov disturbance models. *Ind. Eng. Chem. Res.* **2010**, *49* (17), 7901−7908.

(6) Li, X.; Yang, G. Fault detection for linear stochastic systems with sensor stuck faults. *Optimal Ccontrol Appl. Methods.* **2012**, *33* (1), 61−80.

(7) Du, Y.; Duever, T.; Budman, H. Gerenalized polynomial chaos-based fault detection and classification for nonlinear dynamic processes. *Ind. Eng. Chem. Res.* **2016**, *55* (7), 2069−2082.

(8) Kim, Y.; Lee, S. J.; Park, T.; Lee, G.; Suh, J. C.; Lee, J. M. Robust leak detection and its localization using interval estimation for water distribution network. *Comput. Chem. Eng.* **2016**, *92* (2), 1−17.

(9) Chiang, L. H.; Jiang, B.; Zhu, X.; Huang, D.; Braatz, R. D. Diagnosis of multiple and unknown faults using the causal map and multivariate statistics. *J. Process Control* **2015**, *28*, 27−39.

(10) Feng, Z.; Qin, J. S.; Liang, M. Time-frequency analysis based on Vold-Kalman filter and higher order energy separation for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions. *Renewable Energy* **2016**, *85*, 45−56.

(11) Jiang, Q.; Huang, B.; Ding, S. X.; Yan, X. Bayesian fault diagnosis with asynchronous measurements and its application in networked distributed monitoring. *IEEE Trans. Ind. Electron.* **2016**, *63* (10), 6316−6324.

(12) Du, Y.; Budman, H.; Duever, T. Comparision of stochastic fault detection and classification algorithms for nonlinear chemical processes. *Comput. Chem. Eng.* **2017**, *106*, 57−70.

(13) Isermann, R. *Fault Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*; Springer: Berlin, Germany, 2006.

(14) Du, Y.; Budman, H.; Duever, T. Integration of Fault Diagnosis and Control by Finding a Trade-Off between the Observability of Stochastic Faults and Economics. *The 19th World Congress of the International Federation of Automatic Control, Cape Town, South Africa*; International Federation of Automatic Control: 2014; pp 7388−7393.

(15) Du, Y.; Budman, H.; Duever, T. Integration of fault diagnosis and control based on a trade-off between fault detectability and closed loop performance. *J. Process Control* **2016**, *38*, 42−53.

(16) Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.* **1996**, *58* (1), 267−288.

(17) Osborne, M.; Presnell, B.; Turlach, B. A. On the Lasso and its dual. *J. Comput. Graph. Stat.* **2000**, *9* (2), 319−337.

(18) Osborne, R.; Presnell, B.; Turlach, B. A. A new approach to variable selection in least squares problems. *IMA J. Numer. Anal.* **2000**, *20*, 389−403.

(19) Patton, R. J.; Frank, P. M.; Clark, R. N. *Issues of Fault Diagnosis for Dynamic Systems*; Springer: 2010.

(20) Serradilla, J.; Shi, J. Q.; Morris, A. J. Fault detection based on Gaussian process latent variable models. *Chemom. Intell. Lab. Syst.* **2011**, *109* (1), 9−21.

(21) Kocijan, J.; Girard, A.; Banko, B.; Murray-Smith, R. Dynamic systems identification with Gaussian processes. *Math. and Comput. Model Dynamic. Sys.* **2005**, *11* (2), 411−424.

(22) Pruher, J.; Simandl, M. Gaussian process based recursive system identification. *J. Phys.: Conf. Ser.* **2014**, *570*, 012002.

(23) Kocijan, J.; Murray-Smith, R.; Rasmussen, C.; Girard, A. Gaussian process model based predictive control. *Proceedings of the American Control Conference, Boston, MA*; IEEE: 2004; p 2214. DOI: 10.23919/ACC.2004.1383790.

(24) Cao, G.; Lai, E.; Alam, F. Gaussian process model predictive control of unknown nonlinear system. *IET Control Theory* **2017**, *11* (5), 703−713.

(25) Rasmussen, C. E.; Williams, C. K. L. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, 2006.

(26) Yang, X.; Maciejowski, J. M. Fault tolerant control using Gaussian processes and model predictive control. *Int. J. Appl. Math.Comput. Sci.* **2015**, *25* (1), 133−148.

(27) Ljung, L. *System Identification - Theory for the User*, 2nd ed.; Prentice-Hall: 1999.

(28) Shi, J. Q.; Choi, T. *Gaussian Process Regression Analysis for Functional Data*; Chapman & Hall CRC: London, 2011.

(29) Deisenroth, M. *Efficient Reinforcement Learning Using Gaussian Processes*; KIT Scientific Publishing: Karlsruhe, 2010.

(30) Jones, D. R. A taxonomy of global optimization methods based on response surfaces. *J. Global Optim.* **2001**, *21*, 345−383.
(31) Du, D.; Yang, H.; Ednie, A. R.; Bennett, E. S. Statistical metamodeling and sequential design of computer experiments to model Glyco-altered gating of sodium channels in cardiac myocytes. *IEEE J. Biomed Health Inform.* **2016**, *20* (5), 1439−1452.
(32) Stewart, B. T.; Venkat, A. N.; Rawlings, J. B.; Wright, S. J.; Pannocchia, G. Cooperative distributed model predictive control. *Syst. Control Lett.* **2010**, *59*, 460−469.
(33) Du, Y.; Duever, T.; Budman, H. Stochastic fault diagnosis using a generalized polynomial chaos model and maximum likelihood. *The 9th International Symposium on Advanced Control of Chemical Processes, Whistler, BC, Canada*; International Federation of Automatic Control: 2015; pp 1271−1276.
(34) Bathelt, A.; Ricker, N. L.; Jelali, M. Revision of the Tennessee Eastman process model. *The 9th IFAC Symposium on Advanced Control of Chemical Processes. Whistler, BC, Canada*; International Federation of Automatic Control: 2015.
(35) Downs, J. J.; Vogel, E. F. A plant-wide industrial process control problem. *Comput. Chem. Eng.* **1993**, *17* (3), 245−255.
(36) Ricker, R. L. Decentralized control of the Tennessee Eastman challenge process. *J. Process Control* **1996**, *6* (4), 205−221.
(37) Yin, S.; Ding, S.; Haghani, A.; Hao, H.; Zhang, P. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *J. Process Control* **2012**, *22*, 1567−1581.
(38) Chen, G.; McAvoy, T. Predictive online monitoring of continuous processes. *J. Process Control* **1998**, *8* (5−6), 409−420.
(39) Huang, W.; Zhao, D.; Sun, F.; Liu, H.; Chang, E. Scalable gaussian process regression using deep neural network. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina*; AAAI Press: 2015.
(40) Du, Y.; Du, D. Fault detection and diagnosis using empirical mode decomposition based principal component analysis. *Comput. Chem. Eng.* **2018**, *115*, 1−21.

DOI: 10.1021/acs.iecr.8b01110
Ind. Eng. Chem. Res. 2018, 57, 8962−8977