Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

A Non-optimality Detection Technique for Continuous Processes ⋆

Lingjian Ye * **Yi Cao** ** **Xiushui Ma** * **Zhihuan Song** ***

* *Ningbo Institute of Technology, Zhejiang University, 315100, Ningbo, China (e-mail: lingjian.ye@gmail.com, mxsh63@yahoo.com.cn).*
** *School of Engineering, Cranfield University, Cranfield, Bedford MK43 0AL, UK (e-mail: y.cao@cranfield.ac.uk)*
*** *Institute of Industrial Process Control, Zhejiang University, 310027, Hangzhou, China, (e-mail: zhsong@iipc.zju.edu.cn)*

Abstract: This paper develops a novel non-optimality detection technique for continuous processes based on the theory of process monitoring. Since the optimal statuses have some common features so as to satisfy optimality necessary conditions, the distributions of process variables are found to fall in some subspace that is determined from the disturbances. To detect those statuses that are not optimally operated, non-optimality is considered as a special kind of process fault. An equivalent interpretation is also provided to demonstrate the rationale of proposed analogy. For illustration, the principal component analysis (PCA) is particularly employed as a detection tool, whose effectiveness is verified through an exothermic reactor example.

Keywords: Continuous process, real-time optimization, optimality evaluation, process monitoring

1. INTRODUCTION

Chemical plants are often initially operated in the nominal points, which are optimally determined at the process design phase. However, under various uncertain disturbances (such as raw material quality variation, catalyst aging, etc.), the process will deviate from the true optimum and induce significant economic loss for plant operation. Therefore, it is a natural motivation to increase the economic profit by re-adjusting the operation such that the process is again optimally operated in new circumstances, which is the aim of real-time optimization (RTO).

A typical RTO solution, such as the so called two-step approach (Chachuat et al., 2009), involves two steps, namely disturbance estimation and re-optimization, which are alternatively repeated in sequence. In such a two-step RTO approach, the computation burden is intense because it has to solve large scale nonlinear programming problems in an acceptable time period, in order to quickly adjust the plant operation, *i.e.* update the set-points of control loops in the lower layer of a hierarchical control system in time. However, such computation efforts may unnecessarily be wasted or deficient in the following situations: (1) if the disturbances remain relatively unchanged in certain RTO cycles, the RTO procedures can be avoided because no actions are actually needed; (2) the influences of various disturbances are interacted, their effects may counteract with each other such that the overall operation may still be optimal in spite of the existence of disturbances. In such a case, the computation of RTO is also wasted; (3) In practical applications, the time interval of implementing RTO is often set to be hours (Skogestad, 2004). However, there is no information on how well the process is maintained in optimality between two RTO updates, which is crucial for daily operation.

Motivated by the reasons above, this paper develops a non-optimality detection technique based on the theory of process monitoring. By saying non-optimality, we refer to those system statuses where the overall plant operation is not optimal. Since the process operation is in a continuous sense and the measurements are noisy, in practice, non-optimality can be understood that the system deviates from the optimality over a certain level. In the proposed technique, the non-optimality is to be treated as a special process fault and identified through a multivariate statistical approach. To our best knowledge, such an idea has not been reported in literature. For illustration purpose, the principal component analysis (PCA) is employed as a tool for non-optimality detection. The new technique can be used for on-line monitoring of the system status, so as to inform the control system whether it is necessary to perform RTO, thus it overcomes the aforementioned defects of a traditional RTO strategy.

The rest of the paper is organized as follows: Section 2 describes some features of the optimal statues and then the principle of non-optimality detection is presented. In Section 3, the mathematical formulations of PCA is briefly reviewed. An example of exothermic reactor is studied in

⋆ The first, third and fourth authors gratefully acknowledge the National Natural Science Foundation of China (NSFC) (61304081), Zhejiang Provincial Natural Science Foundation of China (LQ13F030007), Postdoctoral Science Foundation of China (2013M541778), National Project 973 (2012CB720500) and Ningbo Innovation Team (2012B82002).

Section 4. The work is concluded with some discussions in Section 5.

2. PRINCIPLE OF NON-OPTIMALITY DETECTION

2.1 Descriptions of Optimal Statuses

Consider a static optimization problem for continuous processes with uncertainties

$$\min_{u} J(u, d) \tag{1}$$

s.t.

$$g(u, d) \leq 0 \tag{2}$$

with available measurements

$$y = f(u, d) \tag{3}$$

where J is a scalar cost function to be minimized, which is assumed to be convex in this paper; $u \in \mathbb{R}^{n_u}$, $d \in \mathbb{R}^{n_d}$ and $y \in \mathbb{R}^{n_y}$ are the manipulated, disturbance and measurement variables, respectively; $g : \mathbb{R}^{n_u \times n_d} \Rightarrow \mathbb{R}^{n_g}$ and $f : \mathbb{R}^{n_u \times n_d} \Rightarrow \mathbb{R}^{n_y}$ are the operational constraints and measurement equations, respectively.

Denote the solution of the optimization problem (1) as u^{opt}, it can be easily shown that the following necessary conditions of optimality (NCO) should be satisfied at u^{opt} (Ye et al., 2013)

$$\begin{aligned} &g_a = 0, \;\; g_a \in \mathbb{R}^{n_a} \\ &\nabla_r J = \frac{\partial J}{\partial u} V_2 = 0, \;\; \nabla_r J \in \mathbb{R}^{n_u - n_a} \end{aligned} \tag{4}$$

where g_a is a subset of g that are active and $\nabla_r J$ is the reduced gradients. V_2 is defined in the following way: let the singular value decomposition of the Jacobian matrix $\partial g_a/\partial u = USV^T$ and $V = [V_1 \; V_2]$, where V_2 are $n_u - n_a$ right singular vectors corresponding to $n_u - n_a$ zero singular values. Therefore, the reduced gradients $\nabla_r J$ can be intuitively interpreted as the gradients of objective function J with respective to the remained degrees of freedom of u after g_a has been satisfied.

Let $\Pi = \left[g_a^{\text{T}} \; \nabla_r J^{\text{T}}\right]^{\text{T}}$ be the overall set of NCO, which includes a total number of n_u equations to satisfy. Clearly, for optimal statuses, the degree of freedom is n_d, which is determined from the number of system disturbances. Once the disturbance values has been fixed, the optimal status for the system states can be determined via satisfying the NCO (4) and the process model.

Let $x \in \mathbb{R}^{n_x}$ denotes the system states. The relationship between optimal states x^{opt} and the disturbances d can then be represented in the following form

$$F(x^{\text{opt}}, d) = 0 \tag{5}$$

where F consists of the NCO and model equations, which establishes an implicit mapping from d to x^{opt}. For the sake of simplicity in representation and without loss of generosity, in the following derivations, we assume the input u is included in the system state x and all the states are measurable. Therefore, it is possible to approximately write (5) in a linear quasi-steady state as (Qin, 2003)

$$Ax^{\text{opt}} + Bd = 0 \tag{6}$$

$$y^{\text{opt}} = x^{\text{opt}} + n \tag{7}$$

where $A = \partial F/\partial x$ ($A \in \mathbb{R}^{n_d \times n_x}$), $B = \partial F/\partial d$ ($B \in \mathbb{R}^{n_d \times n_d}$), y^{opt} is the measurements in optimal statuses and n is the noises. Denote $B^{\perp}$ as the orthogonal complement of B such that $B^T B^{\perp} = 0$, left multiply the term $(B^{\perp})^T$ for (6) gives

$$(B^{\perp})^T A x^{\text{opt}} \equiv C x^{\text{opt}} = 0 \tag{8}$$

where $C = (B^{\perp})^T A$. From above equation we see that x^{opt} lies in the orthogonal spaces of C, i.e.

$$x^{\text{opt}} = (C^T)^{\perp} s \tag{9}$$

where s denotes the independent component. Inserting above equation into (7) we obtain

$$y^{\text{opt}} = (C^T)^{\perp} s + n \tag{10}$$

from which we know that with the projection matrix C, the observed measurements y^{opt} can be compactly described a lower n_d-dimensional space. In other words, the common features of optimal statuses can be equivalently interpreted by some independent latent variables.

Therefore, using the theory of process monitoring, the non-optimality as a special process fault, for which the relationship in (10) is not satisfied, can be effectively identified.

2.2 An Equivalent Interpretation

Before proceeding to introducing the PCA monitoring tool, in this subsection we provide an equivalent interpretation for the principle of non-optimality detection technique, where it is demonstrated that non-optimality actually amounts to a kind of biased sensor faults.

Consider a hypothesized control system CS1 where all the NCO components, denoted as Π, are selected as the controlled variables with 0 set-points. If we assume the controllers include integral actions, so that $\Pi = 0$ can be satisfied in steady state. Therefore, CS1 is always optimal under any disturbance scenario with all the NCO satisfied. Note CS1 is stated to be hypothesized because it may not be practically realizable due to the measurability of Π.

Now consider a practical control system CS2 where a measurable $\hat{\Pi}$, which could be measurements or their combinations, are the controlled variables, and without loss of generality, they are all assumed with 0 set-points. For comparisons, here we assume CS2 is not optimally operated, *i.e.* there is some error $\Delta = \hat{\Pi} - \Pi \neq 0$ between $\hat{\Pi}$ and Π, hence the integral controllers in CS2 will push $\hat{\Pi} = 0$ and result in NCO violations with a magnitude of Δ.

Both CS1 and CS2 can be uniformly described in another control system CS3, where the sensor for measuring the controlled variables contain possible faults. As illustrated in Fig. 1, CS1 is optimal where the controlled variables are Π and the sensors output Π for the controllers, whereas CS2 is not optimal where the controlled variables are $\hat{\Pi}$ and the sensors output $\hat{\Pi}$. CS3 provides a uniform framework where the sensors contain possible faults. In CS3, the controlled variables are Π, however, the sensors may output Π or $\hat{\Pi}$ for the controllers. When the sensors in CS3 are normal, CS3 is equivalent to CS1 which is optimal. When the sensors in CS3 are faulty with a bias of Δ, CS3 is equivalent to CS2 which is not optimal.

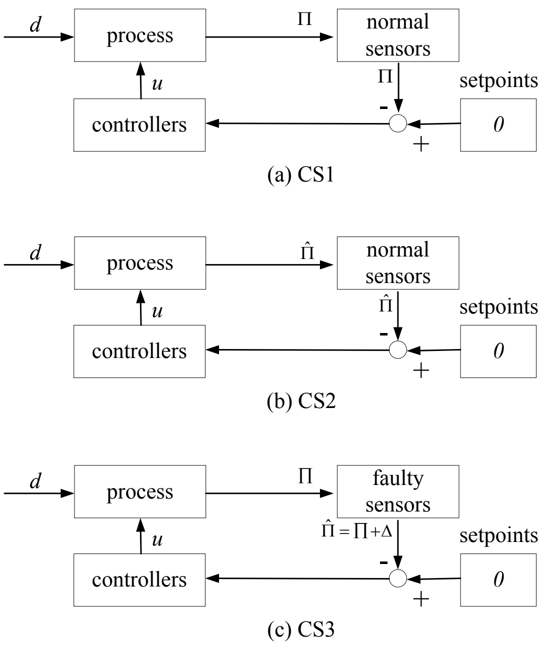

Fig. 1. Analogies between non-optimality detection and sensor fault detection: (a) CS1 with normal sensors (optimal) (b) CS2 with normal sensors (non-optimal) and (c) CS3 with possible faulty sensors (normal: optimal, faulty: non-optimal)

Through above analogies, the non-optimality detection can be interpreted as detecting a kind of biased sensor fault for CS3. Therefore, the procedures for non-optimality detection are as follows:

(1) Collect sampling data for normal conditions in CS3, which can be realized by off-line solving the optimization problem (1) with different disturbance scenarios using a process model and store the corresponding measurements to form a database;
(2) Based on multi-variable statistical approaches, use appropriate monitoring tools to build a monitoring model and calculate corresponding control limits;
(3) On-line monitoring for sensor faults in CS3, which is equivalent to non-optimality detection in this paper.

3. PRINCIPAL COMPONENT ANALYSIS

In this section, we briefly review a monitoring tool, namely the principal component analysis (PCA), which is employed for non-optimality detection in this paper.

Denote the data samples for PCA modelling in a matrix form $Y^{\text{opt}} \in \mathbb{R}^{M\times n_y}$, where M is the number of samples. The row vector $(y_i^{\text{opt}})^{\text{T}}$ in Y^{opt} denotes a group of measurements associated with an optimal state. To obtain Y^{opt}, we can solve the optimization problem (1)-(2) under various disturbance scenarios, and store the corresponding measurements using measurement model (3). Normally Y^{opt} is firstly scaled to have zero mean values and unit variance for each variable. The PCA form is represented as

$$Y^{\text{opt}} = TP^{\text{T}} + E = TP^{\text{T}} + \tilde{T}\tilde{P}^{\text{T}} \tag{11}$$

where $T \in \mathbb{R}^{M\times k}$ and $P \in \mathbb{R}^{n_y\times k}$ is the score and loading matrix of principle components, respectively. $E \in \mathbb{R}^{M\times n_y}$ is the residual matrix, $\tilde{T} \in \mathbb{R}^{M\times(n_y-k)}$ and $\tilde{P} \in \mathbb{R}^{n_y\times(n_y-k)}$ is the score and loading matrix of residual components, respectively. k is the number of principle components, whose value can be determined from cross validation or cumulative percent variance (CPV).

The various matrices in (11) can be obtained through the symmetric eigenvalue decomposition for the covariance matrix of Y^{opt}, $\Sigma = (Y^{\text{opt}})^{\text{T}}Y^{\text{opt}}/(M-1)$

$$\begin{aligned} \Sigma &= \left[P\ \tilde{P}\right]\Lambda\left[P\ \tilde{P}\right]^{\text{T}} \\ T &= Y^{\text{opt}}P \\ \tilde{T} &= Y^{\text{opt}}\tilde{P} \end{aligned} \tag{12}$$

where Λ is a diagonal matrix consists of all eigenvalues of Σ, i.e.

$$\Lambda = \text{diag}\{\lambda_1, \lambda_2, ..., \lambda_{n_y}\} \tag{13}$$

Through above steps, each row vector $(y_i^{\text{opt}})^{\text{T}}$ in Y^{opt} can be projected onto the principle and residual spaces

$$\begin{aligned} \hat{y}_i^{\text{opt}} &= PP^{\text{T}}y_i^{\text{opt}} \\ \tilde{y}_i^{\text{opt}} &= (I - PP^{\text{T}})y_i^{\text{opt}} \end{aligned} \tag{14}$$

where $\hat{y}_i^{\text{opt}}$ and $\tilde{y}_i^{\text{opt}}$ is the projected vector onto the principle and residual spaces, respectively, while satisfying

$$\begin{aligned} (\hat{y}_i^{\text{opt}})^{\text{T}}\tilde{y}_i^{\text{opt}} &= 0 \\ y_i^{\text{opt}} &= \hat{y}_i^{\text{opt}} + \tilde{y}_i^{\text{opt}} \end{aligned} \tag{15}$$

Using the PCA, the original data can be described in the reduced k dimensional uncorrected principle directions, with most of the information retained. For PCA, the T^2 and SPE statistics can be constructed in the principle and residual spaces, respectively. The T^2 statistic indicates the variation extent of data in principle space, which is defined as

$$T^2 = x^{\text{T}}P\Lambda^{-1}P^{\text{T}}x \tag{16}$$

If we assume the distributions of process data is Gaussian, then the T^2 statistic is demonstrated to obey an F distribution with k and $N-k$ as the degrees of freedom in normal conditions. Given a significance level α, the control limit of T^2 statistic can be calculated and the process is monitored as

$$T^2 \leq T_\alpha^2 = \frac{k(N-1)}{N-k}F_{k,N-k;\alpha} \tag{17}$$

Meanwhile, the SPE statistic indicates the distribution of data in the residual spaces, which is defined as the norm of projected residual vector

$$SPE = \|\tilde{x}\|^2 = \|(I - PP^{\text{T}})x\|^2 \tag{18}$$

Similarly, the corresponding SPE control limit δ_α^2 can also be computed as Jackson and Mudholkar (1979)

$$SPE \leq \delta_\alpha^2 = \theta_1\left(1 + \frac{c_\alpha\sqrt{2\theta_2h_0^2}}{\theta_1} + \frac{\theta_2h_0(h_0-1)}{\theta_1^2}\right)^{1/h_0} \tag{19}$$

where

$$\theta_i = \sum_{j=k+1}^{n_y} \lambda_j^i, \quad i = 1, 2, 3 \tag{20}$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \tag{21}$$

and c_α is the normal deviate corresponding to the upper $1-\alpha$ percentile.

4. ILLUSTRATIVE EXAMPLE: A REACTOR CASE

4.1 Process description

This is a continuous stirred-tank reactor (CSTR) where a reversible exothermic reaction occurs in the CSTR, as shown in Fig. 2 (Alstad, 2005; Kariwala, 2007; Ye et al., 2013). The raw material is A and the desired product is B. The inlet temperature, concentrations of A and B in the feed are denoted as T_i, C_{Ai} and C_{Bi} respectively, whilst, the outlet temperature, concentrations of unreacted A and product B in the outlet stream are denoted as T, C_A and C_B respectively.

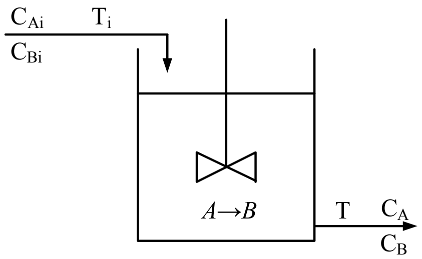

Fig. 2. Exothermic reactor process

The first principle models are composed of differential equations for mass and energy balances

$$\frac{dC_A}{dt} = \frac{1}{\tau}(C_{Ai} - C_A) - r \tag{22}$$

$$\frac{dC_B}{dt} = \frac{1}{\tau}(C_{Bi} - C_B) + r \tag{23}$$

$$\frac{dT}{dt} = \frac{1}{\tau}(T_i - T) + 5r \tag{24}$$

where $\tau = 60$ s is the residence time, and r is the rate of reaction which is determined from other process variables

$$r = 5000e^{-\frac{10000}{1987T}}C_A - 10^6 e^{-\frac{15000}{1987T}}C_B \tag{25}$$

The classifications for manipulated variable, available measured variables and disturbances are given as

$$u = [T_i]$$
$$y = [C_A \ C_B \ T \ T_i]^T$$
$$d = [C_{Ai} \ C_{Bi}]^T$$

For the measurements, the anticipated noises for measured variables are all of Gaussian type, the means are all assumed to be 0 and the standard deviation is considered as 0.01 for the concentration C_A and C_B, and 0.5 for the temperatures T and T_i. Disturbances are considered to be uniformly distributed with varying ranges in $0.5 \leq C_{Ai} \leq 1.5$ and $0 \leq C_{Bi} \leq 0.5$.

The objective of this CSTR operation is to maximize the economic profit, which can be represented as minimizing a cost function

Table 1. Process variables and nominal values for exothermic reactor

Variable	Physical description	Nominal value
C_A	Outlet A concentration	0.498 mol·L ^{1}
C_B	Outlet B concentration	0.502 mol·L ^{1}
T	Outlet steam temperature	426.803 K
T_i	Inlet steam temperature	424.292 K
C_{Ai}	Inlet A concentration	1.0 mol·L ^{1}
C_{Bi}	Inlet B concentration	0 mol·L ^{1}
J	Economic objective	-5149.3 ×10 ^{4}$

$$J = -20090C_B + (0.1657T_i)^2 \tag{26}$$

where the first term of J is the negative profit of product B and the second represents the cost of heating the input stream. The nominal values for various process variables are given in Table 1.

4.2 Non-optimality detection

The data set for building the detection model is generated by off-line solving 500 optimization problems minimizing J under different disturbances, which randomly vary within the predefined ranges. Therefore, the data set obtains distinct features belonging to optimal statuses of the CSTR. For the PCA, the number of principle components is set to be 2, which explained more than 99% of the total variation. The confidence level α is set to be 0.99, the control limits are calculated to be $T^2_{lim} = 9.3330$ and $SPE_{lim} = 0.0552$, as stated in the previous section.

Two different testing data sets are used for performance evaluation. Data set 1 contains 500 optimal samples which are generated the same as the modeling data set, however, with different random disturbances. Data set 2 contains 500 non-optimal samples whose economic costs are larger than the minimal cost by more than 1%. The detection performances are provided in Fig. 3, where the effectiveness of proposed non-optimality technique is verified. It can be found that most samples in Data set 1 fall in the class of optimal status whilst most samples in Data set 2 are confirmed to be not optimal.

For more precise evaluation, the false detection rate for Data set 1 and missing detection rate for Data set 2 are used as the indicators for performance evaluation, the results are summarized in Table 2. For Data set 1, both the $T^2(0.006)$ and $SPE(0.012)$ statistics give very low false rates. For Data set 2, The $T^2(0.11)$ gives a higher missing rate than the $SPE(0.022)$, which is in accordance with the fact that the SPE statistic is generally more suitable for detecting sensor fault than the T^2 statistic, where we have provided an equivalent interpretation of non-optimality detection to be a special kind of biased sensor fault. Besides, the results indicate that the PCA algorithm exhibits a better ability in monitoring the optimal samples than the non-optimal samples, which means that the algorithm is more apt to assign a sample to be optimal to avoid a false alarm. This may be caused by the nonlinearity of the CSTR process and a conservation solution is achieved for non-optimality detection.

4.3 On-line monitoring

For on-line implementing the proposed technique, two control systems are simulated to evaluate their economic per-

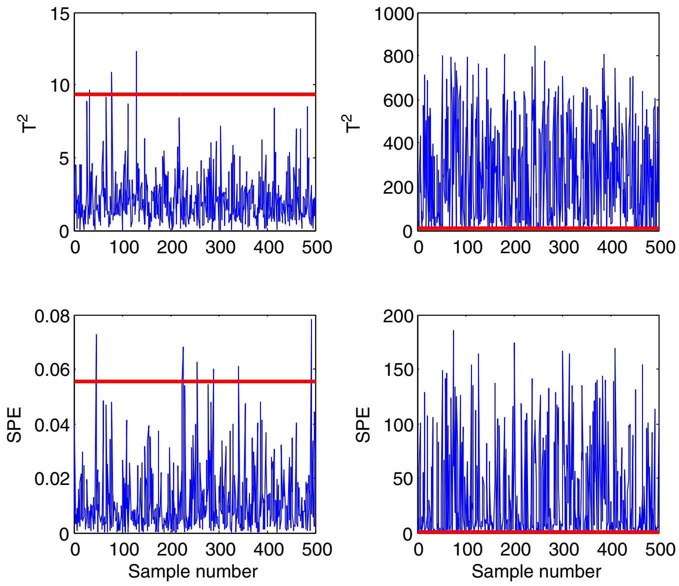

Fig. 3. Detection chart: (a) Data set 1; (b) Data set 2

Table 2. Summary of detection performance

Testing data set	Monitoring statistics	
	T^2	SPE
Data set 1: False rate	0.006	0.012
Data set 2: Missing rate	0.110	0.022

formances, which are configured with different controlled variables.

The controlled variable in control system I is a linear combination of measurements as

$$c_1 = -0.1657C_A + 0.1334C_B + 0.0023T$$

with a set-point of $c_{1,sp} = 0.9771$. The controlled variable in control system II is a quadratic combination of measurements as

$$\begin{aligned} c_2 = &-0.0349C_A - 0.102C_B - 0.00066T + 0.00116C_AC_B \\ &+ 1.3594 \times 10^{-5}C_AT + 0.00029C_BT + 0.00014C_A^2 \\ &- 0.00241C_B^2 + 7.8887 \times 10^{-7}T^2 \end{aligned}$$

with a set-point of $c_{2,sp} = -0.1413$. Both $c_{1,sp}$ and $c_{2,sp}$ are calculated by substituting the nominal values of measurements into their controlled variable expressions. The simulated disturbance scenario is shown in Fig. 4, where abrupt and large magnitude of step changes are considered.

The closed-loop behaviors of the two control systems are evaluated with applying the proposed detection technique. To better verify the correctness of detection results, the on-line trajectory of manipulated variable T_i and its true optimal value (Note this would be unknown for real applications) are also plotted for comparisons.

As shown in Fig. 5 for control system I, the detection technique suggests that phases 0-1000 s, 3000-4000 s and 4000-5000 s are optimal where both T^2 and SPE statistics are within alarm bounds (Strictly speaking, 0-1000 s and 4000-5000 s are suspicious because the SPE statistic lies at the boundary), which coincide with the fact that T_i runs near its optimal value. Phases 1000-2000 s and 2000-3000 s are detected as non-optimal with SPE statistics out

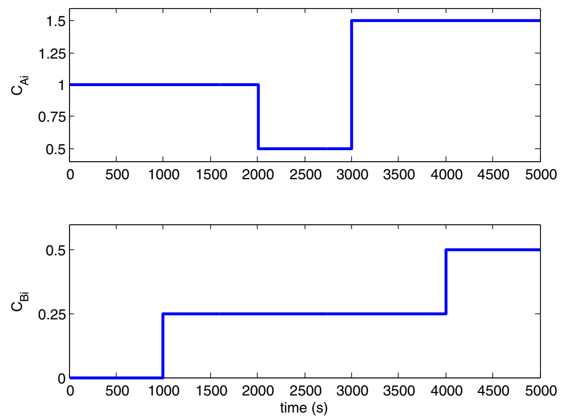

Fig. 4. Simulated disturbance scenario

of range, which are also in agreement with the fact that T_i deviates significantly from its optimal value.

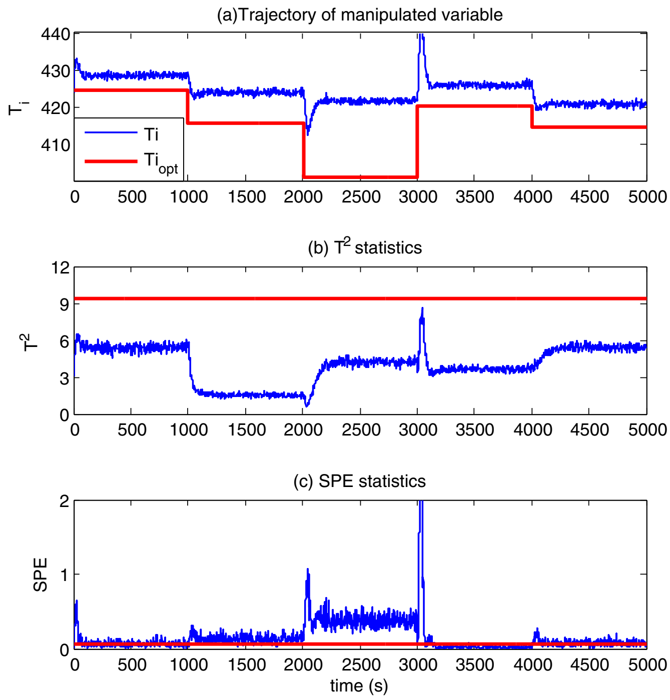

Fig. 5. On-line monitoring results for control system I: (a) trajectory of T_i; (b) T^2 statistic; (c) SPE statistic

Compared with control system I, control system II exhibits better self-optimizing performances under uncertainties, as indicated in Fig. 6. All phases are shown to be optimal with neither T^2 nor SPE statistic violation (except for suspicious phase 4000-5000 s whose SPE statistic lies at the boundary). Note that occasional sudden violations for T^2 and SPE statistics in the entire time range are due to process dynamics, which can be overcome by certain steady state conformation techniques to avoid a false alarm.

It is observed that in both control system I and II, the T^2 statistic has not successfully detected any non-optimal operations, this is because the two control systems are self-optimizing and the process is already near optimally operated. For near optimal operations, the T^2 statistic exhibits a degraded performance compared with SPE statistic. The failure of T^2 for detecting near optimality operation can also be explained by the analogy in Section 2.2, generally, SPE is more sensitive for detecting sensor faults than T^2 metric. On the other hand, it can be verified that for an arbitrary T_i that far away from the optimum, the T^2 statistic will trigger an alarm.

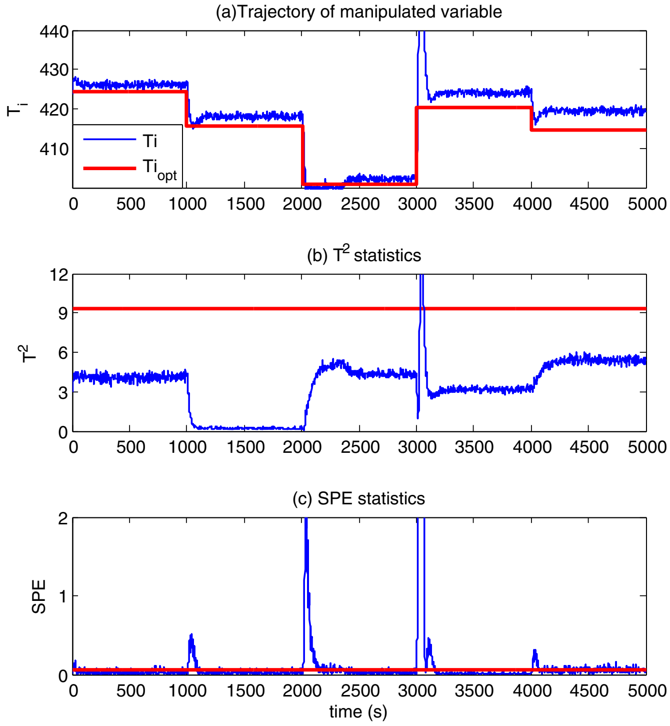

Fig. 6. On-line monitoring results for control system II: (a) trajectory of T_i; (b) T^2 statistic; (c) SPE statistic

Overall, in the above experiments, we conclude that (1)control system II is superior to control system I that it maintains the operation nearer to optimality under uncertainties; (2) the proposed methodology is effective for non-optimality detection.

5. CONCLUSION AND DISCUSSIONS

In this paper, a novel non-optimality detection technique is proposed to evaluate whether the plant is optimally operated. The key idea is to treat non-optimal statuses as abnormal situations, which can be identified through process monitoring means. A CSTR case is simulated to verify the effectiveness of proposed technique, where both off-line and on-line experiments are demonstrated. Concerning the contribution of this work, the following discussions are presented to point out possible directions for future work:

- The commonest monitoring tool PCA is used for demonstration. Since we have interpreted non-optimality as a special kind of process fault, one can naturally conclude that numerous tools developed in the field of process monitoring (Ge et al., 2013) will also applicable. For more complex processes, the PCA may not be sufficient hence more suitable monitoring tools should be sought.
- The currentness of proposed technique relies on an accurate model that describes the real plant. However, there exists model and plant mismatch for most of industrial practices. Under such circumstances, the detection technique may probably give a false alarm that unnecessarily interrupts normal operations. How to develop solutions to the cases with model and plant mismatch, and even those without any prior knowledge on process model, is an open problem.
- One potential usage of proposed technique is as a tool for evaluating the quality of plant operation and optimizing performance of the control system. A natural extension would be combining other techniques to improve the RTO performance. For examples, in the two-step RTO approach, the disturbance estimation and re-optimization may only perform when a non-optimality has been detected, thus reduce unnecessary and repeated computing resources. This promising usage for assisting in RTO can also be advocated from another perspective. Traditional two-step R-TO approach suffer from the “slowness” shortcoming of RTO performances, recently, the self-optimizing control (SOC) (Skogestad, 2000) has gained much attention due to its superior performance in RTO speed. A new type of hierarchical control structure is currently under progress (Ye et al., 2014), where the non-optimality detection runs at the optimization layer and the SOC runs at the regular control layer, both of which work complementarily to improve the RTO performance of control system.

REFERENCES

Alstad, V. (2005). *Studies on selection of controlled variables*. Ph.D. thesis, Norwegian University of Science and Technology.

Chachuat, B., Srinivasan, B., and Bonvin, D. (2009). Adaptation strategies for real-time optimization. *Comput. Chem. Eng.*, 33(10), 1557–1567.

Ge, Z., Song, Z., and Gao, F. (2013). Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, 52(10), 3543–3562.

Jackson, J.E. and Mudholkar, G.S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, 21(3), 341–349.

Kariwala, V. (2007). Optimal measurement combination for local self-optimizing control. *Ind. Eng. Chem. Res.*, 46(11), 3629–3634.

Qin, S.J. (2003). Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 17(8-9), 480–502.

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure. *J. Proc. Control*, 10(5), 487–507.

Skogestad, S. (2004). Control structure design for complete chemical plants. *Computers & Chemical Engineering*, 28(1), 219–234.

Ye, L., Cao, Y., Li, Y., and Song, Z. (2013). Approximating necessary conditions of optimality as controlled variables. *Ind. Eng. Chem. Res.*, 52(2), 798–808.

Ye, L., Cao, Y., Ma, X., and Song, Z. (2014). A novel hierarchical control structure with controlled variable adaptation. *Ind. Eng. Chem. Res.*, submitted.