IFAC-PapersOnLine, Volume 50, Issue 1, 2017, Pages 4672-4677
DOI: 10.1016/j.ifacol.2017.08.691

Global Self-Optimizing Control for Uncertain Constrained Process Systems ⋆

Lingjian Ye * **Yi Cao** ** **Sigurd Skogestad** ***

* *Ningbo Institute of Technology, Zhejiang University, 315100, Ningbo, China; Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), 7491, Trondheim, Norway (e-mail: lingjian.ye@gmail.com)*
** *School of Water, Energy and Environment, Cranfield University, Cranfield, Bedford MK43 0AL, UK (e-mail: y.cao@cranfield.ac.uk)*
*** *Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), 7491, Trondheim, Norway (e-mail: skoge@ntnu.no)*

Abstract: Self-optimizing control is a promising control strategy to achieve real-time optimization (RTO) for uncertain process systems. Recently, a global self-optimizing control (gSOC) approach has been developed to extend the economic performance to be globally acceptable in the entire uncertain space spanned by disturbances and measurement noise. Nevertheless, the gSOC approach was derived based on the assumption of no change in active constraints, which limits the applicability of the approach. To address this deficiency, this paper proposes a new CV selection approach to handle active constraint changes. It ensures that all constraints are within their feasible regions when the selected CVs are maintained at constant setpoints for all expected uncertainties. In particular, constraints of interest are linearized at multiple operating conditions to get better estimates of their values and then incorporated into the optimization formulation when solving the globally self-optimizing CVs. The new CV selection approach is able to ensure an improved operational economic performance without potential constraint violations, as illustrated in an evaporator case study.

Keywords: self-optimizing control, real-time optimization, constrained process, uncertain process

1. INTRODUCTION

Real-time optimization (RTO) refers to a class of technology that maintains/restores plant operational optimality in the face of uncertainties, which is extremely important in a highly competitive market. Among various RTO solutions, self-optimizing control (SOC) strategy (Skogestad, 2000) is a promising approach through feedback control of appropriate controlled variables (CVs) selected. A key feature of SOC is that by choosing what to control in the first place, the strategy actually establishes a link between regulatory control and optimization, which were traditionally considered separately. The link leads to RTO realized along with normal feedback control. An appealing advantage by integrating optimization and regulatory control is that, as compared to other RTO approaches, the optimizing speed is enhanced because RTO is accomplished upon convergence of successive steady-state switching caused by disturbances. Moreover, SOC is also compatible to traditional control structure design problems such as the plant-wide control (Downs and Skogestad, 2011). Note that SOC is a complement to other RTO strategies, including extremum seeking control, as the setpoints for self-optimizing variables are degrees of freedom.

In recent years, SOC has been developed toward different directions to embrace more complex systems that widely exist in chemical industry. Among them, global SOC (gSOC) approach has been recently proposed to enhance the self-optimizing control performance (Ye et al., 2015), where the operating region with acceptable loss was enlarged. In another work, Jaschke and Skogestad used the elimination theory to select polynomial CVs by decoupling unknown disturbances from measurements (Jaschke and Skogestad, 2012). These methods are of potential dealing with nonlinear plants.

For dynamic optimization, Dahl-Olsen et al. (2008) used the maximum gain rule for CV selection of batch processes. Later, a local perturbation control approach was proposed to determine the CV candidates for dynamic SOC problem (Hu et al., 2012b). In other works, the Hamiltonian function was directly controlled to achieve SOC of batch processes (Jaschke et al., 2011; Ye et al., 2013), which however require exact Hamiltonian function analytically derived. More recently, Grema and Cao (2016) developed a data-driven dynamic self-optimizing control approach for oil reservoir water flooding control.

⋆ The author L. Ye gratefully acknowledge the National Natural Science Foundation of China (NSFC) (61673349, 61304081), Ningbo Natural Science Foundation (2015A610151) and China Scholarship Council (No. 201508330751).

Published by Elsevier. This is the Author Accepted Manuscript issued with: Creative Commons Attribution Non-Commercial No Derivatives License (CC:BY:NC:ND 4.0). The final published version (version of record) is available online at DOI:10.1016/j.ifacol.2017.08.691. Please refer to any applicable publisher terms of use.

For simplification, most SOC approaches assume active constraints invariant. Nevertheless, this assumption is invalid in practice. To deal with active constraint varying, the null space method (Alstad and Skogestad, 2007) was extended to cover this situation (Manum and Skogestad, 2012). In this method, the disturbance space was firstly divided into different subregions where active sets remain unchanged within each of them. Then, CV values of neighboring regions were monitored to detect region transition thus implemented CVs are switched online. In another approach for constrained processes, CV was selected to minimize the local average loss while ensuring all the constraints satisfied over the allowable set of uncertainties by explicitly absorbing the constraints in to the optimal CV selection formulation (Hu et al., 2012a). In a price of some conservation, this approach does not require CV switching online, hence the control policy can be simplified. A third approach is to use "back-off" where the setpoints are adjusted optimally in order to avoid reaching new constraints (Govatsmark and Skogestad, 2005).

The perspective of this work is to extend the recently developed gSOC approach (Ye et al., 2015) to constrained processes with active set changes. The gSOC is applicable to non-linear systems but based on invariant active constraints. Since active set changes widely exist in practical plant operations, it would be interesting to extend the gSOC approach to such scenarios. The idea adopted in this paper is similar to the one in (Hu et al., 2012a), that is, incorporating the constraints of interest into the SOC formulation and generating an invariant set of CVs. However, improvements have been made in this paper by linearizing the constraints at multiple operating conditions to get better estimates of constraint values, which makes it more precise to compromise between operational feasibility and economic enhancement.

2. UNCONSTRAINED GSOC

For processes without constraints or with an invariant set of active constraints, the following unconstrained optimization problem is generally applicable.

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

with available measurements

$$\mathbf{y}_m = \mathbf{y} + \mathbf{n} = \mathbf{f}(\mathbf{u}, \mathbf{d}) + \mathbf{n} \tag{2}$$

where J is the scalar cost function to be minimized, which is an economic index for plant operation, $\mathbf{u}$ and $\mathbf{d}$ are the manipulated variables and uncertain disturbances, respectively. $\mathbf{y}_m$, $\mathbf{y}$, $\mathbf{n}$ are the measured, true output variables and measurement noises/errors, respectively, and $\mathbf{f}$ is the input-output mapping function.

Define the economic loss $L \geq 0$ as the difference between J and its optimal value J^{opt}, i.e.

$$L = J(\mathbf{u}, \mathbf{d}) - J^{\text{opt}}(\mathbf{d}) \tag{3}$$

Assuming uncorrelated $\mathbf{d}$ and $\mathbf{n}$, the average loss over the entire uncertain operating region is

$$L_{\text{av}} = E[L] = \int_{\mathbf{d}\in\mathcal{N},\mathbf{n}\in\mathcal{D}} \rho(\mathbf{d})\rho(\mathbf{n}) L \, \mathrm{d}\mathbf{n}\mathrm{d}\mathbf{d} \tag{4}$$

where $\mathcal{D}$ and $\mathcal{N}$ is the variation region spanned by $\mathbf{d}$ and $\mathbf{n}$, respectively, $\rho(\cdot)$ is the probability density of a variable. The objective is to select linear measurement combinations, $\mathbf{c} = \mathbf{H}\mathbf{y}$ as CVs such that when they are maintained at constant setpoints, $\mathbf{c}_s$, the average economic loss L_{av} is minimized over the entire uncertain space.

The global SOC (gSOC) method proposed recently (Ye et al., 2015) approximately minimizes the average loss over all operating scenarios hence exhibits a more guaranteed performance than previous local ones (Alstad and Skogestad, 2007; Alstad et al., 2009; Kariwala, 2007; Kariwala et al., 2008). In gSOC, the CVs (matrix $\mathbf{H}$) and also their setpoints ($\mathbf{c}_s$) are considered together for optimization. This is done by introducing the general CVs, $\hat{\mathbf{c}} \triangleq \mathbf{c} - \mathbf{c}_s$, 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 an artificial measurement (vector $\mathbf{1}$), *i.e.* $\hat{\mathbf{y}} = \begin{bmatrix}\mathbf{1} & \mathbf{y}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$. For simplifying notation, in the remaining part of the paper, $\mathbf{c}$ and $\mathbf{y}$ will be used to representing $\hat{\mathbf{c}}$ and $\hat{\mathbf{y}}$, respectively.

Firstly, the economic loss L is approximated with a quadratic function in terms of $\mathbf{c}$ (Halvorsen et al., 2003)

$$L = \frac{1}{2}\mathbf{e}_c^{\mathrm{T}}\mathbf{J}_{cc}\mathbf{e}_c \tag{5}$$

where $\mathbf{e}_c \triangleq \mathbf{c} - \mathbf{c}^{\text{opt}}$ is the deviation of $\mathbf{c}$ from its optimal value $\mathbf{c}^{\text{opt}}$, $\mathbf{J}_{cc}$ is the Hessian of J with respect to $\mathbf{c}$ at the optimal point, which can be evaluated as $\mathbf{J}_{cc} = (\mathbf{H}\mathbf{G}_y)^{-\mathrm{T}}\mathbf{J}_{uu}(\mathbf{H}\mathbf{G}_y)^{-1}$ ($\mathbf{G}_y$ and $\mathbf{J}_{uu}$ are the sensitivity matrix of $\mathbf{y}$ and Hessian of J respectively, both with respect to $\mathbf{u}$). Since at the optimum, $\mathbf{c}^{\text{opt}} = \mathbf{H}\mathbf{y}^{\text{opt}}$ and considering that the measured CVs, $\mathbf{c}_m$ are controlled as $\mathbf{c}_m = \mathbf{H}\mathbf{y}_m = \mathbf{c}_s = 0$, the true value of $\mathbf{c}$ is $\mathbf{c} = \mathbf{c}_m - \mathbf{H}\mathbf{n} = -\mathbf{H}\mathbf{n}$. Therefore, $\mathbf{e}_c = -\mathbf{H}(\mathbf{y}^{\text{opt}} + \mathbf{n})$.

In local SOC, for an analytical derivation, it is assumed that $\mathbf{y}^{\text{opt}}$ depends linearly on the disturbances $\mathbf{d}$, that is, $\mathbf{y}^{\text{opt}} = \mathbf{F}\mathbf{d}$ where $\mathbf{F}$ is a constant sensitivity matrix (Alstad et al., 2009). Further, assuming that $\mathbf{W}_d^2 = E(\mathbf{d}\mathbf{d}^{\mathrm{T}})$ and $\mathbf{W}_n^2 = E(\mathbf{n}\mathbf{n}^{\mathrm{T}})$ are the covariance matrices of (expected) disturbances the measurement errors (noise), this leads to the following exact value for the average loss

$$L_{\text{av}} = E(L) = \frac{1}{2}\|\tilde{\mathbf{F}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{J}_{cc}^{1/2}\|_{\mathrm{F}}^2 \tag{6}$$

where

$$\tilde{\mathbf{F}} = [\mathbf{F}\mathbf{W}_d \;\; \mathbf{W}_n] \tag{7}$$

Without loss of generality, $\mathbf{J}_{cc} = \mathbf{I}$ can be enforced by incorporating a constraint of $\mathbf{H}\mathbf{G}_y = \mathbf{J}_{uu}^{1/2}$. The optimal CVs, $\mathbf{c} = \mathbf{H}\mathbf{y}$ are then obtained as the solution to (Alstad et al., 2009)

$$\min_{\mathbf{H}} \; L_{\text{av}} = \min_{\mathbf{H}} \; \frac{1}{2}\|\tilde{\mathbf{F}}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\|_{\mathrm{F}}^2, \qquad \text{s.t. } \mathbf{H}\mathbf{G}_y = \mathbf{J}_{uu}^{1/2} \tag{8}$$

gSOC made improvements by accounting for the nonlinearity between $\mathbf{y}^{\text{opt}}$ and $\mathbf{d}$, which somehow complicates the problem and makes L_{av} analytically incalculable. As a solution, Monte Carlo simulations were proposed to evaluate the average loss.

Basically, by introducing matrices $\mathbf{Y}$ and $\tilde{\mathbf{Y}}$ as

$$\mathbf{Y} = \begin{bmatrix}\mathbf{y}_{(1)}^{\text{opt}} & \cdots & \mathbf{y}_{(N)}^{\text{opt}}\end{bmatrix}, \qquad \tilde{\mathbf{Y}} = \begin{bmatrix}\frac{1}{\sqrt{N}}\mathbf{Y} & \mathbf{W}_n\end{bmatrix} \tag{9}$$

where N is the number of sampled disturbances scenarios big enough for a reliable average loss but still computa-

tionally tractable, $\mathbf{y}_{(i)}^{\text{opt}}$ denotes the optimal measurement vector under ith disturbances scenario, see (Ye et al., 2015) for theoretical derivations. The global optimal CV selection problem is finally formulated as

$$\min_{\mathbf{H}} \; L_{\text{av}} = \frac{1}{2}\|\tilde{\mathbf{Y}}^{\text{T}}\mathbf{H}^{\text{T}}\|_{\text{F}}^2, \quad \text{s.t. } \mathbf{H}\mathbf{G}_{y,\text{ref}} = \mathbf{J}_{uu,\text{ref}}^{1/2} \tag{10}$$

where the subscript $(\cdot)_{\text{ref}}$ denotes a chosen reference operating point. This problem is convex and an analytical solution follows in the same form as in local SOC methods (Alstad et al., 2009; Ye et al., 2015)

$$\mathbf{H}^{\text{T}} = (\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^{\text{T}})^{-1}\mathbf{G}_{y,\text{ref}}(\mathbf{G}_{y,\text{ref}}^{\text{T}}(\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^{\text{T}})^{-1}\mathbf{G}_{y,\text{ref}})^{-1}\mathbf{J}_{uu,\text{ref}}^{1/2} \tag{11}$$

In summary, the CVs are selected as follows: (a) Sampling the disturbance space using Monte Carlo method, N disturbance scenarios $\{\mathbf{d}_{(i)}\}$, $i = 1, ..., N$, are generated; (b) For each scenario $\mathbf{d}_{(i)}$, the original optimization problem (1) is solved and corresponding optimal measurements, $\mathbf{y}_{(i)}^{\text{opt}}$ are stored. Then $\mathbf{Y}$ and $\tilde{\mathbf{Y}}$ are constructed according to eq (9); (c) Choose a reference point, the gain matrix, $\mathbf{G}_{y,\text{ref}}$ and Hessian of cost function, $\mathbf{J}_{uu,\text{ref}}$ are evaluated; (d) The combination matrix $\mathbf{H}$ is solved using eq (11).

3. CONSTRAINED GSOC

3.1 Problem formulation

Now consider a general constrained NLP as

$$\min_{\mathbf{u}} \; J(\mathbf{u}, \mathbf{d}), \quad s.t. \; \mathbf{g}(\mathbf{u}, \mathbf{d}) \leq 0 \tag{12}$$

where $\mathbf{g}$ are n_g operational constraints, other symbols are defined in the same way as in last section.

By carrying out optimization on all operating scenarios sampled in the whole operating space, the elements in $\mathbf{g}$ can be classified into three types: (1) $\mathbf{g}_1$: always active; (2) $\mathbf{g}_2$: always inactive; (3) $\mathbf{g}_3$: vary between active and inactive in terms of different operating conditions. The $\mathbf{g}_1$ constraints are firstly assigned with the same number of DOF in $\mathbf{u}$ and $\mathbf{g}_1$ is assumed to be directly controlled. In this way, one reformulates the NLP and select the CVs in the reduced space, which is exactly what has previously been done in traditional SOC. The $\mathbf{g}_2$ constraints can be omitted because they have no influence on the optimal solution. In contrast, $\mathbf{g}_3$ is the changing active constraints and remains a challenge. In the following, we are allowed to assume without loss of generality that $\mathbf{g}$ in problem (12) all fall into type $\mathbf{g}_3$.

For constrained processes, we want to select $\mathbf{H}$ such that the economic loss is minimized by controlling $\mathbf{c} = \mathbf{H}\mathbf{y}$ at constant setpoints and meanwhile, the constraints $\mathbf{g}(\mathbf{u}, \mathbf{d}) \leq 0$ are satisfied in the whole operating space, thus guaranteing feasibility. To this end, write

$$\begin{aligned} L_{\text{con}} = J - J_{\text{con}}^{\text{opt}} &= \underbrace{J - J_{\text{uncon}}^{\text{opt}}}_{L_{\text{uncon}}} - \underbrace{(J_{\text{con}}^{\text{opt}} - J_{\text{uncon}}^{\text{opt}})}_{\varepsilon(\mathbf{d})} \\ &= L_{\text{uncon}} - \varepsilon(\mathbf{d}) \end{aligned} \tag{13}$$

where $J_{\text{uncon}}^{\text{opt}}$ and $J_{\text{con}}^{\text{opt}}$ is the minimal cost function for constrained and unconstrained problem (1) and (12) respectively, satisfying $J_{\text{uncon}}^{\text{opt}} \leq J_{\text{con}}^{\text{opt}}$. The equality holds when none of $\mathbf{g}$ is active for a particular $\mathbf{d}$. L_{uncon} is the economic loss for the *same optimization problem* but, without considering $\mathbf{g}$. The nonnegative term $\varepsilon(\mathbf{d})$ is defined as the residual between the minimal cost of constrained and unconstrained process. Note that $\varepsilon(\mathbf{d})$ is only a function of $\mathbf{d}$ for a given optimization problem.

The average constrained loss $L_{\text{av,con}}$ is

$$L_{\text{av,con}} = \frac{1}{N}\sum_{i}^{N}(L_{\text{uncon},(i)} - \varepsilon_{(i)}) = L_{\text{av,uncon}} - \frac{1}{N}\sum_{i}^{N}\varepsilon_{(i)} \tag{14}$$

Then the CV combination matrix $\mathbf{H}$ is obtained as

$$\min_{\mathbf{H}} L_{\text{av,con}} = \min_{\mathbf{H}} \left[L_{\text{av,uncon}} - \frac{1}{N}\sum_{i}^{N}\varepsilon_{(i)} \right] \tag{15}$$

$$s.t. \; \mathbf{g}_{(i)}|_{\mathbf{c}=0} \leq 0, \;\; \forall\, i = 1, \ldots, N \tag{16}$$

where $\mathbf{g}_{(i)}$ is satisfied for all N sampled operating conditions. We note that $\varepsilon_{(i)}$ has no connection with $\mathbf{H}$, hence it can be dropped without affecting solving for the optimal $\mathbf{H}$. Therefore, we equivalently consider

$$\min_{\mathbf{H}} L_{\text{av,uncon}} \tag{17}$$

$$s.t. \; \mathbf{g}_{(i)}|_{\mathbf{c}=0} \leq 0, \;\; \forall\, i = 1, \ldots, N \tag{18}$$

where the criterion to be minimized is the same as in unconstrained gSOC problem, but with the addition of constraint satisfaction. Therefore, following the gSOC algorithm presented in the previous section, we have

$$\begin{aligned} \min_{\mathbf{H}} \; & L_{\text{av,uncon}} = \frac{1}{2}\|\tilde{\mathbf{Y}}^{\text{T}}\mathbf{H}^{\text{T}}\|_{\text{F}}^2 \\ \text{s.t.} \quad & \mathbf{H}\mathbf{G}_{y,\text{ref}} = \mathbf{J}_{\text{uu,ref}}^{1/2} \\ & \mathbf{g}_{(i)}|_{\mathbf{c}=0} \leq 0, \;\; \forall\, i = 1, \ldots, N \end{aligned} \tag{19}$$

where the optimal measurements in $\tilde{\mathbf{Y}}$ are generated by solving the unconstrained optimization problem (1).

3.2 Constraint estimation

In a local method (Hu et al., 2012a), constraints are linearized at the single nominal point, that is, the obtained linear model is used to get an estimate of constraint in the whole operating space. That is,

$$\mathbf{g} = \mathbf{g}^{*} + \mathbf{g}_u\Delta\mathbf{u} + \mathbf{g}_d\Delta\mathbf{d} \tag{20}$$

where $\mathbf{g}^{*}$ denotes the constraint value at the nominal point, Δ is the deviation of a variable from the nominal point, $\mathbf{g}_\nu$ stands for the first order derivative of $\mathbf{g}$ in terms of ν evaluated at the nominal point, i.e. the i^{th} row and j^{th} column, $[\mathbf{g}_\nu]_{i,j} = \frac{\partial g_i}{\partial \nu_j}$.

This local approach is extended for the gSOC to consider all disturbance scenarios. For an arbitrary $\mathbf{d}$, the first-order approximation of constraints around the optimum with respect to $\mathbf{c}$ is

$$\mathbf{g} = \mathbf{g}^{\text{opt}} + \mathbf{g}_c\mathbf{e}_c \tag{21}$$

with $\mathbf{g}^{\text{opt}}$ as the constraint value at the optimum for a specific disturbance i ($i = 1, \ldots, N$), $\mathbf{e}_c = \mathbf{c} - \mathbf{c}^{\text{opt}}(\mathbf{d})$ as defined earlier. Here, the use of linearization in terms of $\mathbf{c}$, rather $\mathbf{u}$, facilitates later close-loop analysis. Around the optimum, a local relationship holds as $\Delta\mathbf{c} = \mathbf{H}\mathbf{G}_y\Delta\mathbf{u}$, we have

$$\mathbf{g}_c = \frac{\partial \mathbf{g}}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{c}} = \mathbf{g}_u(\mathbf{H}\mathbf{G}_y)^{-1} \tag{22}$$

Recall that $\mathbf{e}_c = -\mathbf{H}(\mathbf{y}^{\text{opt}} + \mathbf{n})$, substituting above results into (21) follows that

$$\mathbf{g} = \mathbf{g}^{\text{opt}} - \mathbf{g}_u(\mathbf{H}\mathbf{G}_y)^{-1}\,\mathbf{H}(\mathbf{y}^{\text{opt}} + \mathbf{n}) \tag{23}$$

For a particular disturbance scenario $\mathbf{d}_{(i)}$, the constraints $\mathbf{g}_{(i)}$ should be in principle estimated independently, i.e. all related terms including $\mathbf{g}^{\text{opt}}$, $\mathbf{G}_y$, $\mathbf{g}_u$, $\mathbf{y}^{\text{opt}}$ and $\mathbf{n}$ are taken as their values occurred at the optimum of $\mathbf{d}_{(i)}$.

However, we note that the term $(\mathbf{H}\mathbf{G}_y)^{-1}$ requires intensive calculation of an inverse matrix. To simplify the problem one step further, we consider $\mathbf{g}_u(\mathbf{H}\mathbf{G}_y)^{-1}$ is constant for all disturbance conditions, by taking its value at the reference point. Use the relationship of $\mathbf{H}\mathbf{G}_{y,\text{ref}} = \mathbf{J}_{\text{uu,ref}}^{1/2}$, we define $\mathbf{G}_{\text{ref}}^{\mathbf{g}} \coloneqq (\mathbf{g}_u)_{\text{ref}}\,\mathbf{J}_{\text{uu,ref}}^{-1/2}$, which is required to be evaluated only once at the reference point. Consequently, the approximated constraints are

$$\mathbf{g}_{(i)} = \mathbf{g}_{(i)}^0 - \mathbf{G}_{\text{ref}}^{\mathbf{g}}\,\mathbf{H}(\mathbf{y}_{(i)}^{\text{opt}} + \mathbf{n}_{(i)}),\quad i = 1, \ldots, N \tag{24}$$

Here, the slope is only evaluated at a single point, however, corrections are made for the intercept ($\mathbf{g}^{\text{opt}}$) and bias ($\mathbf{e}_c$) from point to point, hence the method here is still stronger than Hu et al. (2012a). Denote the jth element(row) of $\mathbf{g}$ and $\mathbf{G}_{\text{ref}}^{\mathbf{g}}$ as $\mathbf{g}_j$ and $\mathbf{G}_{\text{ref},j}^{\mathbf{g}}$, respectively. Then

$$\mathbf{g}_j \le \bar{\mathbf{g}}_j = \max_{i\in[1,N]}(\mathbf{g}_{(i),j}^0 - \mathbf{G}_{\text{ref},j}^{\mathbf{g}}\mathbf{H}\mathbf{y}_{(i)}^{\text{opt}}) + |\mathbf{G}_{\text{ref},j}^{\mathbf{g}}\mathbf{H}\mathbf{W}_n|_1 \tag{25}$$

where $\bar{\mathbf{g}}_j$ is defined as the worst case constraint which is expected to be bounded less than 0.

To sum up, the optimal CV selection for constrained process is finally formulated as

$$\min_{\mathbf{H}}\ \frac{1}{2}\|\tilde{\mathbf{Y}}^{\text{T}}\mathbf{H}^{\text{T}}\|_{\text{F}}^2,\quad \text{s.t.}\ \begin{cases}\mathbf{H}\mathbf{G}_{y,\text{ref}} = \mathbf{J}_{\text{uu,ref}}^{1/2}\\ \bar{\mathbf{g}}_j \le 0,\ \ \forall\, j = 1, \ldots, n_g\end{cases} \tag{26}$$

which can be solved numerically.

4. CASE STUDY: AN EVAPORATOR EXAMPLE

4.1 Operation requirements

A forced-circulation evaporator (Newell and Lee, 1989) is investigated, as shown in Figure 1. The process involves 3 state variables and 20 process variables, for detailed model equations, we refer to other publications (Newell and Lee, 1989; Kariwala et al., 2008).

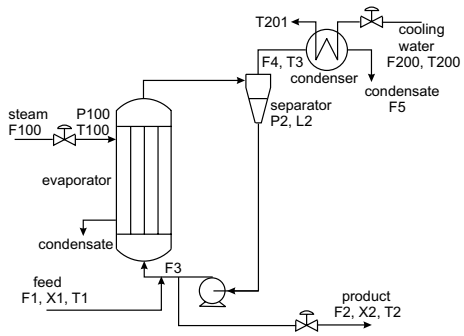

Fig. 1. Evaporator

The cost function to be minimized is

$$J = 600F_{100} + 0.6F_{200} + 1.009(F_2 + F_3) + 0.2F_1 - 4800F_2 \tag{27}$$

The manipulated variables and disturbances are

$$\mathbf{u} = [F_{200}\ F_1\ F_2\ F_3\ P_{100}]^{\text{T}} \tag{28}$$

$$\mathbf{d} = [X_1\ T_1\ T_{200}]^T \tag{29}$$

where the variation ranges for disturbances are defined as $\pm 5\%$ for X_1 and $\pm 20\%$ for T_1 and T_{200} of their nominal values. The following measurements are available

$$\mathbf{y} = [P_2\ T_2\ T_3\ F_2\ F_{100}\ T_{201}\ F_3\ F_5\ F_{200}\ F_1]^T \tag{30}$$

with expected noise magnitudes of 2%, 2.5% and 1°C (uniform) for flowrates, pressures and temperatures, respectively.

The following constraints should be satisfied

$$X_2 \ge 35.5\% \tag{31}$$

$$40 \le P_2 \le 80\ kPa \tag{32}$$

$$P_{100} \le 400\ kPa \tag{33}$$

$$0 \le F_{200} \le 400\ kg/min \tag{34}$$

$$0 \le F_1 \le 20\ kg/min \tag{35}$$

$$0 \le F_3 \le 100\ kg/min \tag{36}$$

Among these process constraints, two for X_2 and P_{100} are always active in the whole disturbance variation regions, hence they should be controlled at their boundaries in the first place. Including the separator level which has no steady state effect, 3 degrees of freedom are consumed. On the other hand, the constraints for F_{200}, F_1 and F_3 are always inactive, hence they can be ignored as long as the process is controlled around the optimum. In contrast, (32) (including two constraints for upper and lower bounds) is found to optimally vary between active or inactive depending on the disturbance. Therefore, without loss of generality, we consider the reduced problem by using

$$\mathbf{u} = [F_{200}\ F_1]^{\text{T}} \tag{37}$$

as the manipulated variables.

4.2 Results and discussions

To apply the gSOC algorithm, Monte Carlo simulation were performed to sample the expected disturbance space with a sequence of 500 random disturbances. Optimal measurements are obtained accordingly by numerical optimization in terms of these disturbances. Note that the constraint (32) does not need to be considered at this stage. All related matrices, including $\mathbf{Y}$, $\mathbf{W}_n$, $\tilde{\mathbf{Y}}$, are obtained. The reference point is chosen as the nominal point, then $\mathbf{G}_{y,\text{ref}}$ and $\mathbf{J}_{\text{uu,ref}}$ are evaluated by using the finite difference method.

For comparisons of different schemes, consider the following possible methods. **Method 1**: gSOC + a single linear constraint model: eq (20); **Method 2**: gSOC + simplified linear constraint model: eq (24); **Method 3**: gSOC + multiple linear constraint model: eq (23). In all cases, the *fmincon* routine of Matlab 2013a is used in Windows 7 OS, hardwares are Intel Core i5 @1.7GHz CPU and 8 GB RAM.

To better explore features of various approaches, we will consider first the case when there is no measurement noise. Furthermore, we investigate the following two measurement subsets for demonstration purpose

$$\mathbf{y}_1 = [F_2\ F_{100}\ F_{200}]^{\text{T}}$$

Table 1. Losses for 100 random disturbances (no measurement error)

		$L_{\text{av,uncon}}$	nonlinear loss	violation (times)	comput. time(s)
$\mathbf{y}_1$	Method 1	15.20	10.64	0	0.342
	Method 2	11.57	8.28	0	2.49
	Method 3	8.53	6.13	1	8.80
$\mathbf{y}_2$	Method 1	7.27	4.82	2	0.812
	Method 2	5.46	2.22	5	5.42
	Method 3	6.17	2.83	0	18.40

$$\mathbf{y}_2 = [P_2\ F_2\ F_5\ F_{200}]^{\mathrm{T}}$$

Results without measurement error. Following the approaches proposed, different combination matrices, $\mathbf{H}$ were obtained for Methods 1–3. The resulting CVs, $[-\mathbf{c}_s\ \mathbf{c}] = \mathbf{Hy}$, are tested for 100 groups of random disturbances under closed-loop control. To make the comparison fair, the same random disturbances are used for different methods, the results are summarized in Table 1.

For subset $\mathbf{y}_1$, the numerically obtained optimal combination matrices are as (the first columns associated with the artificial measurement $\mathbf{1}$ give the optimized setpoints $-\mathbf{c}_s$ for each $\mathbf{H}$)

$$Method\ 1: \mathbf{H} = \begin{bmatrix} -7.14 & -59.87 & 7.53 & 0.075 \\ -27.59 & -5.40 & 4.35 & -0.028 \end{bmatrix}$$

$$Method\ 2: \mathbf{H} = \begin{bmatrix} -7.15 & -59.62 & 7.50 & 0.074 \\ -26.44 & -36.92 & 8.33 & -0.024 \end{bmatrix}$$

$$Method\ 3: \mathbf{H} = \begin{bmatrix} -5.52 & -70.24 & 8.84 & 0.076 \\ -24.56 & -47.72 & 9.70 & -0.023 \end{bmatrix}$$

The losses $L_{\text{av,uncon}}$ for the three methods are 15.20, 11.57 and 8.53, respectively, which imply that the proposed methods made economic improvements. The nonlinear validations confirm with losses 10.64, 8.28 and 6.13, respectively. Note that these losses are calculated through original nonlinear process model incorporating with the constraints, hence they are generally smaller than $L_{\text{av,uncon}}$, which is evaluated using the unrealistic unconstrained case.

For the 100 disturbances (subset $\mathbf{y}_1$), the constraints have all been satisfied in Method 1 and 2. However, one case of constraint violation is observed in Method 3. As shown in Figure 2 (a), Method 1 restricts P_2 within a quite narrow region (roughly between 40–60 kPa), which explains why the loss with Method 1 are relatively large. In other words, using a single linear constraint model in this case gives a conservative result because it overestimates the influence of constraints. When the CVs are obtained with Method 2, P_2 varies in a much wider range as shown in the figure. This allows one to obtain a smaller economic loss while still ensuring constraint satisfaction. Method 3 further relaxes P_2 so that the economic loss is reduced as compared to Method 2. However, a disadvantage is that one case of constraint variation occurs with $P_2 = 39.85\ kPa$, which is very slightly under the limit of 40 kPa. This happens because a linear function is still an approximation of the real plant. However, it does not necessarily imply that the proposed method is less reliable than using a single linear model, as explained shortly.

The results are somewhat different for subset $\mathbf{y}_2$ (numerical CVs are not provided for brevity). While the economic

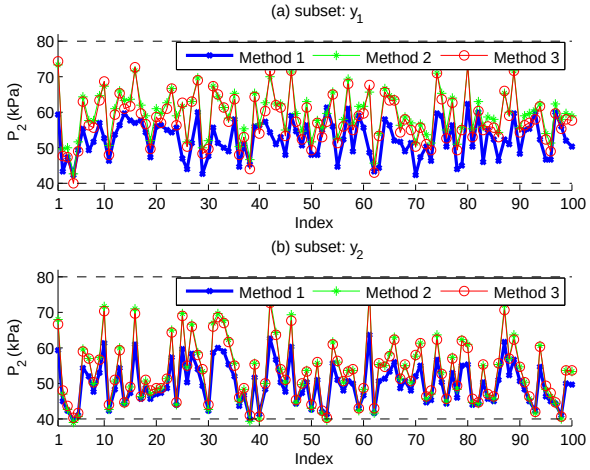

Fig. 2. Pressure P_2 with controlling different CVs for 3 methods (no measurement error)

Table 2. Losses for 100 random disturbances (with measurement error)

		$L_{\text{av,uncon}}$	nonlinear loss	violation (times)	comput. time(s)
$\mathbf{y}_1$	Method 1	19.13	14.42	0	2.35
	Method 2	16.84	12.31	0	3.21
	Method 3	14.43	11.41	0	14.89
$\mathbf{y}_2$	Method 1	12.30	8.97	0	1.46
	Method 2	11.34	6.57	0	7.40
	Method 3	11.58	6.94	0	36.35

loss for Method 1 is still the largest, this time, Method 2 achieves the smallest loss. On the other hand, Method 2 leads to 5 constraint violations. Method 3 in this case obtained the best overall results, that is, it resulted in no constraint violations and a small economic loss.

Results with measurement error. For subset $\mathbf{y}_1$, the economic losses show the same trend as in the noise-free case, namely, the losses are gradually reduced from Method 1 to 3 (19.13, 16.84 and 14.43 for loss $L_{\text{av,uncon}}$, 14.42, 12.31 and 11.41 for nonlinear losses). However, by incorporating measurement errors, no constraint violations occur for all three methods (notice from Fig.3 that P_2 is restricted in a much narrower region compared to the noise-free case), which is obtained at the price of economic losses increased.

For subset $\mathbf{y}_2$, the loss trends are also the same as in the noise-free case. Method 1 gives the largest loss, whereas Method 2 gives the smallest loss, and this time, Method 2 results in no constraint violations. This is because the conservative consideration of worst case measurement error results in sufficient constraint margin, which compensates for the effect of disturbances.

Implications of including measurement error. In the experiments investigated above, the influence of measurement error is similar to setting a back-off in the constraint, which is a common solution to compensate the estimation error. This is the advantageous effect. Therefore, no constraint violations occurred in the closed-loop validations. However, these results were obtained under the condition that the worst case regarding to measurement error has not happened. When we manually set such a worst case

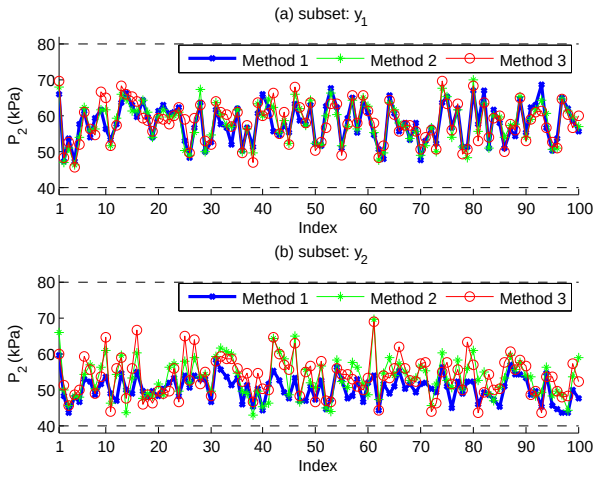

Fig. 3. Pressure P_2 with controlling different CVs for 3 methods (with measurement error)

scenario in the simulation, some constraint violations occurred which is similar to the noise-free case. However, one may argue that such a worse case combination of measurement error is unlikely in practice, especially if there are many measurements. Another fact is that the bandwidth of measurement error is often much higher than process dynamics. Hence, even when the noise under worst case indeed happens, it soon updates to a less severe scenario before the process settles into the steady state. In other words, the process itself acts like a "filter" to counteract some effect of measurement error.

Computational issue. To investigate this issue, average computation times were calculated for all cases by optimizing 100 times over random initial $\mathbf{H}$, see the last column in Table 1 and 2. The results indicate that the consumed computation time ranks as Method 3 > 2 > 1. For example, in the noise-free case with subset $\mathbf{y}_1$, the smallest computation time is 0.342 s for Method 1 while the largest time is 8.80 s for Method 3. The same trend is observed for $\mathbf{y}_2$, however, the computation times all increase (roughly two times), because more measurements involve more calculations. The computation times further increase in the noisy case, because constraint variations are additionally caused by measurement noise. The largest computation time occurs in Method 3 (subset $\mathbf{y}_2$), which averagely requires 36.35 s to find the optimal CVs.

5. CONCLUSIONS

In this paper, we extended the gSOC method to handle constrained processes with active set changes. The perspective is seeking an invariant solution for self-optimizing control, by incorporating constraints satisfaction in the optimization formulation for solving CVs. To account for process nonlinearity, it was proposed to linearize the constraints at multiple operating conditions to get better constraint estimates. Two alternatives were proposed, one was to make full corrections for the constraint model (Method 3), whereas the other remained the gain constant to simplify computation (Method 2). As investigated in the evaporator case, both the two methods improved the economic performance while ensuring the operation safety. Method 3 provided best constraint estimate at the price of intensive computations, while Method 2 achieved acceptable estimate accuracy with less computation time.

REFERENCES

Alstad, V. and Skogestad, S. (2007). Null space method for selecting optimal measurement combinations as controlled variables. *Ind. Eng. Chem. Res.*, 46(3), 846–853.

Alstad, V., Skogestad, S., and Hori, E.S. (2009). Optimal measurement combinations as controlled variables. *J. Process Control*, 19(1), 138–148.

Dahl-Olsen, H., Narasimhan, S., and Skogestad, S. (2008). Optimal output selection for control of batch processes. In *Proc. of American Control Conference*.

Downs, J.J. and Skogestad, S. (2011). An industrial and academic perspective on plantwide control. *Annual Reviews in Control*, 35(1), 99–110.

Govatsmark, M.S. and Skogestad, S. (2005). Selection of controlled variables and robust setpoints. *Industrial & engineering chemistry research*, 44(7), 2207–2217.

Grema, A.S. and Cao, Y. (2016). Optimal feedback control of oil reservoir waterflooding processes. *International Journal of Automation and Computing*, 13(1), 73–80.

Halvorsen, I.J., Skogestad, S., Morud, J.C., and Alstad, V. (2003). Optimal selection of controlled variables. *Ind. Eng. Chem. Res.*, 42(14), 3273–3284.

Hu, W., Umar, L.M., Xiao, G., and Kariwala, V. (2012a). Local self-optimizing control of constrained processes. *J. Proc. Contr.*, 22, 488–493.

Hu, W., Mao, J., Xiao, G., and Kariwala, V. (2012b). Selection of self-optimizing controlled variables for dynamic processes. In *Advanced Control of Chemical Processes*, volume 8, 774–779.

Jaschke, J. and Skogestad, S. (2012). Optimal controlled variables for polynomial systems. *J. Process Control*, 22(1), 167 – 179.

Jaschke, J., Fikar, M., and Skogestad, S. (2011). Self-optimizing invariants in dynamic optimization. In *CDC-ECE*, 7753–7758.

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

Kariwala, V., Cao, Y., and Janardhanan, S. (2008). Local self-optimizing control with average loss minimization. *Ind. Eng. Chem. Res.*, 47(4), 1150–1158.

Manum, H. and Skogestad, S. (2012). Self-optimizing control with active set changes. *J. Process Control*, 22(5), 873–883.

Newell, R.B. and Lee, P. (1989). *Applied process control : a case study*. Prentice-Hall.

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

Ye, L., Cao, Y., and Xiao, Y. (2015). Global approximation of self-optimizing controlled variables with average loss minimization. *Ind. Eng. Chem. Res.*, 54(48), 12040–12053.

Ye, L., Kariwala, V., and Cao, Y. (2013). Dynamic optimization for batch processes with uncertainties via approximating invariant. In *8th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 1786–1791. IEEE.