

# Retrofit Self-Optimizing Control: A Step Forward Toward Real Implementation

Lingjian Ye, Yi Cao, *Senior Member, IEEE*, Xiaofeng Yuan, and Zhihuan Song

**Abstract**—After 15 year development, it is still hard to find any real application of the self-optimizing control (SOC) strategy, although it can achieve optimal or near optimal operation in industrial processes without repetitive real-time optimization. This is partially because of the misunderstanding that the SOC requires to completely reconfigure the entire control system, which is generally unacceptable for most process plants in operation, even though the current one may not be optimal. To alleviate this situation, this paper proposes a retrofit SOC methodology aiming to improve the optimality of operation without change of existing control systems. In the new retrofitted SOC systems, the controlled variables (CVs) selected are kept at constant by adjusting setpoints of existing control loops, which therefore constitutes a two-layer control architecture. CVs made from measurement combinations are determined to minimize the global average losses. A subset measurement selection problem for the global SOC is solved through a branch and bound algorithm. The standard testbed Tennessee Eastman process is studied with the proposed retrofit SOC methodology. The optimality of the new retrofit SOC architecture is validated by comparing two state of art control systems by Ricker and Larsson *et al.*, through steady-state analysis as well as dynamic simulations.

**Index Terms**—Control systems, optimization methods, process control.

## I. INTRODUCTION

OPTIMAL operation is the ultimate goal of all industrial plants. Nevertheless, achieving optimality is challenging due to inevitable uncertainties involved in any real plant operation. Real-time optimization (RTO) has been widely adopted by industry for decades to address this challenge, however, inade-

quately. In the solution, RTO is repeatedly carried out based on updated online detection and estimation of operational uncertainties. Then, the optimal solution is implemented as a setpoint update sent down to lower layer control loops in cascade. Optimization based on estimated uncertainties is an open-loop solution and its performance purely depends on the accuracy of uncertainty estimation, hence, is difficult to be robust. Meanwhile, the RTO has to wait for an operation reaching to a steady state for an optimality evaluation before applying any new change to the plant. This makes the RTO slow and not suitable to counteract fast changing disturbances [1]. To address the weakness of the RTO, the concept of self-optimizing control (SOC) was proposed in 2000 [2]. Instead of repetitively carrying out RTO online, SOC aims to select a set of controlled variables (CVs) offline such that when these CVs are controlled at certain constant setpoints online, the corresponding operation is automatically optimal or near optimal in spite of the existence of various uncertainties. Since then, the concept and theory of SOC have been advanced in many aspects, such as, instead of individual measurements [3], CV can be measurement combinations [4]; super-optimality of the average loss criterion over the worst-case criterion [5]; local linear-model-based CV selection [6] advanced to global SOC (gSOC) based on Monte Carlo simulation data over the entire operation space, which can be further enhanced by recent advancement in artificial intelligence [7], [8]; and efficient algorithms for subset measurements selection [9]–[12]. For more details of SOC development, readers are referred to [8].

In spite of aforementioned advancement, the SOC still has rarely been adopted by real industrial processes, while the traditional RTO has been widely accepted in industry. One of the main reasons is that the RTO update setpoints of existing control systems directly, hence, can work with any plant in operation, while the SOC was originally proposed to configure the entire control structure, hence, was misunderstood that it could only work with a completely new design of the entire control system of existing plants. Clearly, this is generally unacceptable for any real process in operation even though such an operation may not be optimal at all.

To alleviate this situation, this paper proposes a retrofit SOC methodology to show SOC can be implemented without reconfiguring existing control loops, but by updating the setpoint of these loops. In such a retrofitted system, two kinds of control requirements, namely the fundamental regulatory control requirements and operational optimality requirements are met by the existing control layer and a new retrofit SOC layer,

Manuscript received July 19, 2016; revised October 4, 2016; accepted December 11, 2016. Date of publication February 14, 2017; date of current version May 10, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61673349 and Grant 61304081 and in part by the Ningbo Natural Science Foundation under Grant 2015A610151.

L. Ye is with the Ningbo Institute of Technology, Zhejiang University, Ningbo 315100, China (e-mail: lingjian.ye@gmail.com)

Y. Cao is with the School of Water, Energy and Environment, Cranfield University, Cranfield, MK43 0AL, U.K. (e-mail: y.cao@cranfield.ac.uk)

X. Yuan is with the School of Information Science and Engineering, Central South University, Changsha 410006, China (e-mail: yuanxf@csu.edu.cn)

Z. Song is with the Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: zhsong@iipc.zju.edu.cn)

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIE.2017.2668991

respectively. A four step procedure is proposed to implement the retrofit SOC. This procedure ensures the previous control system is preserved, but some of process setpoints are automatically adjusted to achieve optimal or near optimal operation. Although the retrofit SOC is similar to RTO in terms of adjusting setpoints to achieve optimum, the retrofit SOC inherits feedback and dynamic control features from the traditional SOC, hence is more robust and faster than the RTO.

Preliminary works on the retrofit SOC method and measurement subset selection for gSOC problem for the Tennessee Eastman process (TEP) have been recently reported in two conferences [13], [14], respectively. This paper consolidates both with enhancements by, elaborating a systematic four-step procedure, which is easy to follow and potentially applicable to practical industrial processes, presenting a theoretical derivation to link the gSOC subset selection problem with the existing PB<sup>3</sup> algorithm, and updating the case study with improved results. Besides of this, the new contributions in this paper show a way how various mathematical tools developed up to date are integrated to solve a realistic large-scale problem such as the TEP.

The content of this paper is organized as follows. Section II elaborates the proposed retrofit SOC methodology, while the details of a particular CV selection method and a measurement subset screening algorithm are presented in Section III. Section IV reviews the TEP. Application of the retrofit SOC to the TEP is given in Section V, and finally, the work ends up with some conclusions drawn in Section VI.

## II. RETROFIT SOC METHODOLOGY

The retrofit SOC methodology proposed is concerned with such a process plant in operation with a workable control system, but requiring an improvement in its operational optimality. In order to work with the existing control system in operation, the retrofit SOC provides a cascade solution to adjust setpoints of certain existing control loops such that self-optimizing CVs constructed from measurement combinations are maintained at zero. This, in turn, leads to an improvement in operational optimality. It is assumed that the existing control system has already met the following requirements:

- 1) the closed-loop control system is stable;
- 2) core operational specifications, such as production grade and rate are properly controlled to follow given setpoints;
- 3) all operational constraints are handled appropriately under all scenarios to be considered.

The remaining degrees of freedom will be used to improve economic optimality. Assuming a mathematical/simulation model representative of the process plant is available, the proposed retrofit SOC methodology involves the following four main steps:

- 1) initial assessment of the economic performance of the existing control system;
- 2) selection of self-optimizing CVs;
- 3) configuration of retrofit SOC control structure;
- 4) performance validation.

### A. Step 1. Initial Economic Performance Assessment

Assume the ultimate goal is to achieve the optimal operation by solving the problem as follows:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J(\mathbf{u}, \mathbf{x}, \mathbf{d}) \\ \text{s.t.} \quad & \mathbf{f}(\mathbf{u}, \mathbf{x}, \mathbf{d}) = 0, \quad \mathbf{g}(\mathbf{u}, \mathbf{x}, \mathbf{d}) \leq 0 \end{aligned} \quad (1)$$

where  $J$  is the cost function to be minimized,  $\mathbf{u} \in \mathbb{R}^{n_u}$ ,  $\mathbf{d} \in \mathbb{R}^{n_d}$ , and  $\mathbf{x} \in \mathbb{R}^{n_x}$  are the manipulated variables (MVs), disturbances, and state variables, respectively.  $\mathbf{f}$  and  $\mathbf{g}$  are the steady-state process model and operational constraints, respectively.

In order to justify the necessity of adopting a retrofit control system, it is essential to assess the economic performance of the existing system. To facilitate such an assessment, first, the operational data record will be reviewed to identify different operational scenarios. There are two kind of operational scenarios, namely, scheduled operational condition changes, such as production rate or grade changes, which can be identified by changes in setpoint variables,  $\mathbf{c}_s$ , and undesired disturbances,  $\mathbf{d}$ , which may variate in operation, however, were not introduced by operators, such as physical and chemical properties of raw materials changes and downstream pressure fluctuations. Both the operating conditions and disturbances define a particular operating scenario, noted as  $\mathbf{\Pi} = \{\mathbf{c}_s, \mathbf{d}\}$  in the remainder of this study.

Under a particular operating scenario  $\mathbf{\Pi}_{(i)}$  identified from historical dataset, the economic cost function,  $J_i$  is evaluated using corresponding historical data. Meanwhile, the theoretically optimal cost,  $J_i^{\text{opt}}$  is obtained by solving problem (1). Then, the economic performance is quantified by a loss function  $L$  as follows:

$$L_i = J_i - J_i^{\text{opt}}. \quad (2)$$

The overall economic performance for  $N$  scenario is identified in the entire historical dataset is then calculated as

$$L_{\text{av}} = \sum_{i=1}^N \rho_i L_i \quad (3)$$

where  $\rho_i$  is the probability of scenario  $\mathbf{\Pi}_{(i)}$ , which can be estimated from the historical dataset.  $L_{\text{av}}$  is an indication of the potential benefit by adopting the retrofit SOC.

### B. Step 2. Selection of Self-Optimizing CVs

If the existing system has a significant loss,  $L_{\text{av}}$  to improve, then a set of self-optimizing CVs are to be selected. Although many CV selection approaches for SOC are available, traditional local SOC approaches may not be appropriate because the economic performance improvement achievable by a local SOC approach is much smaller than the gSOC approach, which is recommended and will be described in next section.

In a plant-wide process, there may exist numerous available measurements for CV construction. However, it is not necessary to adopt all available measurements for CVs. The tradeoff between the economic performance and CV complexity leads a measurement subset selection problem, which is a combinatorial challenge in nature and computationally intractable when

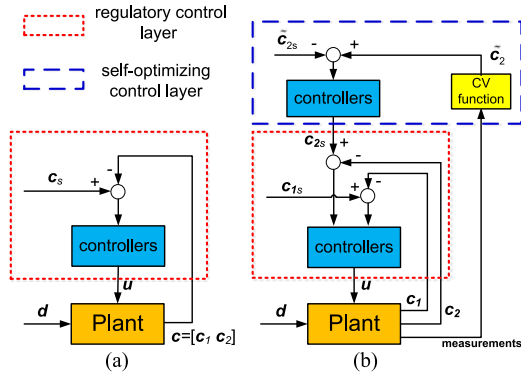


Fig. 1. Retrofit self-optimizing control architecture. (a) Original control system. (b) Retrofitted SOC system.

the number of available measurements is large. A tailored branch and bound approach is to be proposed to solve this problem in Section III.

### C. Step 3. Configuration of Retrofit SOC Structure

In the existing control system, the primary CVs are first divided into two classes: 1)  $c_1$ , including CVs that have no steady-state effects (e.g., liquid levels) and active constraints (equality constraints); and 2)  $c_2$ : other CVs. On top of the existing control loops associated with  $c_2$ , a new upper control layer (SOC layer) will be configured such that the values of  $c_2$  will be adjusted to maintain new CVs  $\tilde{c}_2$ , selected in Step 2, at zero. The retrofit control architecture, together with the existing control system, is illustrated in Fig. 1.

The proposed control system has the following features:

- 1) regulatory control tasks are naturally addressed by the preserved existing control system;
- 2) by controlling the self-optimizing CVs on the upper layer, the economic performance is improved.

### D. Step 4. Performance Validation

Finally, the economic performance of the retrofitted SOC system can be validated through steady-state evaluation and dynamic simulation for operational scenarios identified. The steady-state evaluation is mainly to verify the economic improvement of the retrofitted system over the existing system based on the nonlinear process model. The verified results should be in high agreement with estimations made in the CV selection step. For dynamic simulation, it is devoted to verify the online controllability of selected CVs to ensure the dynamic performance is satisfactory. Results obtained from dynamic simulation could provide a guidance for control tuning in the new SOC layer.

## III. PLANT-WIDE CV SELECTION FOR SOC

### A. Global SOC Method

With appropriate pretreatments, we should be able to formulate an unconstrained static optimization problem as

$$\min_{\mathbf{u}} J(\mathbf{u}, \mathbf{\Pi}) \quad (4)$$

with available measurements

$$\mathbf{y}_m = \mathbf{y} + \mathbf{n} = \mathbf{f}_y(\mathbf{u}, \mathbf{\Pi}) + \mathbf{n} \quad (5)$$

where  $\mathbf{y}_m$ ,  $\mathbf{y}$ , and  $\mathbf{n}$  are the measured, true, and noise values of measurements, respectively,  $\mathbf{f}_y$  is the measurement model. The objective is to select  $\mathbf{c} = \mathbf{H}\mathbf{y}$  as CVs such that the average economic loss [defined in (3)] is minimized over the given set of operating scenarios,  $\{\mathbf{\Pi}_{(i)}\}$  when CVs are maintained at constant setpoints,  $\mathbf{c}_s$ . Note, in this section, we use the symbol  $\mathbf{c}$  instead of  $\tilde{\mathbf{c}}_2$  for a shorthand of notation. In gSOC, the CVs and their setpoints are considered together for optimization. This is done by inducing the general CVs  $\hat{\mathbf{c}} = \mathbf{c} - \mathbf{c}_s$ , which should be controlled at zero. The optimal value of  $\mathbf{c}_s$  can then be obtained together with  $\mathbf{c}$  by expanding measurements  $\mathbf{y}$  with a constant 1, e.g., introducing  $\hat{\mathbf{y}} = [1 \ \mathbf{y}^T]^T$ . For simplifying notation, in the remaining part of this paper,  $\mathbf{c}$  and  $\mathbf{y}$  will be used to represent  $\hat{\mathbf{c}}$  and  $\hat{\mathbf{y}}$ , respectively.

The gSOC method [8] aims to minimize the average loss over all operating scenarios, hence, leads to losses much smaller than what achievable with existing local approaches, e.g., [5], [6], [15], and [16]. In the gSOC, first, the economic loss  $L$  is approximated with a quadratic function in terms of  $\mathbf{c}$  [4]

$$L = \frac{1}{2} \mathbf{e}_c^T \mathbf{J}_{cc} \mathbf{e}_c \quad (6)$$

where  $\mathbf{e}_c \triangleq \mathbf{c} - \mathbf{c}^{\text{opt}}$ , the deviation of  $\mathbf{c}$  from optimum,  $\mathbf{c}^{\text{opt}}$ ,  $\mathbf{J}_{cc}$  is the Hessian of  $J$  with respect to  $\mathbf{c}$ , obtained as  $\mathbf{J}_{cc} = (\mathbf{H}\mathbf{G}_y)^{-T} \mathbf{J}_{uu} (\mathbf{H}\mathbf{G}_y)^{-1}$  with  $\mathbf{G}_y$ , the sensitivity matrix of  $\mathbf{y}$  and  $\mathbf{J}_{uu}$ , the Hessian of  $J$ , both with respect to  $\mathbf{u}$ . Since  $\mathbf{c}^{\text{opt}} = \mathbf{H}\mathbf{y}^{\text{opt}}$  and through control,  $\mathbf{c}_m = \mathbf{H}\mathbf{y}_m = 0$ , we have  $\mathbf{c} = \mathbf{c}_m - \mathbf{H}\mathbf{n} = -\mathbf{H}\mathbf{n}$  and  $\mathbf{e}_c = -\mathbf{H}(\mathbf{y}^{\text{opt}} + \mathbf{n})$ . Hence [8],

$$L_{\text{av}} = E(L) = E(L^{\text{II}}) + E(L^n) \quad (7)$$

$$L^{\text{II}} = \frac{1}{2} (\mathbf{y}^{\text{opt}})^T \mathbf{H}^T \mathbf{J}_{cc} \mathbf{H} \mathbf{y}^{\text{opt}}, \quad L^n = \frac{1}{2} \text{tr}(\mathbf{n}\mathbf{n}^T \mathbf{H}^T \mathbf{J}_{cc} \mathbf{H}) \quad (8)$$

where  $\text{tr}(\cdot)$  represents the trace of a matrix.  $L^{\text{II}}$  and  $L^n$  are losses caused by uncertainties and noises, respectively.

In general, the optimization problem to find  $\mathbf{H}$  that minimizes the average loss (7) is nonconvex. In [8], a short-cut algorithm was proposed to simplify the optimization problem by assuming the Hessian term  $\mathbf{J}_{cc}$  is constant over different operating scenarios such that an analytical solution can be readily derived, as presented later. Moreover in order to ensure the uniqueness of  $\mathbf{H}$ ,  $\mathbf{J}_{cc} = \mathbf{I}$  is enforced by incorporating a constraint of  $\mathbf{H}\mathbf{G}_{y,r} = \mathbf{J}_{uu,r}^{1/2}$ , where the subscript  $(\cdot)_r$  denotes a chosen reference operating point. With this constraint, the loss function is simplified as

$$L^{\text{II}} = \frac{1}{2} (\mathbf{y}^{\text{opt}})^T \mathbf{H}^T \mathbf{H} \mathbf{y}^{\text{opt}}, \quad L^n = \frac{1}{2} \text{tr}(\mathbf{H}\mathbf{n}\mathbf{n}^T \mathbf{H}^T). \quad (9)$$

Furthermore, by introducing  $\mathbf{W}^2 \triangleq E(\mathbf{n}\mathbf{n}^T)$ ,  $\mathbf{Y}$  and  $\tilde{\mathbf{Y}}$  as

$$\mathbf{Y}^T = [(\mathbf{y}_{(1)}^{\text{opt}}) \cdots (\mathbf{y}_{(N)}^{\text{opt}})], \quad \tilde{\mathbf{Y}} = \begin{bmatrix} \frac{1}{\sqrt{N}} \mathbf{Y} \\ \mathbf{W} \end{bmatrix} \quad (10)$$

the gSOC CV selection problem is finally formulated as

$$\min_{\mathbf{H}} L_{\text{av}} = \frac{1}{2} \|\tilde{\mathbf{Y}}\mathbf{H}^T\|_{\text{F}}^2, \quad \text{s.t. } \mathbf{H}\mathbf{G}_{y,r} = \mathbf{J}_{uu,r}^{1/2} \quad (11)$$

which is convex and an analytical solution follows as [8]

$$\mathbf{H}^T = (\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}})^{-1} \mathbf{G}_{y,r} (\mathbf{G}_{y,r}^T (\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}})^{-1} \mathbf{G}_{y,r})^{-1} \mathbf{J}_{uu,r}^{1/2}. \quad (12)$$

In summation, the CVs are selected by the following steps.

1) Sampling the operational space using Monte Carlo simulation, a dataset corresponding to disturbance scenarios  $\{\mathbf{\Pi}_{(i)}\}$ ,  $i = 1, \dots, N$ , is generated.

2) Choose a reference point, the sensitivity matrix,  $\mathbf{G}_{y,r}$  and Hessian,  $\mathbf{J}_{uu,r}$  are evaluated.

3) For each scenario  $\mathbf{\Pi}_{(i)}$ , optimization problem (4) is solved and corresponding values,  $\mathbf{y}_{(i)}^{\text{opt}}$  are stored. Then, construct  $\mathbf{Y}$  and  $\tilde{\mathbf{Y}}$ .

4) Obtain combination matrix  $\mathbf{H}$  in (12).

### B. Algorithm for Measurement Subset Selection

The aforementioned gSOC method is based on a given measurement set. However, a real plant may contain substantial measurements, which are not necessarily be contained in the CV function. The subset selection problem is combinational in nature and may be computationally intractable to search exhaustively when the number of measurements is large. Kariwala and Cao [9]–[11] have developed a series of bidirectional branch and bound algorithms for fast subset measurement selection based on the local SOC. For example, the widely used PB<sup>3</sup> algorithm [11] was based on the local average loss criterion [5]. In these algorithms, candidate subsets are divided into branches and evaluated against upwards and downwards pruning criteria. Branches, which satisfy either upwards or downwards pruning criteria will be fixed or removed from candidate list, respectively. In this way, most nonoptimal candidate subsets will be eliminated without further evaluations so that the optimal subset can be efficiently identified.

In order to use the PB<sup>3</sup> algorithm to solve the gSOC subset selection problem, the subset selection criterion is derived as follows. For a given measurement subset  $\mathbf{y}$ , the minimal global average loss can be calculated by inserting the optimal combination matrix derived in (12) into (11), as

$$\begin{aligned} L_{\min,y} &= \frac{1}{2} \|\tilde{\mathbf{Y}}\mathbf{H}^T\|_{\text{F}}^2 \\ &= \frac{1}{2} \|\tilde{\mathbf{Y}}(\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}})^{-1} \mathbf{G}_{y,r} (\mathbf{G}_{y,r}^T (\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}})^{-1} \mathbf{G}_{y,r})^{-1} \mathbf{J}_{uu,r}^{1/2}\|_{\text{F}}^2 \\ &= \frac{1}{2} \text{tr} \left( \mathbf{J}_{uu,r}^{1/2} (\mathbf{G}_{y,r}^T (\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}})^{-1} \mathbf{G}_{y,r})^{-1} \mathbf{J}_{uu,r}^{1/2} \right) \\ &= \frac{1}{2} \sum_{i=1}^{n_u} \lambda_i^{-1} \left( \tilde{\mathbf{G}}^T (\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}})^{-1} \tilde{\mathbf{G}} \right) \end{aligned} \quad (13)$$

where  $\lambda_i$  is  $i$ th largest eigenvalues of a matrix,  $\tilde{\mathbf{G}} = \mathbf{G}_{y,r} \mathbf{J}_{uu,r}^{-1/2}$ . Equation (13) indicates that the achievable minimum average loss,  $L_{\min,y}$  is independent from  $\mathbf{H}$ , but depends on measurement set  $\mathbf{y}$  adopted for combination. For a subset,  $X \subset \mathbf{y}$ , denote  $\mathbf{G}_X$  a submatrix of  $\tilde{\mathbf{G}}$  with rows selected corresponding to  $X$

and  $\mathbf{Y}_X$  a submatrix of  $\tilde{\mathbf{Y}}$  with columns selected corresponding to  $X$ . Then, the subset of measurements selection problem can be represented as follows:

$$\min_{X \subset \mathbf{y}} \sum_{i=1}^{n_u} \lambda_i^{-1} \left( \mathbf{G}_X^T (\mathbf{Y}_X^T \mathbf{Y}_X)^{-1} \mathbf{G}_X \right). \quad (14)$$

Based on the aforementioned analysis and preparations, the core part of the existing PB<sup>3</sup> algorithm [11] can be directly used to solve the gSOC subset selection problem although some modifications to the original code are necessary to prepare appropriate matrices and to ensure constant 1 is always included. For more details in general principles and computational complexities of bidirectional branch and bound and the PB<sup>3</sup> algorithm, readers are referred to [9]–[11].

## IV. TEP CASE STUDY

### A. Brief Process Description

The retrofit SOC technique is applied to a well-known benchmark example, the TEP [17]. The plant-wide TEP consists of four reactions with  $A, C, D, E$  are the reactants,  $G$  and  $H$  are the desired products and  $F$  is the byproduct. Besides, there exists an inert component  $B$  in the material circle, which is contained in the feed and removed through the purge to maintain inventory balance. The process includes five major operating units: a gas–liquid two phase reactor, a product condenser, a vapor–liquid separator, a recycle compressor, and a product stripper. The flowchart of the TE plant can be found in [17].

**1) Process Variables:** The TEP has 12 MVs and 41 measurements as described in [17]. All MVs have been scaled within 0–100 limits in percentage, which can be viewed as valve positions accordingly (11 valves and 1 agitation rate). For the measurements, they were defined with different sampling frequency/dead time together with random noises to keep consistency with industrial practice, see [17] for more details.

**2) Operating Constraints:** Five process operating constraints were present for safe operation and equipment protections including reactor pressure and temperature, and three liquid levels (reactor, product separator, and stripper base). Among them, the reactor pressure is most likely to be violated because of the open-loop instability. In particular, this pressure should be constrained within 2895 kPa and triggers a shutdown action when it exceeds a limit of 3000 kPa. The reactor temperature is constrained within 150 °C. The three liquid levels are limited within 50–100% for the reactor level and 30–100% for others. Besides of these five process constraints, all MVs are physically restricted, which can be however directly handled in operation.

**3) Operating Conditions:** In general, the process involves six possible operating modes, which are defined by the desired mixup of product (G/H mass ratio) and the production rate. Among them, the first one (Mode 1) was considered as the base case and has been most intensively studied. In this model, the G/H mass ratio is 50/50 with a fixed production rate, 7038 kg · h<sup>-1</sup> for both G and H. In this study, we also consider Mode 1 as an illustration for the proposed methodology.

Besides, Downs and Vogel [17] have introduced several test scenarios with step setpoint changes, which were proposed as testing benchmarks for different control strategies. These step change scenarios are as follows:

- 1) production rate change by  $-15\%$ ;
- 2) product mixup change from 50 G/50 H to 40 G/60 H;
- 3) reactor pressure change from 2705 to 2645 kPa;
- 4) B in purge by  $+2\%$ .

Furthermore, the process suffers from 20 possible disturbances [IDV(1–20)], 15 of which are caused by identifiable reasons, while the sources of others are unknown. IDV(1–15) are further distinguished from *step*, *random variation*, *slow drift*, and *valve sticking* types, see [17] for details.

**4) Operational Objective:** The operational objective of TEP is to minimize the operating cost,  $J$  consisting of four parts, *i.e.*, costs of purge, product, compressor, and steam determined by the corresponding unit costs,  $r_i$  and flow rates,  $Q_i$ , as

$$J = \sum_{i=1}^4 r_i Q_i \quad (15)$$

which gave a total steady-state cost of about  $170 \text{ \$} \cdot \text{h}^{-1}$  at the base case provided by Downs and Vogel. This operating cost was encouraged to be further reduced using whatever appropriate optimization techniques.

### B. Two State of Art Control Structures

Although substantial control solutions have been developed for the TEP in the literature, in this subsection, we focus on two state of art control structures proposed in [3] and [18], which are both very efficient for controlling the TEP. For the sake of convenience, they will be denoted as “CS\_Ricker” and “CS\_Skoge,” respectively, throughout this paper (named after their corresponding authors). The summaries given later highlight the choice of primary CV selection, which is of particular interest in this study. According to the control policy of CS\_Ricker [18], the following nine process variables were selected as controlled variables.

- 1) Separator level and stripper level for stable operation.
- 2) Production rate and mole %G in product.
- 3) Five active constraints were: reactor pressure (maximum and level (minimum), compressor recycle valve (closed), stripper steam valve (closed), and agitator speed (maximum), with physical interpretations why these constraints are active at the optimum provided in [19].

For the remaining three unconstrained degrees of freedom (DOF), Ricker [18] chose to control the reactor temperature, %A and %C in the feed heuristically. The setpoints of these 12 CVs were optimized in [19], with an operating cost of about  $114 \text{ \$/h}$ . For dynamic control, all feedback controllers are proportional-integral (PI) type and parameters were tuned reasonably. As has been intensively demonstrated, this designed control system, CS\_Ricker was indeed efficient and nicely achieved various control targets proposed by Downs and Vogel [17].

Although the operating point in CS\_Ricker was optimized in [19], the problem with operational uncertainties was not addressed. Obviously, the plant operation may lose optimality

when disturbances occur or under other unconsidered operating conditions. In this context, Larsson *et al.* [3] reconsidered the TE challenge problem by applying the SOC methodology [2] to improve the operational economic performance to cover some more scenarios in addition to the nominal condition. Following the methodology in [2], a systematic procedure for the plant-wide control structure design was carried out thoroughly. Apart from the same nine primary CVs as adopted in CS\_Ricker, a series of mathematic evaluations combined with ad hoc logics were carried out to finally identify three most promising self-optimizing CVs [3]: the reactor temperature, recycle flowrate, and %C in the purge. CS\_Skoge gave an overall economic loss of  $3.8 \text{ \$/h}$  for all considered scenarios, as compared to a loss of  $9.8 \text{ \$/h}$  in CS\_Ricker.

## V. APPLICATION OF RETROFIT SOC TO THE TEP

### A. Step 1. Evaluation of Existing Control Structures

To follow the retrofit procedure, the economic performance of CS\_Ricker and CS\_Skoge are first evaluated.

The first step is to identify operating scenarios for study. For disturbances, although Downs and Vogel have defined 20 disturbance scenarios, CS\_Skoge [3] only considered IDV(1) and IDV(2) for the purpose of self-optimizing control, this is because other disturbances are either with no steady-state effects, or too severe to be handled consistently (IDV(6): loss of feed A, which was recommended to be handled using override control). Besides of disturbances, they also included the situations when the setpoint of production rate (throughput) is changed by  $\pm 15\%$ . In this paper, we will include two additional situations as were posed by Downs and Vogel [17]: 1) when the product mixup changes from 50 G/50 H to 40 G/60 H; and 2) a step change of reactor pressure setpoint to be 2645 kPa. Hence, including the normal operating condition, we have seven operating scenarios under considerations.

For each of seven scenarios considered previously, optimization to minimize the cost function in (15) is performed using the SQP algorithm. Since there are only three unconstrained DOF for optimization, the optimization uses the setpoints of reactor temperature, %A and %C in the feed as the decision variables while keeping other associated variables actively controlled. A summary of optimal steady-state values for some variables are given in Table I. For all seven operating scenarios, the economic losses for both CS\_Ricker and CS\_Skoge are calculated at steady state and shown in the last two rows in Table I. The closed-loop results are in high agreement with ref [3] investigated, namely, CS\_Ricker gives some big losses particularly when throughput  $+15\%$  ( $6.74 \text{ \$/h}$ ). In contrast, CS\_Skoge outperforms CS\_Ricker mainly in the case of throughput changes and reduces the overall loss. However, in the new operating condition when the reactor pressure setpoint is 2645 kPa, CS\_Skoge gives an unexpected poorer economic loss of  $5.29 \text{ \$/h}$  than CS\_Ricker ( $0.41 \text{ \$/h}$ ). As an overall assessment, the total loss for CS\_Ricker and CS\_Skoge is  $13.01$  and  $9.38 \text{ \$/h}$ , respectively. Therefore, if an overall  $6 \text{ \$/h}$  is considered as an acceptable loss [3], both of them need to be improved. In the following study we will, without loss of generality, consider a retrofit configuration

**TABLE I**  
OPTIMAL STEADY-STATE VALUES FOR DIFFERENT OPERATING SCENARIOS

Variables	Normal	IDV(1)	IDV(2)	Thrgpt. +15%	Thrgpt. -15%	40 G/ 60H	Ret. Press. 2645 kPa
MV1, %	62.97	62.88	63.06	72.74	53.28	50.60	62.95
MV2, %	53.13	53.07	53.94	61.51	44.85	63.85	53.47
MV3, %	26.26	76.00	29.59	31.48	21.51	27.12	27.28
MV4, %	60.57	57.15	61.10	69.91	51.29	59.55	60.59
MV5, %	0	0	0	0	0	0	0
MV6, %	25.82	24.68	46.92	31.39	20.57	29.62	32.70
MV7, %	37.26	37.10	37.14	43.31	31.33	37.04	37.44
MV8, %	46.44	46.39	46.27	53.66	39.30	45.65	46.36
MV9, %	0	0	0	0	0	0	0
MV10, %	35.94	35.89	35.35	42.06	30.09	33.18	35.77
MV11, %	12.39	12.20	12.16	16.32	9.22	12.58	12.59
MV12, %	100	100	100	100	100	100	100
XMEAS9, °C	122.9	123.0	124.1	124.3	121.5	123.2	123.6
XMEAS23, %	32.15	32.28	30.63	33.35	30.90	32.99	32.01
XMEAS25, %	18.66	18.78	17.82	19.66	17.49	18.92	18.82
XMEAS31, %	13.01	13.19	11.62	13.57	12.34	13.52	12.87
XMEAS5, kscmh	32.37	32.44	31.43	31.58	33.19	32.25	30.17
Optimal cost, \$/h	114.01	111.26	168.99	140.50	90.98	129.08	134.93
loss (CS_Ricker), \$/h	0.0	0.06	2.52	6.74	2.45	0.83	0.41
loss (CS_Skoge), \$/h	0.0	0.02	1.77	1.49	0.56	0.25	5.29

upon CS\_Ricker by using it as a base control system in the new control architecture.

**B. Step 2. Selection of Self-Optimizing CVs**

Three measurement combinations are to be selected as CVs for three unconstrained DOF. The gSOC algorithm requires the following information.

- 1) Optimized measurement values for all operating conditions of interest, which have been calculated when assessing CS\_Ricker and CS\_Skoge.
- 2) The sensitivity matrix,  $G_{y,r}$  and Hessian,  $J_{uu,r}$  at a reference point. The normal operating condition is selected as the reference point. The setpoints of reactor temperature, %A and %C in the feed are taken as the MVs, hence,  $G_{y,r}$  and  $J_{uu,r}$  are evaluated in terms of them using the finite-difference method.
- 3) Magnitudes of measurement errors were estimated using the historical data.

**1) Measurement Subset Selection:** Combining all XMEAS and XMV variables gives us a set of  $41 + 12 = 53$  candidates. However, there have been nine closed loops with constants setpoints, which cannot be used as measurement candidates including: XMEAS(7, 8, 12, 15, 17, 19, 40) and XMV(5, 9, 12). Note XMEAS(19) (steam flow) and XMV(9) (steam valve) actually the same because steam valve is closed, hence both of

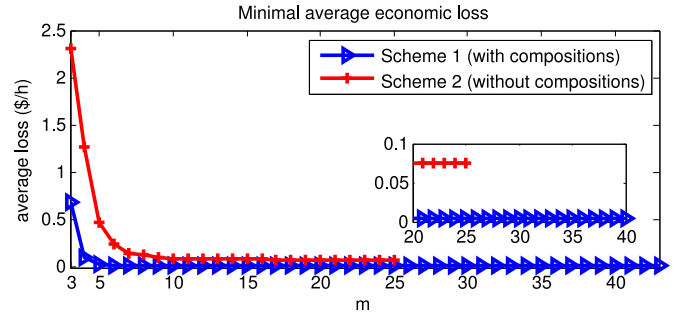


Fig. 2. Minimal average loss against subset size using PB<sup>3</sup> algorithm.

them are eliminated. It is unnecessary to use all remaining 43 measurements for CVs. Hence, we consider a small size of subset to balance the economic performance and CV complexity. Next, we further differ two options: with and without composition variables (Option 1 and 2) as candidate measurements. This is motivated by the fact that compositions are more difficult to measure. Furthermore, since analyzers may lose function in daily operation, it was highly encouraged to avoid relying on compositions as much as possible [17].

Option 1 contains 43 candidate measurements, while Option 2, by eliminating XMEAS(23–41), contains 25, which is still substantial in number and intractable to search exhaustively. Instead of using a heuristic approach together with quantitative analysis to eliminate unpromising candidates as in [3], which would be time consuming and very much rely on the designer’s expert knowledge, in this paper, the tailored PB<sup>3</sup> algorithm is applied based on the global average loss criterion.

All possible sizes of measurement subset ( $m$ ) are considered, with promising subsets automatically identified quickly by using the PB<sup>3</sup> algorithm (within 40 s for a PC with Intel i5 CPU, 8-GB RAM). The trends of minimal average losses against  $m$  for both Option 1 and 2 are illustrated in Fig. 2. Evidently, the minimal average losses show a same decreasing trend as  $m$  increases. However, the achievable minimal loss for Option 1 is much smaller than Option 2 for the same  $m$ , which is not actually surprising because Option 1 has more general measurements to choose. Overall, Fig. 2 suggests that a choice of  $m = 5$  or 6 for Option 1 and  $m = 8$  or 9 for Option 2 offers a reasonable compromise.

As a final choice, we select 2 schemes: the best subset when  $m = 6$  [XMEAS(4, 18, 20, 24, 31, 34)] and  $m = 9$  [XMEAS(1, 3, 11, 14, 16, 18, 20, 21) and XMV(6)] in Option 1 and 2, respectively.

The optimal combination matrices for the selected two schemes are calculated as shown at the bottom of the page,

$$\mathbf{H}_1 = \begin{bmatrix} -42.7 & 0.88 & 0.33 & 0.02 & 0.15 & 0.11 & 0.79 \\ -45.4 & 0.82 & 0.38 & 0.07 & -0.26 & -0.12 & -0.36 \\ -26.9 & 0.48 & 0.26 & 0.0014 & 0.02 & 0.50 & -0.34 \end{bmatrix}$$

$$\mathbf{H}_2 = \begin{bmatrix} -73.4 & -1.32 & 0.0005 & 0.16 & 0.59 & 0.007 & 0.65 & -0.04 & -0.13 & -0.05 \\ -11.0 & -0.59 & 0.0009 & 0.25 & 0.18 & -0.004 & 0.21 & 0.09 & -0.45 & 0.02 \\ 227.1 & -6.1 & 0.006 & -2.39 & -0.91 & 0.03 & 4.66 & -0.39 & -2.90 & -0.03 \end{bmatrix}$$

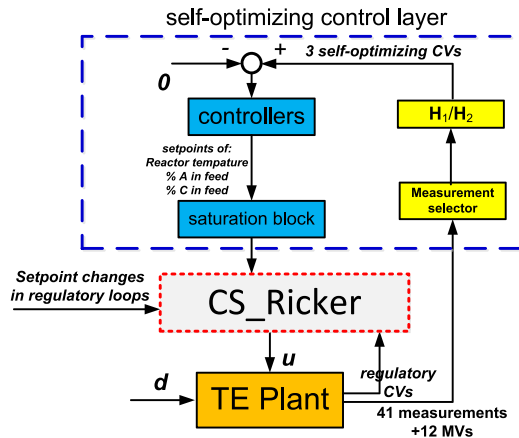


Fig. 3. Retrofit self-optimizing control strategy for the TEP.

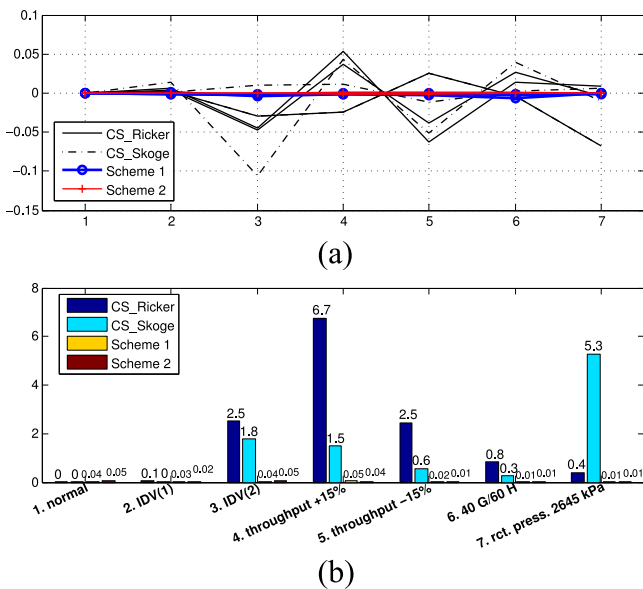


Fig. 4. Optimums over seven operating scenarios: (a) optimal CV values and (b) economic loss comparison.

where the first columns of both  $\mathbf{H}_1$  and  $\mathbf{H}_2$  correspond to the virtual measurement, constant 1.

### C. Step 3. Retrofit Control Structure

With obtained self-optimizing CVs, it is straightforward to configure a retrofit SOC architecture following the design approach as has been presented. Using CS\_Ricker as the existing control system, the three self-optimizing CVs are now the primary CVs to replace reactor temperature, %A and %C in feed, whose setpoints are now manipulated to control the new self-optimizing CVs, as indicated in Fig. 3.

### D. Step 4. Performance Validation

**1) Steady-State Evaluation:** The steady-state performances for selected control schemes is evaluated, which are calculated through the nonlinear process model of TEP. The results are summarized in Fig. 4. To get a more intuitive overview

of the results, Fig. 4(a) and (b) shows various optimal measurement values under all seven operating scenarios. (In the figure, measurement values are first scaled by values at the normal condition, then centered to 0.) This is for comparison on the criterion to select variables with optimal values insensitive to disturbances [2]. As shown in Fig. 4(a), both CS\_Ricker and CS\_Skoge selected measurements with optimal values varying as operating condition changes. Nevertheless, the desired insensitivity of CVs can be achieved by measurement combinations. All six optimal CV values of Scheme 1 and Scheme 2 are almost invariant to operating condition changes. Fig. 4(b) clearly shows that the economic losses is significantly reduced comparing to CS\_Ricker and CS\_Skoge. For both retrofit schemes, none of a loss larger than 0.05 is observed. However, their losses (0.04–0.05) are slightly bigger than CS\_Ricker and CS\_Skoge at the normal point. This is because in order to minimize the global average loss we allow some extra loss occur at a particular operating point. The total loss for Scheme 1 and 2 is calculated as 0.16 and 0.19 \$/h, respectively.

**2) Dynamic Simulations:** Finally, dynamic simulations are performed upon the retrofit control structure. For steady-state decoupling at the reference point, the combination matrix,  $\mathbf{H}$  is modified by left-multiplying a nonsingular matrix,  $\mathbf{M}$  such that the resulting gain matrix,  $\mathbf{MHG}_{y,r} = \mathbf{I}$ . This gives,  $\mathbf{M} = (\mathbf{HG}_{y,r})^{-1}$ . Note that the SOC performances of  $\mathbf{H}$  and  $\tilde{\mathbf{H}} = \mathbf{MH}$  are equivalent. With this treatment, the combination matrices used in dynamic simulation for Scheme 1 and 2 are modified accordingly. PI controllers are employed for all loops configured in the upper control layer, where controller parameters were tuned with the SIMC rule [20], which ensures the stability. For best insurance of safe operation, the outputs for these controllers are constrained within 123–125 °C, 30–34%, 18–20%, respectively. These ranges are selected based on experiences of TE operation.

A sequence of operating scenarios is designed as follows: (1) initially, the system is operated under normal condition; (2) at 10 h, the setpoint of throughput is increased by +15%; and (3) At 80 h, the throughput is reset back, while the reactor pressure is changed from 2800 to 2645 kPa. All setpoint changes are ramped within 10 h to avoid abrupt fluctuations.

The designed Scheme 1 and 2, together with CS\_Ricker and CS\_Skoge, are applied to the aforementioned arranged scenarios. Obtained dynamic responses show that regulatory control tasks can be completed nicely for all four cases, however, their economic behaviors are different, as indicated in Fig. 5. (More detailed responses of various variables are however omitted here due to space limitation.) For CS\_Ricker, it is noted that at about 20–80 h (throughput +15%), the operating cost runs above the minimal cost line (red line). However, when the reactor pressure is maintained at 2645 kPa (90–150 h), the operating cost is around the minimal cost. Finally, in the whole 150-h simulation time, the overall economic cost for CS\_Ricker is calculated as 21088 \$. While for CS\_Skoge, it is observed that during 20–80 h, the economic cost, however, is maintained virtually around the minimal value, which indicates that the economic loss is not large. In the phase of 90–150 h, the economic cost approxi-

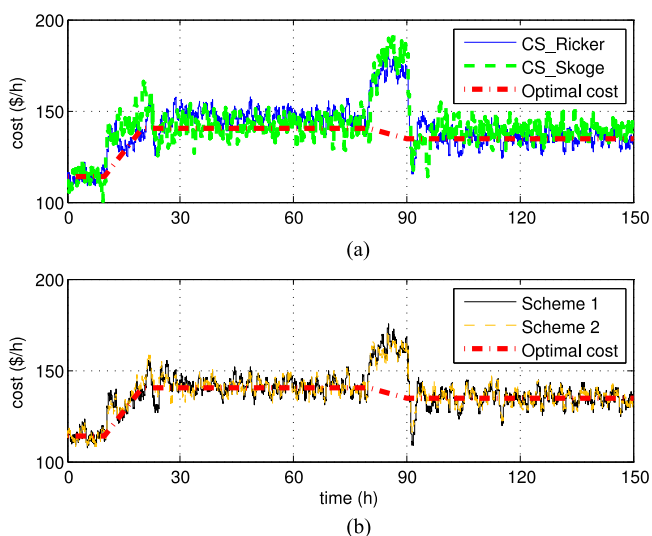


Fig. 5. Dynamic responses of economic costs: (a) CS\_Ricker and CS\_Skoge and (b) Scheme 1 and 2.

mately runs above minimal red line, thus a notable loss occurs. The final operating cost for this control system is 21195 \$, which is even worse than CS\_Ricker for the arranged operating scenarios. As comparisons, the economic costs for Scheme 1 and 2 are automatically maintained around the minimum over the whole simulation time, which indicate good economic performances. The overall costs is calculated as 20640 \$ and 20639 \$, respectively, for Scheme 1 and 2. It is noticeable that for all cases, the cost during 80–90 h experienced a dramatic increase, which was mainly caused by an increase of the purge rate in order to reduce reactor pressure. How to reduce such transient cost generally requires research on dynamic SOC in future. In summary, since the economic costs here have been significantly reduced (about 450–550 \$) as compared to two state of art control systems, one sees that the adopted retrofit strategy is basically successful.

## VI. CONCLUSION

In this paper, we made an effort to push the theory of self-optimizing control into practical applications considering the current industrial practice. The proposed retrofit SOC methodology aims to improve the economic performance of control systems that have been in usage for chemical plants. In general, the approach is to add a new self-optimizing control layer on top of the existing control system, where the derived self-optimizing CVs are controlled by adjusting setpoints of the lower layer control loops. A four-step procedure is given for configuration of a retrofit control structure.

The methodology was applied to a plant-wide TEP testbed by configuring the retrofit SOC control system upon CS\_Ricker. The self-optimizing CVs were selected as measurement combinations using the gSOC method, and the measurement subsets were identified by a tailored PB<sup>3</sup> algorithm. Two different schemes were considered with and without incorporating the composition variables in the self-optimizing CVs, which are controlled by adjusting setpoints of reactor temperature, %A

and %C in feed. The optimality of the designed control architecture is validated both through steady-state analysis and dynamic simulations to demonstrate a great potential of the proposed retrofit methodology. To take full advantage from the advancement of artificial intelligence (AI), various AI tools will be adopted for CV selection in future work. On the other hand, application of the proposed method to real industrial process is also currently under consideration.

A package contains MATLAB code and model to generate results presented in this paper is available in MATLAB File Exchange [21].

## REFERENCES

- [1] L. Ye, Y. Cao, X. Ma, and Z. Song, "A novel hierarchical control structure with controlled variable adaptation," *Ind. Eng. Chem. Res.*, vol. 53, no. 38, pp. 146 95–147 11, 2014.
- [2] S. Skogestad, "Plantwide control: The search for the self-optimizing control structure," *J. Process Control*, vol. 10, no. 5, pp. 487–507, 2000.
- [3] T. Larsson, K. Hestetun, E. Hovland, and S. Skogestad, "Self-optimizing control of a large-scale plant: The Tennessee Eastman process," *Ind. Eng. Chem. Res.*, vol. 40, no. 22, pp. 4889–4901, Oct. 31, 2001.
- [4] I. J. Halvorsen, S. Skogestad, J. C. Morud, and V. Alstad, "Optimal selection of controlled variables," *Ind. Eng. Chem. Res.*, vol. 42, no. 14, pp. 3273–3284, 2003.
- [5] V. Kariwala, Y. Cao, and S. Janardhanan, "Local self-optimizing control with average loss minimization," *Ind. Eng. Chem. Res.*, vol. 47, no. 4, pp. 1150–1158, 2008.
- [6] V. Alstad, S. Skogestad, and E. S. Hori, "Optimal measurement combinations as controlled variables," *J. Process Control*, vol. 19, no. 1, pp. 138–148, 2009.
- [7] L. Ye, Y. Cao, Y. Li, and Z. Song, "Approximating necessary conditions of optimality as controlled variables," *Ind. Eng. Chem. Res.*, vol. 52, no. 2, pp. 798–808, 2013.
- [8] L. Ye, Y. Cao, and Y. Xiao, "Global approximation of self-optimizing controlled variables with average loss minimization," *Ind. Eng. Chem. Res.*, vol. 54, no. 48, pp. 120 40–120 53, 2015.
- [9] Y. Cao and V. Kariwala, "Bidirectional branch and bound for controlled variable selection: Part I. Principles and minimum singular value criterion," *Comput. Chem. Eng.*, vol. 32, no. 10, pp. 2306–2319, 2008.
- [10] V. Kariwala and Y. Cao, "Bidirectional branch and bound for controlled variable selection: Part II. Exact local method for self-optimizing control," *Comput. Chem. Eng.*, vol. 33, no. 8, pp. 1402–1412, 2009.
- [11] V. Kariwala and Y. Cao, "Bidirectional branch and bound for controlled variable selection: Part III. Local average loss minimization," *IEEE Trans. Ind. Informat.*, vol. 6, no. 1, pp. 54–61, Feb. 2010.
- [12] R. Yelchuru and S. Skogestad, "Convex formulations for optimal selection of controlled variables and measurements using mixed integer quadratic programming," *J. Process Control*, vol. 22, pp. 995–1007, 2012.
- [13] L. Ye, Y. Cao, X. Yuan, and Z. Song, "Retrofit self-optimizing control of Tennessee Eastman process," in *Proc. 17th IEEE Int. Conf. Technol.*, Taipei, Taiwan, 2016, pp. 866–871.
- [14] L. Ye, Y. Cao, X. Yuan, and Z. Song, "Subset measurement selection for globally self-optimizing control of tennessee eastman process," in *Proc. 11th IFAC Symp. Dynamics Control Process Syst., Including Biosyst.*, Trondheim, Norway, 2016, pp. 121–126.
- [15] V. Alstad and S. Skogestad, "Null space method for selecting optimal measurement combinations as controlled variables," *Ind. Eng. Chem. Res.*, vol. 46, no. 3, pp. 846–853, 2007.
- [16] V. Kariwala, "Optimal measurement combination for local self-optimizing control," *Ind. Eng. Chem. Res.*, vol. 46, no. 11, pp. 3629–3634, 2007.
- [17] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, vol. 17, no. 3, pp. 245–255, 1993.
- [18] N. Ricker, "Decentralized control of the Tennessee Eastman challenge process," *J. Process Control*, vol. 6, no. 4, pp. 205–221, 1996.
- [19] N. Ricker, "Optimal steady-state operation of the Tennessee-Eastman challenge process," *Comput. Chem. Eng.*, vol. 19, no. 9, pp. 949–959, 1995.
- [20] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *J. Process Control*, vol. 13, pp. 291–309, 2003.
- [21] Y. Cao and L. Ye, "TEP Retrofit gSOC Case Study," MATLAB File Exchange, Jan. 2017. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/61071>





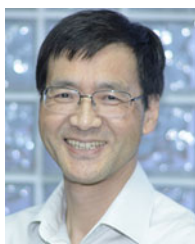
**Lingjian Ye** received the B.Eng. degree in biochemical engineering and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2006 and 2011, respectively.

Since 2011, he has been with the School of Information Science and Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo, China, where he is currently an Associate Professor. His research interests include control structure design, plant-wide control, real-time optimization, and their applications in process systems.



**Xiaofeng Yuan** received the B.Eng. degree in automation and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2011 and 2016, respectively.

He was a Visiting Scholar with the Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB, Canada, from November 2014 to May 2015. He is currently an Associate Professor with the School of Information Science and Engineering, Central South University, Changsha, China. His research interests include industrial process soft sensor modeling, process data analysis, and machine learning.



**Yi Cao** (M'96–SM'10) received the M.Sc. degree in control engineering from Zhejiang University, Hangzhou, China, in 1985, and the Ph.D. degree in engineering from the University of Exeter, Exeter, U. K., in 1996.

He is a Reader with the School of Water, Energy and Environment, Cranfield University, Cranfield, U.K. His research interests include advanced process control, plant-wide control, nonlinear system identification, nonlinear model predictive control, and process monitoring.



**Zhihuan Song** received the B.Eng. and M.Eng. degrees in industrial automation from Hefei University of Technology, Anhui, China, in 1983 and 1986, respectively, and the Ph.D. degree in industrial automation from Zhejiang University, Hangzhou, China, in 1997.

Since 1997, he has been with the Department of Control Science and Engineering, Zhejiang University, where he was at first a Postdoctoral Research Fellow, then an Associate Professor, and is currently a Professor. His research interests include the modeling and fault diagnosis of industrial processes, embedded control systems, and advanced process control technologies. He has published more than 200 papers in journals and conference proceedings.