Computers and Chemical Engineering 106 (2017) 609–620

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Probabilistic uncertainty based simultaneous process design and control with iterative expected improvement model

Lester Lik Teck Chan, Junghui Chen*

Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taoyuan, Taiwan, 32023, Republic of China

A R T I C L E I N F O

Article history:
Received 13 March 2017
Received in revised form 29 July 2017
Accepted 30 July 2017
Available online 3 August 2017

Keywords:
Design and control
Expected improvement
Gaussian process
Probabilistic modeling

A B S T R A C T

The simultaneous design and control aims to achieve economic profits and smooth operation of the process even under uncertainties. However, the over-estimation of the uncertainties leads to conservative design decisions. Because of the disturbance inputs, the cost is not easily evaluated. Unlike the past work of design and control, the proposed probabilistic approach framework directly uses the Gaussian process (GP) model to represent the uncertainty in the input. The GP model that acts as the cost function model is trained by an iterative approach. The variability can be evaluated statistically by the GP model. In addition, the expected improvement optimization is employed to select the representative data, so no redundant data are used in the modeling. The expected improvement searches for the most probable operating condition for improvement based on the predictive distribution from the GP model. The applicability of the proposed method is tested on a mixing tank.

1. Introduction

In order to achieve maximum economic benefit with the minimum cost the plants are to be operated in a flexible manner. Traditionally, the process and control design are carried out sequentially. In the design phase the plant structure and operating conditions are calculated considering economic objectives at steady state with process constraints. The control system is then designed to achieve the desired dynamic behaviour. The problem of this approach is that the designed operational conditions and the steady-state based economic objective of a process flow sheet may not be optimal or may not result in good plant-wide dynamic performance when met with external disturbances and parametric or model uncertainties. The disturbances to the highly nonlinear chemical processes can destabilize the control system. As a result, the process design can affect the dynamic characteristics of the process and thus the control performance of the process.

The application of simultaneous design and control in chemical process aims to identify the design condition that generates maximum benefit with good dynamic performance of the control system even under the influence of disturbances and the existence of uncertainties. The interactions between process design and process control have been documented since 1940s (Ziegler and Nichols, 1943) and have motivated a number of works that have provided theoretical background about properties such as controllability, operability, stability and the selection of measurements and manipulated variables in process control design. Initial work (Morari and Stephanopoulos, 1980; Skogestad and Morari, 1987; Stephanopoulos et al., 1979) dealt mostly with controllability assessment and its incorporation into process synthesis as well as the selection of the control structure. On the other hand, the flexibility and the operability properties were considered (Dimitriadis and Pistikopoulos, 1995; Grossmann and Straub, 1996). These methods mainly adopted the steady-state economics with the respective criterion.

With the availability of improved computational resources allowing more powerful optimization methods and advanced control strategies, a wide variety of integrated design and control methodologies has been reported. The different concepts of the integration of design and control philosophy are evidenced in the reviews of the state of the art (Ricardez-Sandoval et al., 2009; Sharifzadeh, 2013; Yuan et al., 2012). Ricardez-Sandoval et al. (2009) adopted a classification of the methods based on the dynamic behaviour description and the description of the cost functions in the optimization framework. These classifications include (1) controllability index-based approach, (2) dynamic optimization-based approach and (3) robust model-based approach.

The controllability index approach evaluates the variability in the process variables using open-loop controllability metrics such

* Corresponding author.
E-mail address: jason@wavenet.cycu.edu.tw (J. Chen).

http://dx.doi.org/10.1016/j.compchemeng.2017.07.011

as the condition number (Luyben and Floudas, 1994) and the relative gain array (Alhammadi and Romagnoli, 2004). The process output variability estimated from these indicators was then used to assess the optimal process design. The dynamics of the system were mostly represented by linear systems or linearized models and therefore the method was inadequate for handling nonlinear processes. Dynamic optimization approach proposed a formal dynamic optimization problem that aims to estimate the critical profile in the disturbance that produces the worst-case scenario (Bansal et al., 2002; Mohideen et al., 1996). The resulting optimization formulations, which include the rigorous mechanistic process model, are computationally demanding even for simple process systems. Thus, the applicability of these techniques to optimally design chemical processes in the presence of disturbances and parameter uncertainty is challenging or even prohibitive. Model based approach enables the efficient computation of the worst-case scenario. A study is conducted to investigate the system's dynamics and the effect of the disturbances that may be affecting the process during its normal operation. The estimation of the worst-case variability in process variables due to disturbances is then used for computation of the worst-case scenario design and control (Chawankul et al., 2007; Trainor et al., 2013). The method alleviates the computational burden imposed by the dynamic optimization methods and can be used to perform the optimal design of large-scale dynamic systems (Muñoz et al., 2011; Ricardez-Sandoval et al., 2011). On the other hand, the variability of the systems can be described by a probabilistic approach as reported in recent work. Ricardez-Sandoval (2012) introduced a distribution analysis on the worst-case variability in the integrated design framework. The worst case variability is approximated by normal distribution functions in order to estimate the largest variability expected for the process variables at a user-defined probability limit. It requires Monte Carlo sampling at various conditions. The Monte Carlo sampling learns the complete region of the cost function. Intuitively, only certain regions will provide solution to the problem. It follows that learning only these representative regions is sufficient in the evaluation of the optimum process operation condition. The design and control research is an actively researched area. Vega et al. (2014) provided an extensive review and classification of the current state of design and control while Pistikopoulos and Diangelakis (2016) discussed the current application of design and control. Rasoulian and Ricardez-Sandoval (2016) applied the stochastic nonlinear model predictive control to a thin film deposition process. Rafiei-Shishavan et al. (2017) and Mehta and Ricardez-Sandoval (2016) proposed the use of approximating the function with power series expansions for simultaneous design and control. The low-order polynomials approximations are parametric models and the order of the models need to be determined for good performance.

Due to the disturbance inputs, it is not easy to evaluate the cost. This work proposed the representation of the variability of the process using a probabilistic framework. In the proposed method, the Gaussian process (GP) model that acts as the cost function model is trained by an iterative approach. The variability can be evaluated statistically using the GP model. GP model is a nonlinear modeling method and it also naturally accounts for the disturbance distribution description through the uncertain input. In addition, the expected improvement (EI) optimization is adapted. EI is a probabilistic method that provides a systematic way to explore new conditions based on the expected return. It facilitates the selection of representative data and thus no redundant data are used in the modeling. The use of EI in this work allows the pinpointing of the region that provides the greatest possibility for improvement statistically. Furthermore, the stability and feasibility constraint is considered simultaneously in the optimization to ensure that the designed condition is stable and feasible. The remainder of this paper is organized as follows. In the next section a problem definition of this study is presented and is followed by details of the proposed method. Next, the case studies demonstrate the features of proposed method and the article ends with some concluding remarks.

2. Problem definition

The objective of the simultaneous design and control is to minimize the cost such that controller parameters and the manipulated variables ensure process stability. This can be formulated as

$$\min CF = \min \left(CF_{nom}\left(\bar{\mathrm{u}}, \bar{\mathrm{d}}, \mathbf{q}\right) + VC\left(\bar{\mathrm{u}}, \mathbf{d}, \mathbf{q}\right)\right) \tag{1}$$

s.t.

$$\begin{aligned} &\Omega(\mathbf{u},\mathbf{d},\mathbf{y},\mathbf{q}) = 0\\ &\Psi(\mathbf{u},\mathbf{d},\mathbf{y},\mathbf{q}) \geq 0\\ &\mathbf{u} = \mathbf{c}\left(\boldsymbol{\lambda}, \mathbf{y}_c\right)\\ &\Gamma(\mathbf{u},\mathbf{d},\mathbf{y},\mathbf{q}) = 0\\ &\mathbf{u}^l \leq \mathbf{u} \leq \mathbf{u}^u\\ &\mathbf{y}^l \leq \mathbf{y} \leq \mathbf{y}^u \end{aligned} \tag{2}$$

The objective function to be minimized is represented by the cost function (CF) (Eq. (1)) that is the combination of nominal cost function,CF_{nom} and the variability cost (VC). CF_{nom} refers to the overall cost at nominal condition including the nominal values of the input disturbance variables. The term, $\bar{\mathrm{d}}$ refers to the input disturbance variables at the nominal value. The variability on the other hand is evaluated considering the distribution of the input disturbance variables represented by $\mathbf{d}$. In the modeling, the cost function is directly modeled as the output term, CF. This is because the cost in this case is considered as a function of output qualities as shown in Fig. 2. On the other hand, the output, $\mathbf{y}$, is a function of the inputs, $\mathbf{u}$. As a result, the cost function is written as such to represent the relation, $CF(\mathbf{u})$. $\bar{\mathrm{u}}$ denotes the nominal steady state values of the inputs and $\bar{\mathrm{y}}$ represents nominal outputs. $\bar{\mathrm{u}}$ is the designed condition and the variability is evaluated at design condition. $\bar{\mathrm{u}}$ is constrained on the range of the input/the actual available input of the physical process. In Eq. (2) Γ expresses additional requirements for the process performance such as stability constraints and the superscripts of l and u of the variables $\mathbf{u}$ and $\mathbf{y}$ indicate the process input and output variable bounds. Ω and Ψ represent the equality and inequality equations of the process and the mechanistic models are assumed to be available. The mechanistic models consist of manipulated variables $\mathbf{u}$, output variables $\mathbf{y}$, disturbances $\mathbf{d}$ and design variables $\mathbf{q}$. The output variables are partitioned as $\mathbf{y} = \begin{bmatrix} \mathbf{y}_{ol} & \mathbf{y}_c \end{bmatrix}$ where $\mathbf{y}_{ol}$ and $\mathbf{y}_c$ represents the open loop and controlled output variables respectively. The manipulated variables $\mathbf{u}$ ($\mathbf{u} = \begin{bmatrix} u_1 & \cdots & u_i & \cdots & u_I \end{bmatrix}$) are paired to the corresponding output variables ($\mathbf{y}_c = \begin{bmatrix} y_{c,1} & \cdots & y_{c,i} & \cdots & y_{c,I} \end{bmatrix}$). Fig. 1 shows I control loops where C_i, $i = 1, \ldots, I$ denotes the controllers, $y_{sp,i}$, $i = 1, \ldots, I$ represents the set points and e_i, $i = 1, \ldots, I$ the feedback errors. The manipulated variables are determined by the controllers represented as

$$\mathbf{u} = \mathbf{c}\left(\boldsymbol{\lambda}, \mathbf{y}_c\right) \tag{3}$$

where $\boldsymbol{\lambda}$ is the vector of controller parameters and the vectorized form of $\mathbf{c}$ denotes the controller equations of the particular $u_i - y_{c,i}$, $i = 1, \cdots, I$ pairs. The cost function is affected by disturbance, making its variability difficult to be evaluated. The disturbances considered are assumed to follow a specific probability density function for a given period of time. In this work, the disturbance is considered to be Gaussian distributed with known mean and variance. The GP is a probabilistic model and provides a way to for evaluating the variability as it naturally accounts for the disturbance distribution description through the uncertain input. Moreover, the direct