CONSTRAINED SELF-OPTIMIZING CONTROL VIA DIFFERENTIATION [1]

Yi Cao [*,2]

[*] *School of Engineering, Cranfield University, UK*

Abstract: A new approach using differentiation to design "self-optimizing" (Skogestad, 2000) control system is proposed and applied to the evaporation process of Newell and Lee (1989). Using the chain rule of differentiation, an explicit expression of gradient in terms of system's Jacobian matrices has been derived for the first-order optimal condition of a constrained optimization problem. This gradient function can directly be used as a controlled variable to achieve self-optimization. To cope with conditionally active constraints, a cascade control structure has been proposed. With this structure, the optimal condition and conditionally active constraints can automatically switch each other to be active or inactive depending on disturbances so that both are satisfied. Both ideas have been demonstrated with the evaporator system. For the evaporation process, it is also shown that a traditional engineering judgement for level control structure selection may lead to a wrong decision. Simulation results show that the proposed control system does achieve self-optimization with various disturbances.

Keywords: Plantwide Control, Self-Optimizing Control, Control Schemes, Decentralized Control, Cascade Control, Automatic Differentiation.

1. INTRODUCTION

Chemical process plants are always controlled in different layers. For example, several local control layers are designed to maintain local controlled variables at the desired operating point whilst a plantwide optimization layer is responsible to adjust the setpoint to the local layers according to different situations (disturbances). Traditionally, these two layers are designed separately for different (economic and dynamic) objectives although they need working together. Recently, Skogestad (2000) introduced an important concept of "self-optimizing control", which provides a new link between these two layers. Self-optimization is a control strategy where by controlling certain specially selected variables at their nominal setpoints, it automatically achieves the optimal (or acceptable) operating conditions without re-optimization even in the presence of disturbances.

The optimality of a self-optimizing control system is strongly related to the control structure selected. In his seminal work, Skogestad (2000) proposed a design procedure to select a self-optimizing control structure. This procedure has been applied to several chemical processes such as the Tennessee Eastman process (Larsson *et al.*, 2001) and the evaporation process of Newell and Lee (Govatsmark and Skogestad, 2001). However, in the above work, the controlled variables considered are limited to the existing measurements. Therefore, only suboptimal performance is achievable with the control structure selected, i.e. an average loss is always expected in the control system.

[1] Supported by the EPSRC UK under grant GR/R57324
[2] Email: y.cao@cranfield.ac.uk

In this work, the local-optimal condition of a self-optimizing control system is derived. This condition can be represented in a combination of existing measurements. Therefore, the optimal condition can be used as controlled variables to achieve local self-optimization. Moreover, to deal with "soft-constrained" measurements, a cascade control structure is proposed, which can automatically cope with both requirements of optimality and constraints. This approach has been applied to the evaporation process of Newell and Lee (1989). With a carefully designed decentralized cascade structure, the control system does achieve self-optimization without violating process constraints. Another issue with self-optimizing control is the back-off from the nominally optimal setpoints (Heath *et al.*, 2000). The case study shows that the optimality of the self-optimizing system is very sensitive to the size of back-off. Therefore, a tighter back-off is always looked for. This is achieved in the case study by using the particularly designed decentralized control structure.

The paper is organized as follows: The local-optimal conditions with process constraints for self-optimizing control is derived in section 2. To cope with conditionally active constraints, a cascade control structure to satisfy both optimality and constraint requirements is proposed in section 3. The self-optimizing control condition and structure are applied to the evaporation process in section 4, where a decentralized cascade control structure is design to achieve the minimum back-off and to cope with process constraints. Simulation results of various control schemes are compared in terms of the self-optimality in section 5. The paper is concluded in section 6.

2. LOCAL SELF-OPTIMIZING CONDITIONS

Consider the following optimization problem:

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

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$ and $d \in \mathbb{R}^{n_d}$ are state, input and disturbance variables respectively. For a given disturbance, d, the solution of the above optimization problem is denoted as, x^* and u^*. Assume that at the optimal point, the following equalities hold:

$$F(x^*,u^*,d) = \begin{bmatrix} f(x^*,u^*,d) \\ g_1(x^*,u^*,d) \end{bmatrix} = 0 \tag{2}$$

where $f(\cdot)$ and $g_1(\cdot)$ are vector-valued functions with dimensions of n_f and n_1 respectively. If $m = (n_x + n_u) - (n_f + n_1) \neq 0$, then according to the Kuhn-Tucker conditions, there are m first-order optimal conditions. Denote $u^* = [u_1^T \ u_2^T]^T$ with $u_2 \in \mathbb{R}^m$, $z = [x^{*T} \ u_1^T]^T$ and $v = u_2$. Then the optimization problem (1) can be re-stated as:

$$\begin{aligned} \min_{z,v} J &= \phi(z,v,d) \\ \text{s.t.} \quad & F(z,v,d) = 0 \end{aligned} \tag{3}$$

The first-order optimal conditions of the above optimization problem are:

$$J_v = \phi_v + \frac{\partial z}{\partial v}\phi_z = 0 \tag{4}$$

$$F_v + \frac{\partial z}{\partial v}F_z = 0 \tag{5}$$

If the Jacobian matrix, F_z is not singular, then the second condition (5) gives:

$$\frac{\partial z}{\partial v} = -F_v F_z^{-1} \tag{6}$$

Inserting (6) into the first condition (4) leads to the following m-dimension optimal condition:

$$G(z,v,d) := J_v|_{F=0} = \phi_v - F_v F_z^{-1}\phi_z = 0 \tag{7}$$

Normally, the left-hand-side of the above condition is a function of x^*, u^* (u_1, and u_2) and d. For a given disturbance, d, equation (7) corresponds to an unique solution of $v = u_2$, from which all rest system variables, x^* and u_1 can be determined.

If $F(x^*,u^*,d) = 0$ is the only active constraints for all possible disturbances, then it is clear that $G(z,v,d) = 0$ is the only condition which must be maintained to ensure the process operation is optimal. In other words, if condition $G(z,v,d) = 0$ is retained by the control system, then optimal operation can be achieved without re-optimization for different disturbances, i.e. the plant is self-optimizing controlled.

For an optimization problem, the solution of $G(z,v,d) = 0$ is independent from the selection of variable, v, if F_z is non-singular. However, the selection may have an effect on the actual cost achieved because different v correspond to different flatness of J around the optimal point. The flatter the cost function at the optimal point, the less sensitive to control error, thus the better (Skogestad, 2000). For example, in a single degree of freedom case, *i.e.* both v and z are scalar, similar to (7), if both F_v and F_z are non-singular and z is chosen as the independent variable, then the gradient is

$$J_z|_{F=0} = \phi_z - F_z F_v^{-1}\phi_v = -F_z F_v^{-1} J_v|_{F=0} \tag{8}$$

Therefor, in a proximity of the optimal point, if $|F_z| > |F_v|$, then $|J_v|_{F=0}| < |J_z|_{F=0}|$, *i.e.* the curve of J against v is flatter then the one of J against z around the optimum. Hence, z should be chosen such that $|F_z| > |F_v|$ at the optimal point.

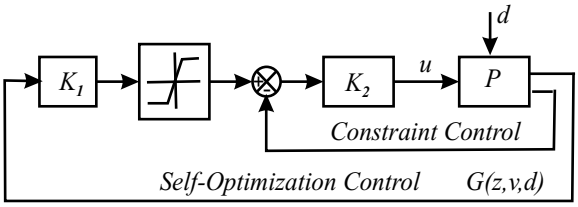

Fig. 1. Cascade structure for self-optimizing and constraint control

For a small system, the gradient function, $G(z, v, d)$ can be derived analytically. Therefore, it should have no difficulty to implement it as a nonlinear soft measurement for feedback control. For a large or complicated process, it may not be a trivial task to get an analytical expression of the gradient function. However, according to equation (7), the gradient function is composed of the first-order derivatives of the cost function and the nonlinear system model functions. Therefore, the numerical values of the gradient function for different disturbances can be calculated by performing on-line linearization at individual sampling time. For this purpose, the automatic differentiation techniques developed in recent years (Griewank, 2000) will play an important role.

3. SELF-OPTIMIZING CONTROL STRUCTURE

According to the above analysis, the control structure of a self-optimizing plant should be selected as follows:

- Stabilization control. Variables related to unstable modes of the plant must be controlled.
- Constraint control. Those included in $g_1 = 0$ need to be controlled.
- Self-optimizing control. The gradient function, $G(z, v, d) = 0$ is to be controlled.

However, active constraints of a process plant may not always be the same. Some output constraints, such as temperature and pressure limits may becomes active under certain circumstances. Traditionally, these variables are always selected as controlled variables. However, by controlling these variables at their nominal setpoints, the plant operation will not be optimal at most times.

To satisfy both requirements of self-optimization and operation constraints, a cascade control structure is proposed as shown in Figure 1.

In Figure 1, an inner loop is closed for constraint control. The setpoint of the inner loop is determined by the outer loop, which is designated for self-optimizing control by maintaining the gradient function $G(z, v, d)$ at zero. Within the feasible range of the process constraint, the setpoint of the inner loop is floating as a manipulated variable to satisfy the self-optimizing condition. However, when disturbances cause the process towards outside of the constraints, the saturation block will limit the setpoint within the constraint so that the controlled variable of the inner loop will be kept within feasible range. In this way, the self-optimizing control and constraint control loops alternatively become active and inactive to achieve constrained self-optimization.

The limits in the saturation block may also need to be consider with a suitable back-off to avoid dynamic violation of the constraints in the worst disturbance.

The cascade structure is more or less similar to traditional multilevel optimizing control system where the optimization layer calculates the optimal setpoint for lower level process control loops. However, in the self-optimizing control configuration, the optimal setpoint is produced by a normal process control loop without realtime online optimization. In a self-optimizing control system, gradient is dynamically calculated. Therefore, the setpoint produced by gradient control changes smoothly from one operating point to another. This is preferable to step setpoint change derived by static optimization and may lead to better results as shown by the case study in the following sections.

4. EVAPORATOR SELF-OPTIMIZING CONTROL

4.1 Gradient function

The constrained self-optimizing control strategy is applied to the evaporation process of Newell and Lee (1989), shown in Figure 2. This is a "forced-

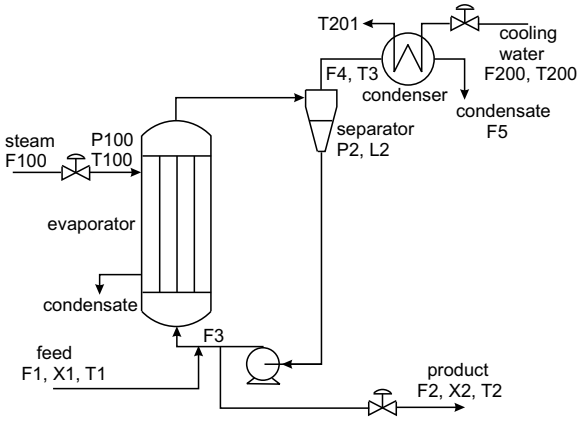

Fig. 2. Evaporator System

circulation" evaporator, where the concentration of dilute liquor is increased by evaporating solvent from the feed stream through a vertical heat exchanger with circulated liquor. The process variables are listed in Table 1 and model equations are given in Appendix A.

Table 1. Variables and Optimal Values

Var.	Description	Value	Units
F_1	Feed flowrate	10	kg/mim
F_2	Product flowrate	1.41	kg/mim
F_3	Circulating flowrate	23.05	kg/mim
F_4	Vapor flowrate	8.59	kg/mim
F_5	Condensate flowrate	8.59	kg/mim
X_1	Feed composition	5	%
X_2	Product composition	35.5	%
T_1	Feed temperature	40	^{o}C
T_2	Product temperature	91.22	^{o}C
T_3	Vapor temperature	83.61	^{o}C
L_2	Separator level	1	meter
P_2	Operating pressure	56.42	kPa
F_{100}	Steam flowrate	10.02	kg/mim
T_{100}	Steam temperature	151.52	^{o}C
P_{100}	Steam pressure	400	kPa
Q_{100}	Heat duty	366.63	kW
F_{200}	Cooling water flowrate	230.54	kg/mim
T_{200}	Inlet C.W. temperature	25	^{o}C
T_{201}	Outlet C.W. temperature	45.5	^{o}C
Q_{200}	Condenser duty	330.77	kW

The economic objective is to minimize the operational cost [$/h] related to steam, cooling water and pump work (Heath *et al.*, 2000; Wang and Cameron, 1994):

$$J = 600F_{100} + 0.6F_{200} + 1.009(F_2 + F_3) \quad (9)$$

The process has the following constraints related to product specification, safety and design limits:

$$X_2 \geq 35 + 0.5\% \quad (10)$$

$$40 \text{ kPa} \leq P_2 \leq 80 \text{ kPa} \quad (11)$$

$$P_{100} \leq 400 \text{ kPa} \quad (12)$$

$$F_{200} \leq 400 \text{ kg/min} \quad (13)$$

$$0 \text{ kg/min} \leq F_3 \leq 100 \text{ kg/min} \quad (14)$$

Note a 0.5% back-off has been enforced on X_2 to ensure the variable remaining feasible for all possible disturbances. The process model has three state variables, L_2, X_2 and P_2 with eight degrees of freedom. Four of them are disturbances, F_1, X_1, T_1 and T_{200}. The rest four degrees of freedom are manipulable variables, F_2, P_{100}, F_3 and F_{200}. The optimization problem of (9) with process constraints, (10) to (14) has been solved under nominal disturbances:

$$d = \begin{pmatrix} F_1 & X_1 & T_1 & T_{200} \end{pmatrix}^T = \begin{pmatrix} 10 & 5 & 40 & 25 \end{pmatrix}^T \quad (15)$$

The minimum cost obtained is 6178.2 $/h and corresponding values of process variables are shown in Table 1.

At the optimal point, there are two active process constraints, $X_2 = 35.5\%$ and $P_{100} = 400$ [kPa]. These two constraints will keep active within whole disturbance region, which is defined as $\pm 20\%$ of the nominal disturbances. Physically, the first active constraint is because a higher outlet composition requires more solvent to be evaporated, therefore needs more steam, cooling water and pump cost. For the second constraint, since heater duty, Q_{100} is determined by both steam pressure, P_{100} and circulating flowrate, F_3, reducing P_{100} will increase F_3 due to energy balance. However, the sensitivity of P_{100} to steam cost is much lower than that of F_3. Hence, an optimal operation should keep X_2 at its lower bound and P_{100} at its higher bound.

These two active constraints plus the separator level, which has no steady-state effect on the plant operation, but must be stabilized at its nominal setpoint, consume three degrees of freedom. Therefore, the first-order optimal condition has one degree of freedom. Choose cooling water flowrate, F_{200} as v and rest manipulated variables and state variables as z, i.e.

$$z = \begin{pmatrix} L_2 & X_2 & P_2 & F_2 & P_{100} & F_3 \end{pmatrix}^T$$

By using (7), the following gradient function is obtained (see Appendix B):

$$G = 0.6 - 0.5538\frac{T_{201} - T_{200}}{F_{200}} \times \left(6.306\frac{0.16(F_1 + F_3) + 0.07F_1}{T_{100} - T_2} + \frac{42F_1}{36.6}\right) \quad (16)$$

4.2 Self-optimizing control structure

Among process constraints lised from (10) to (14), X_2 and P_{100} are actively constrained; F_{200} and F_3 are manipulated variables. Therefore, only P_2 needs to be controlled in a cascade structure illustrated in Figure 1. The primarily controlled variables should include variables for stabilization, variable for active constraint control and the gradient function for self-optimizing control. Since one of the active constraints, P_{100} is a manipulable variable, it is kept at its maximum value in an open loop. Thus the control structure has three primary measurements, L_2, X_2 and the gradient, G of (16), one secondary measurement, P_2 and three manipulated variables, F_2, F_3 and F_{200}.

The next step is to pair these inputs and outputs to construct a decentralized control scheme. The evaporation process has been considered by many researchers for decentralized control since the model was published. Since the process dynamic model has an integrator, the steady-state gain matrix cannot directly be calculated. This blocked people to use the RGA to select input/output pairing. Most researchers including the original authors have made their decisions based on the following engineering heuristic judgement, using product flowrate, F_2 to control separator level, L_2 (Newell and Lee, 1989; Heath *et al.*, 2000; Kookos and Perkins, 2001; Kookos and Perkins, 2002). This judgement was also confirmed by applying pole direction analysis (Govatsmark and Skogestad, 2001). However, this decision will result in

a decentralized control system, which has strong interactions and requires a big back-off for the product composition. The big back-off will increase operating costs significantly.

Actually, input-output pairing of an integrating process can still be determined using the RGA index with a specially determined steady-state gain matrix (McAvoy, 1998). In this work, it is determined via directly model analysis. From model equations, it can be identified that F_2 is the only manipulated variable which can affect X_2 if all other loops are open, i.e. the model represents a semi-decentralized system without any decoupling compensator (Wang and Cameron, 1994). Therefore, it is a natural choice to pair F_2 with X_2. Since other loops will not impose any interactions to this loop, the composition can be tightly controlled so that the back-off of X_2 can be significantly reduced. The rest loops are configure as follows: L_2 controlled by F_3, gradient function G controlled by the setpoint of P_2, which is in turn controlled by F_{200} via the cascade structure of Figure 1. It can be shown from the open-loop transfer function of the process that this configuration results in a near-triangular system, i.e. a system has only one-way interactions. Therefore, it is relatively easy to tune loop controllers to get satisfied performance. It is worth to point out that this configuration can also be obtained by applying the techniques proposed by Samyudia *et al.* (1995).

5. SIMULATION RESULTS AND COMPARISON

The above decentralized and cascade self-optimizing control scheme is implemented with PI controllers in a MATLAB/Simulink model. The parameters of four PI controllers are shown in Table 2. In

Table 2. PI controller parameters

Loop	Gain	Integral time [min]
(L_2,F_3)	200	5
(X_2,F_2)	36.74	4.6619
(P_2,F_{200})	200	6.667
(G,P_2 Setpoint)	1000	2000

the simulation all disturbances are modelled as a step signal passing through a first-order delay. The amplitudes of step changes are randomly produced within the ±20% range of the nominal values. The changing intervals and time constants of the first-order delays are shown in Table 3. With

Table 3. Disturbance model parameters

Disturbance	Interval [min]	Time constant [min]
F_1	120	20
X_1	6	2
T_1	15	5
T_{200}	15	5

Table 4. Alternative configurations and operating costs

Structure	Self-opt. loop	Cost I [$]	Cost II [$]
FF	F_{200}/F_1	120,918	120,627
F	F_{200}	120,947	120,656
G	T_{201}	120,952	120,660
C	P_2	121,653	121,356
B	F_3	122,800	121,596
This work	Gradient	120,916	120,625

the above configuration, simulation results of a 20-hour operation are shown in Figure 3. The total operating cost is $120,916. From Figure 3 (c) and (i) it can be seen that around hour 4 and hour 19:30 there are two periods where the operating pressure constraint becomes active. During these periods, the gradient function presents big offsets, i.e. the process has to sacrifice some cost to ensure operating safety.

It also can be observed from Figure 3 (b) that the product composition X_2 has been tightly controlled within 35.48% to 35.52% range. This is much better than expected at design stage. Therefore, a further reduction on the back-off value is achievable. That means a further reduction in the total operating cost.

Govatsmark and Skogestad (2001) have also investigated the self-optimizing control problem for the same evaporation process. They proposed seven most promising self-optimizing control structures. Five of these structures (G, FF, F, C and B) are listed in Table 4. Two structures using T_2 and T_3 for self-optimization are omitted for comparison because they are equivalent to structure C using P_2. These structures are based on that separator level L_2 is controlled by F_2. Thus, these configurations cannot achieve 0.5% back-off of X_2 under the same disturbances as above. For comparison, these structures are modified to use F_2 to control X_2. A modification to use the cascade structure to ensure P_2 within its boundary is also imposed on these structures. With the modifications in place, these configurations plus gradient control are simulated under the same disturbances conditions as above. The total 20-hour operating costs of these structures are shown in Table 4 (Cost I). For comparison, the total costs with reduced composition back-off (0.05%) are also listed (Cost II). The table clearly shows that the gradient control indeed is the best configuration. The costs of other three configurations, FF, F and G are only few dollars more than the optimal operation. Since these variables are normal measurements, in practical system, they are the best alternative choices for self-optimizing control. Structure B is the worst not only in cost bust also in constraint. Simulation results (not shown in the paper) indicate that it cannot maintain P_2 within limits. The relative loss of structure C is about 0.6% whilst the absolute loss is over $800 per day. Therefore, it

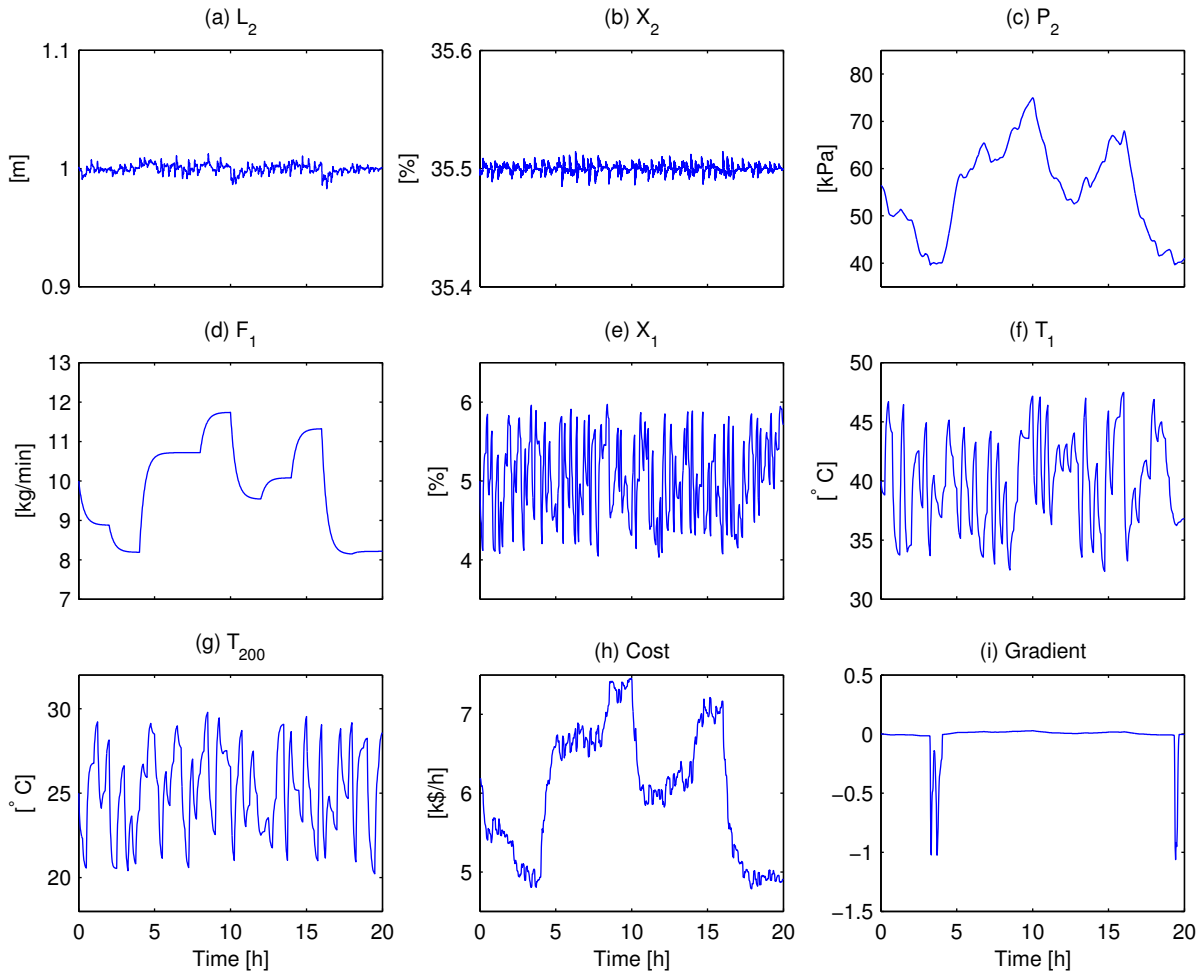

Fig. 3. Self-optimizing control simulation results: (a) Separator level, (b) Product composition, (c) Operating pressure, (d) Feed flowrate, (e) Feed composition, (f) Feed temperature, (g) Cooling water inlet temperature, (h) Operating cost, (i) Gradient function.

is not suitable for self-optimizing control in long-term operating. It is worth to point out that this comparison is based on constraint control on X_2 and P_{100}, *i.e.* all structures are at least partly optimal. This might be the reason why the costs of alternative configurations are close to each other. Even through, the gradient control still provides a benchmark for achievable operating cost for different configurations. The reduction in composition back-off can save about \$350 per day in total operating cost. It is significant in the long-term operating point view. The benefit is due to the input-output pairing configured appropriately.

The optimality of gradient control is also compared with traditional two-layer optimizing control configuration, where the setpoint of P_2 is determined by a steady-state optimizer. The comparison is based on the assumption that disturbances in F_1, X_1, T_1 and T_{200} change randomly within $\pm 20\%$ of their nominal values. All disturbances change simultaneously at every 5 hours. It has been found that it is necessary to pre-filter the step change in the setpoint of P_2 to avoid overshot of P_2 and to reduce the operating cost.

Table 5. Operating cost comparison

Optimal setpoint with filter			Gradient
$\tau = 30$ min	$\tau = 45$ min	$\tau = 60$ min	control
287,756.01	287693.81	287,713.69	287,665.21

The overall operating cost is very sensitive to the time constant of the filter. Fifty-hour operation costs with three different filters are compared with that of gradient control in Table 5. The results in Table 5 show that setpoint filter with time constant of 45 [min] is the best for direct P_2 setpoint change. However, even with this carefully designed filter, the direct setpoint change based on steady-state optimization is still not as good as the gradient control configuration. A dynamic performance comparison between gradient control and optimal P_2 setpoint control with prefilter of 45 min is shown in Figure 4. Figures 4 (a) and (c) show that the main difference between these two configurations are at 0 hour and 40 hour when P_2 decreasing to the lower bound, where P_2 has faster response with gradient control than with directly P_2 setpoint optimizing control. This comparison reveals that the gradient control is more or less equivalent to traditional two-layer directly

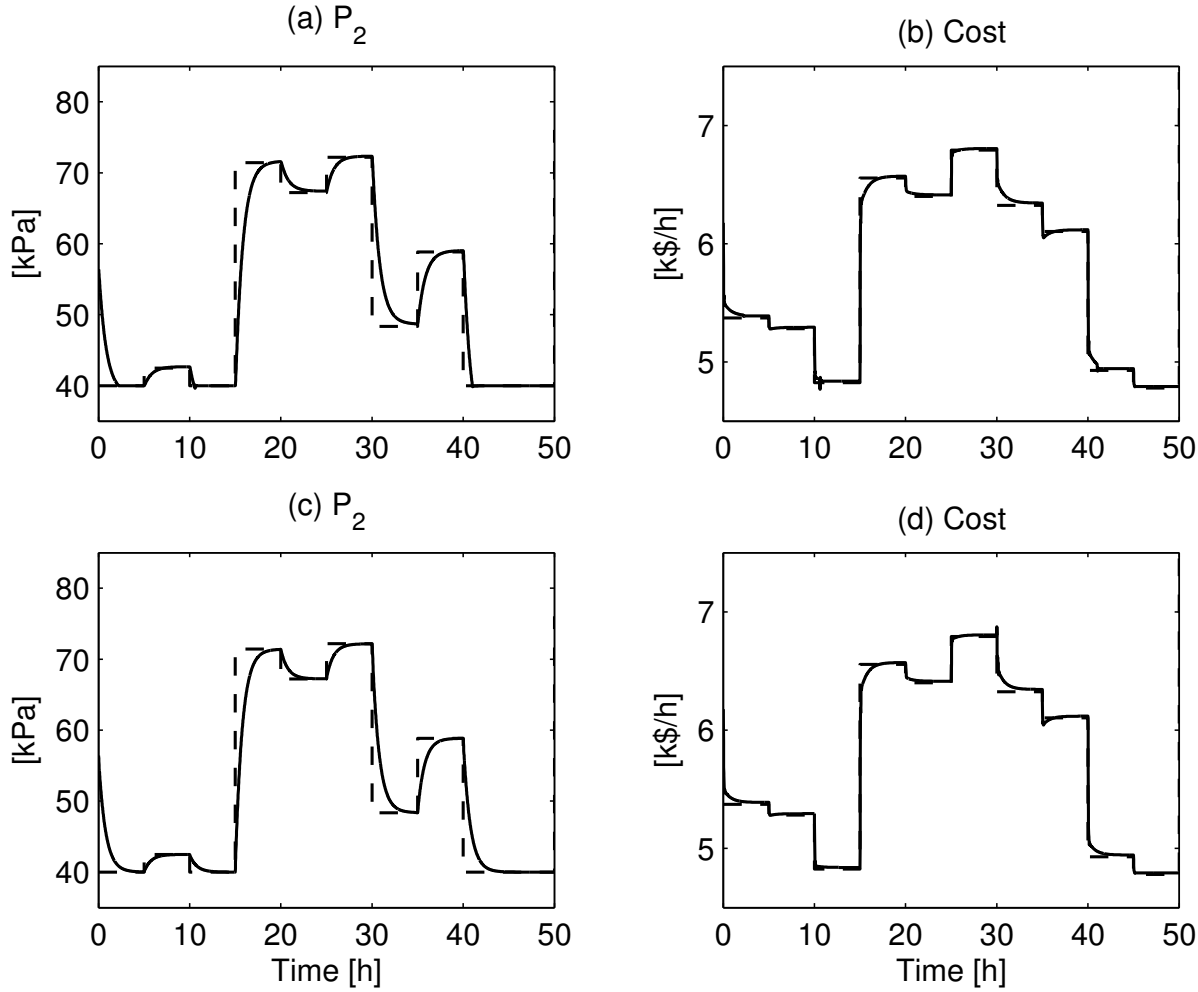

Fig. 4. Dynamic optimality comparison: (a) and (b) gradient control, (c) and (d) directly P_2 setpoint optimizing control. In all figures, solid lines are dynamic simulation results and dashed lines are steady-state optimization results.

setpoint optimizing control with a setpoint prefilter. However, the advantage of gradient control is that it has a varying time constant adaptively adjusted according to disturbances and current process states. This feature makes the gradient control is more promising than traditional two-layer directly setpoint optimizing control.

6. CONCLUSIONS

The concept of self-optimizing control has been scrutinized. Using the chain rule of differentiation, an explicit expression of gradient in terms of system's Jacobian matrices has been derived for the first-order optimal condition of a constrained optimization problem. This gradient function can directly be used as a controlled variable to achieve self-optimization. To cope with conditionally active constraints, a cascade control structure has been proposed. With this structure, the optimal condition and conditionally active constraints can automatically switch each other to be active or inactive depending on disturbances so that both are satisfied. Both ideas have been demonstrated with the evaporator system. For the evaporation process, it is also shown that a traditional engineering judgement for level control structure selection may lead to a wrong decision.

REFERENCES

Govatsmark, M.S. and S. Skogestad (2001). Control structure selection for an evaporation process. In: *Proceedings of ESCAPE'11*. Kolding, Denmark. pp. 657–662.

Griewank, A. (2000). *Evaluating Derivatives*. SIAM. Philadelphia, PA.

Heath, J.A., I.K. Kookos and J.D. Perkins (2000). Process control structure selection based on economics. *AICHE J.* **46**(10), 1998–2016.

Kookos, I.K. and J.D. Perkins (2001). An algorithm for simultanous process design and control. *Ind. Eng. Chem. Res.* **40**, 4079–4088.

Kookos, I.K. and J.D. Perkins (2002). An algorithmic method for the selection of multivariable process control structures. *J. Process Control* **12**, 85–99.

Larsson, T., K. Hesterun, E. Hovland and S. Skogestad (2001). Self-optimization control of a large-scale plant: The tennessee eastman process. *Ind. Eng. Chem. Res.* **40**, 4889–4901.

McAvoy, T.J. (1998). A methodology for screening
level control structures in plantwide control systems. *Comput. Chem. Eng.* **22**(11), 1543–1552.

Newell, R.B. and P.L. Lee (1989). *Applied Process Control – A Case Study*. Prentice Hall. Englewood Cliffs, NJ.

Samyudia, Y., P.L. Lee, I.T. Cameron and M. Green (1995). A new approach to decentralized control design. *Chem. Eng. Sci.* **50**(11), 1695–1706.

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

Wang, F.Y. and I.T. Cameron (1994). Control studies on a model evaporation process – constrainted state driving with conventional and higher relative degree systems. *J. Process Control* **4**(2), 59–75.

Appendix A. MODEL EQUATIONS

$$\frac{\mathrm{d}L_2}{\mathrm{d}t} = \frac{F_1 - F_4 - F_2}{20} \tag{A.1}$$

$$\frac{\mathrm{d}X_2}{\mathrm{d}t} = \frac{F_1 X_1 - F_2 X_2}{20} \tag{A.2}$$

$$\frac{\mathrm{d}P_2}{\mathrm{d}t} = \frac{F_4 - F_5}{4} \tag{A.3}$$

$$T_2 = 0.5616 P_2 + 0.3126 X_2 + 48.43 \tag{A.4}$$

$$T_3 = 0.507 P_2 + 55.0 \tag{A.5}$$

$$F_4 = \frac{Q_{100} - 0.07 F_1 (T_2 - T_1)}{38.5} \tag{A.6}$$

$$T_{100} = 0.1538 P_{100} + 90.0 \tag{A.7}$$

$$Q_{100} = 0.16 (F_1 + F_3)(T_{100} - T_2) \tag{A.8}$$

$$F_{100} = Q_{100}/36.6 \tag{A.9}$$

$$Q_{200} = \frac{0.9576 F_{200}(T_3 - T_{200})}{0.14 F_{200} + 6.84} \tag{A.10}$$

$$T_{201} = T_{200} + \frac{13.68 (T_3 - T_{200})}{0.14 F_{200} + 6.84} \tag{A.11}$$

$$F_5 = \frac{Q_{200}}{38.5} \tag{A.12}$$

Appendix B. GRADIENT FUNCTION

The gradient is derived based on that L_2 and X_2 are perfectly controlled. Therefore, equations (A.1)and (A.2) are simplified as follows:

$$F_2 = F_1 \frac{X_1}{X_2} \tag{B.1}$$

$$F_4 = F_1 \left(1 - \frac{X_1}{X_2}\right) \tag{B.2}$$

Linking (B.2) with (A.6) gives

$$\frac{\partial F_3}{\partial P2} = \frac{0.5616[0.16(F_1 + F_3) + 0.07 F_1]}{0.16 (T_{100} - T_2)}$$

The derivative of F_{100} to P_2 is determined from (A.9) and (A.8):

$$\frac{36.6}{0.16} \frac{\partial F_{100}}{\partial P_2} = \frac{\partial F_3}{\partial P_2}(T_{100} - T_2) - 0.5616(F_1 + F_3)$$

From equations (A.10) and (A.12), it can be derived that

$$\frac{\partial F_5}{\partial P_2} = \frac{(0.9576 \times 0.507) F_{200}}{38.5(0.14 F_{200} + 6.84)} \tag{B.3}$$

$$\frac{\partial F_5}{\partial F_{200}} = \frac{(0.9576 \times 6.84)(T_3 - T_{200})}{38.5(0.14 F_{200} + 6.84)^2} = \frac{(0.9576 \times 0.5)(T_{201} - T_{200})}{38.5(0.14 F_{200} + 6.84)} \tag{B.4}$$

Hence, the partial derivatives of the cost function given in (9) are:

$$\frac{\partial J}{\partial F_{200}} = 0.6$$

$$\frac{\partial J}{\partial P_2} = 600 \frac{\partial F_{100}}{\partial P_2} + 1.009 \frac{\partial F_3}{\partial P_2} = 0.5616 \times \left(6.306 \frac{0.16(F_1 + F_3) + 0.07 F_1}{T_{100} - T_2} + \frac{42 F_1}{36.6}\right)$$

Denote $f = F_4 - F_5$ in (A.3). Then

$$\frac{\partial f}{\partial P_2} = -\frac{\partial F_5}{\partial P_2}$$

$$\frac{\partial f}{\partial F_{200}} = -\frac{\partial F_5}{\partial F_{200}}$$

Combining with (B.3) and (B.4) gives

$$\frac{\partial f}{\partial F_{200}} \left(\frac{\partial f}{\partial P_2}\right)^{-1} = 0.9862 \frac{T_{201} - T200}{F_{200}} \tag{B.5}$$

To select a variable from P_2 and F_{200} to derive the gradient of J, the magnitude of the above ratio is to be checked. For simplicity, the check is only done at the nominal operating point, where $0.9862(T_{201} - T_{200})/F_{200} = 0.0877 < 1$. Hence, choosing F_{200} as independent variable leads to a flatter gradient, therefore is less sensitive to control error than choosing P_2.

Choosing F_{200} as the independent variable and using (7), the gradient function can be calculated as

$$G = \frac{\partial J}{\partial F_{200}} - \frac{\partial f}{\partial F_{200}} \left(\frac{\partial f}{\partial P_2}\right)^{-1} \frac{\partial J}{\partial P_2} \tag{B.6}$$

Inserting all relevant partial derivatives into the above equation gives the gradient function in (16).