Self-Optimizing Control in Chemical Recycle Processes

Julian Straus * Sigurd Skogestad *

* Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: {julian.straus; sigurd.skogestad}@ntnu.no)

Abstract: An engineer always has to make assumptions about the system boundary. In this paper, the impact of neglected dependencies of manipulated variables on the disturbance variables as example for said assumptions is investigated in the context of self-optimizing control. The feedback through dependent disturbances influences both the optimal operating point and the combination of measurements. As a case study, we consider an ammonia synthesis reactor with a simplified model for the ammonia separation and the recycle. The disturbance dependency changes the optimal selection matrices through the recycle. However, we find that it is possible to neglect the recycle in the selection of the controlled variables for this example if the setpoint is adjusted.

© 2018, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Self-optimizing control, Integrated systems, Feedback through recycles, Control of ammonia reactors.

1. INTRODUCTION

Control structure design, and in particular what to control, is important for both economic performance and stabilization of a process. The aim of a good control structure is to operate at the economic optimum while satisfying safety and environmental constraints in the presence of disturbances. In order to achieve economic optimal operation, different methods can be utilized. One can generally distinguish between online and offline optimization methods. Examples for online optimization include real time optimization (RTO) and economic nonlinear model predictive control (E-NMPC), whereas offline methods include self-optimizing control (SOC).

The aim of this paper is to investigate, how the dependency of disturbances may influence the theoretical performance of a self-optimizing control structure. Section 2 recapitulates SOC with focus on the applied exact local method (Halvorsen et al., 2003), whereas Section 3 looks into the effect of dependent disturbances in the calculation of the optimal selection matrix \( \mathbf{H} \). Section 4 investigates the influence of the feedback on a case study representing an ammonia reactor with a simplified recycle loop.

2. SELF-OPTIMIZING CONTROL

Self-optimizing control (SOC) is the selection of controlled variables \( \mathbf{c} \) which when kept constant in the case of a disturbance, result in an acceptable economic loss (Skogestad, 2000). The starting point is a steady-state optimization problem given by

\[
\min_{\mathbf{x}, \mathbf{u}} J(\mathbf{x}, \mathbf{d}, \mathbf{u})
\]

s.t. \( 0 = \mathbf{g}(\mathbf{x}, \mathbf{d}, \mathbf{u}) \)

\( 0 \geq \mathbf{h}(\mathbf{x}, \mathbf{d}, \mathbf{u}) \) \hspace{1cm} (1)

in which \( \mathbf{x} \in \mathbb{R}^{n_x} \) denote the state variables, \( \mathbf{d} \in \mathbb{R}^{n_d} \) the disturbance variables, and \( \mathbf{u} \in \mathbb{R}^{n_u} \) the steady-state degrees of freedom. The process model itself is given by \( \mathbf{g} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_d} \times \mathbb{R}^{n_b} \to \mathbb{R}^{n_y} \) whereas \( \mathbf{h} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_d} \times \mathbb{R}^{n_b} \to \mathbb{R}^{n_h} \) denote the operational constraints given by the process. The cost function \( J : \mathbb{R}^{n_x} \times \mathbb{R}^{n_d} \times \mathbb{R}^{n_b} \to \mathbb{R} \) describes an economic cost of the system.

For given disturbances \( \mathbf{d} \), we assume that there exists an input \( \mathbf{u}_{\text{opt}}(\mathbf{d}) \) which minimizes the optimization problem (1). If different values than the optimal input \( \mathbf{u}_{\text{opt}} \) are chosen for the manipulated variables \( \mathbf{u} \), there will be a steady-state loss

\[
L = J(\mathbf{u}, \mathbf{d}) - J(\mathbf{u}_{\text{opt}}(\mathbf{d}), \mathbf{d})
\]
The aim of self-optimizing control is then to find controlled variables \( c \) which when kept constant give a \( u \) that minimize this loss for expected disturbances.

One direct solution to self-optimizing control is to control the gradient of the cost function \( J \) with respect to the inputs \( u \) \((J_u)\) to 0 as this would imply that the cost function is always at an extremum. The corresponding model-free approach of controlling the measured gradient to zero is called extremum seeking control and dates back to 1922 (Tan et al., 2010). However, in general, the gradient cannot be measured. In certain cases it is possible to express the gradient of the cost function as a direct function of the measurements and control it to 0 (Jäschke and Skogestad, 2014).

As it is frequently not possible to obtain the gradient of the cost function as a simple expression of the measurements \( y \in \mathbb{R}^{n_y} \)

\[
y = h_y(x, d, u)
\]

it is necessary to define the controlled variables \( c \) as a function of the available measurements as

\[
c = h_c(y)
\]

in which \( h_c : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_c} \) may be a function of any type. Frequently, linear measurement combinations are used resulting in

\[
c = Hy
\]

in which \( H \in \mathbb{R}^{n_x \times n_y} \).

### 2.1 Linearization of the process model and cost function

The majority of the self-optimizing control methods are based on a local analysis at the nominal optimal operation point. This results in a linearization of the measurements

\[
y = G^y u + G^y_d d
\]

where \( G^y \in \mathbb{R}^{n_y \times n_u} \) and \( G^y_d \in \mathbb{R}^{n_y \times n_d} \) are the process and disturbance gain matrices, respectively. The cost is approximated through a second order Taylor expansion around the nominal operation point \((u^*, d^*)\)

\[
J(u, d) = J(u^*, d^*) + \left[ J_u \right]^T \left[ \Delta d \right] + \frac{1}{2} \left[ \left[ \Delta d \right]^T \left[ J_{uu} J_{ud} J_{dd} \right] \left[ \Delta d \right] \right]
\]

with \( \Delta d = d - d^* \) and \( \Delta u = u - u^* \). Note, that the derivatives \( J_u, J_{dd}, J_{uu}, J_{ud} \) and \( J_{dd} \) are evaluated at the nominal point \((u^*, d^*)\). Combining (2) with (7) and utilizing that \( J_u = 0 \) at the optimum, we can calculate the loss for disturbances \( d = d^* \) as

\[
L = \frac{1}{2} \left( u - u_{opt}(d) \right)^T J_{uu} \left( u - u_{opt}(d) \right)
\]

### 2.2 Calculation of the selection matrix \( H \)

Several methods exist to obtain optimal measurement combinations, \( c = Hy \). The reader is referred to (Jäschke et al., 2017) for a concise review of the different methods which can be utilized. In this study, the exact local method as developed by Halvorsen et al. (2003) and simplified by Yelchuru and Skogestad (2012) is utilized. In order to make a statement about the loss, Halvorsen et al. (2003) introduced diagonal scaling matrices for the disturbances \( W_d \) and measurement errors \( W_{n_y} \) as

\[
\Delta d = W_d d'; \quad n_y = W_{n_y} n_y
\]

in which the vectors \( d' \) and \( n_y' \) are assumed to satisfy

\[
\left\| \left[ \begin{array}{c} d' \\ n_y' \end{array} \right] \right\|_2 \leq 1
\]

For a given selection matrix \( H \), the linearized model (6), and the general loss expression (8), it is possible to derive the worst-case loss (Halvorsen et al., 2003) and the average expected loss (Kariwala et al., 2008) as

\[
L_{WC}(H) = \frac{1}{2} \sigma (M)^2
\]

\[
L_{avg}(H) = \frac{1}{2} \|M\|^2_F
\]

in which the loss matrix \( M \) is shown to be

\[
M = J_{uu}^{1/2} (HG_y)^{-1} HY
\]

The optimal sensitivity matrix for the measurements \( F \) can be obtained numerically or calculated from the linearized model (Halvorsen et al., 2003)

\[
F = \frac{\partial y_{opt}}{\partial d} = - (G^y J_{uu}^{-1} J_{ud} - G^y_d)
\]

The optimal measurement combination \( H \) can now be calculated as the solution which minimizes the average (12) and worst case (11). Both these optimization problems have the same optimal solution (Kariwala et al., 2008) which can be obtained by solving

\[
\min_H \| J_{uu}^{1/2} (HG_y)^{-1} HY \|_F
\]

The analytical solution to this problem was first described by Alstad et al. (2009) and later simplified by Yelchuru and Skogestad (2012) to

\[
H^T = (YY^T)^{-1} G^y
\]

From (18) and (14), we can see that the required model information is \( G^y \) and \( F \), where the latter can be calculated using (16). In practice, if a nonlinear process model is utilized, it is simpler to calculate \( F \) numerically from (15). Similarly, the loss \( L \) can be calculated using the nonlinear model and optimization problem (1).

### 3. DEPENDENT DISTURBANCES

Consider the block diagram in Figure 1, where “Local plant” represents our submodel (ammonia reactor in our case study) and “Remaining plant” represents the neglected part of the process (the recycle in our case).

The first question now is: Assume that we optimize our “Local plant” with a fixed value of \( d_0 \), that is, we neglect the effect \( u \) has on \( d_0 \) through, for example, the recycle. Is this acceptable? Of course, the answer is generally no.

The second question is: Assume now that we find controlled variables that is, find \( H_0 \) based on considering our “Local plant”. Is this acceptable? Again, the answer is generally no, but in practice the answer may be “yes” if the local cost function is the same as the overall one. To
better understand this, let us consider how the matrices used to find $\mathbf{H}$ in (13) and (14) may change.

To see the difference between $\mathbf{G}^\gamma$ (based on the overall plant) and $\mathbf{G}_0^\gamma$ (based on the local plant), we can look at the total differential,

$$\mathbf{G}^\gamma \triangleq \left( \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \right)_d \quad = \quad \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \bigg|_d + \frac{\partial \mathbf{y}}{\partial \mathbf{u}} \bigg|_{\mathbf{u}_d} \frac{\partial \mathbf{u}_d}{\partial \mathbf{d}} \frac{\partial \mathbf{y}_d}{\partial \mathbf{d}} \quad \frac{\partial \mathbf{d}}{\partial \mathbf{u}} \bigg|_{\mathbf{d}_0} \frac{\partial \mathbf{y}_d}{\partial \mathbf{d}} \quad \frac{\partial \mathbf{d}}{\partial \mathbf{d}} \bigg|_{\mathbf{d}_0} \frac{\partial \mathbf{y}_d}{\partial \mathbf{d}} \quad (19)$$

with $\mathbf{y}_d$ corresponding to the outlet variables of the local plant which affect the neglected part, see Figure 1. In our case, these are the outlet flow, pressure, temperature, and composition. The gain $\mathbf{G}_0^\gamma$ is the previously neglected feedback and can be obtained from the submodel of the remaining plant. The gain $\mathbf{G}_d^\gamma$ corresponds to the change in the outlet variables with changing input.

A similar analysis can be conducted for the Hessian of the cost function ($\mathbf{J}_{\mathbf{uu}}$) and the disturbance gain $\mathbf{G}_d^\gamma$.

4. CASE STUDY - AMMONIA SYNTHESIS LOOP

The core of the case study is a three-bed ammonia reactor previously described by Morud and Skogestad (1998) and utilized by Straus and Skogestad (2017) in the application of economic nonlinear model predictive control. In this model, the disturbances ($\mathbf{d}_0$) are the inlet variables to the system

$$\mathbf{d}_0 = [\dot{m}_{\text{Feed}_0}, p_{\text{Feed}_0}, T_{\text{Feed}_0}, w_{\text{NH}_3,\text{Feed}_0}]$$

There exist 3 input variables ($\mathbf{u}$), which correspond to the split ratios to the three reactor beds. The cost function for the ammonia reactor is to maximize the extent of reaction $\xi$, i.e.

$$\mathbf{J} = -\xi = -\dot{m}_{\text{Feed}_0} (w_{\text{NH}_3,\text{Rea}} - w_{\text{NH}_3,\text{Feed}_0}) \quad (21)$$

As the reaction is limited by the thermodynamical equilibrium, a recycle is necessary to utilize the unreacted hydrogen and nitrogen. The reactor is connected to the recycle through the inlet stream $\mathbf{d}_0$ and the outlet stream $\mathbf{y}_d$. This recycle stream ($\mathbf{d}_r$) corresponds to 75% of the mass of the feed to the reactor. Hence, the impact of the dependency of the neglected remaining plant is expected to be large in this case study.
Additionally, a model equation similar to a valve coefficient has to be added for the pressure drop after the separator
\[ 0 = \dot{n}_{\text{Sep}} - k \sqrt{p_{\text{Feed}} - p_{\text{Sep}}} \]  
with a given pressure drop coefficient \( k \) (kmol/(s√bar)). The compressor duty of an isothermal compressor is (e.g. Skogestad 2008)
\[ W = \frac{\dot{n}_{\text{Sep}} RT_{\text{Feed}}}{\eta} \ln \left( \frac{p_{\text{Feed}}}{p_{\text{Sep}}} \right) \]  
As there is no purge flow and the product is pure ammonia, all of the feed has to be converted. The system will therefore operate with a constant extent of reaction, and hence, it cannot be used anymore as cost function as it was the case in the local reactor system. Instead, the new economic cost function corresponds to minimizing the compressor duty of the recycle loop, i.e.
\[ J = W \]  
As mentioned beforehand, this change in cost function may affect SOC variables defined for the reactor system. The new cost function aims at minimizing the flow within the recycle. This corresponds to minimizing the feed flow to the reactor while maintaining a constant extent of reaction. This implies that the actual optimal operation point with recycle does not fulfill requirement (10). This can be partly explained by an increase in the process fraction \( w \) 3.

### 4.2 Application of SOC

This brings us to the second question in Section 3: is it possible to optimize the reactor neglecting the recycle? With the new cost function and the modified system, the optimal nominal inlet conditions of the reactor are given in Table 1. Unsurprisingly, it is not possible to neglect the recycle in the optimization. Especially the reactor inlet mass flow \( \dot{m}_{\text{Feed0}} \) changes a lot due to the recycle. This is caused by a positive feedback. A higher conversion per pass corresponds to more ammonia separated, and hence, a lower recycle flowrate. This in turn increases the residence time in the beds and hence increases the conversion per pass. The ammonia mass fraction experiences negative feedback due to the assumption of a constant split factor. Hence, its value changes only by a small value. Due to the aforementioned assumptions, the inlet pressure and temperature of the system are the same with and without the recycle stream.

### Table 1. Nominal (optimal) inlet conditions for the reactor

<table>
<thead>
<tr>
<th>Recycle</th>
<th>( \dot{m}_{\text{Feed0}} ) [kg/s]</th>
<th>( p_{\text{Feed0}} ) [bar]</th>
<th>( T_{\text{Feed0}} ) [°C]</th>
<th>( w_{\text{NH}_3, \text{Feed0}} ) [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without</td>
<td>70.0</td>
<td>200</td>
<td>250</td>
<td>8.0</td>
</tr>
<tr>
<td>With</td>
<td>61.8</td>
<td>200</td>
<td>250</td>
<td>8.3</td>
</tr>
</tbody>
</table>

The changes in the optimal sensitivity matrices \( F_i \) (not shown) are even more pronounced, especially for the two disturbances with different values in the nominal optimal case; the inlet flowrate \( \dot{m}_{\text{Feed0}} \) \((\dot{m}_{\text{Feed}})\) and the inlet mass fraction \( w_{\text{NH}_3, \text{Feed0}} \) \((w_{\text{NH}_3, \text{Feed}})\). Based on these findings, it can be concluded that the linearization (not surprisingly) loses validity through the introduction of a recycle stream. This can be caused by the change in the optimal inlet flowrate of the reactor as shown in Table 1. \( \dot{m}_{\text{Feed0}} \) is reduced by 12 % and should be outside the linear range of the nonlinear model.
Fig. 3. Loss as a function of the disturbance for both cases (H₀ and H). The setpoints for the local selection matrices H₁,₀ are not adjusted to optimal setpoints of the global recycle system.

Adjusting the reactor inlet in the model without recycle to the new optimal value allows the verification of this claim. The plant gains Gᵢ are in this situation similar to the ones with recycle. Furthermore, Jₐᵤ is similar except for a scalar multiplier. This can be explained by the total differential (19). In the case of the ammonia reactor with a maximized extent of reaction, wNH₃,Rea is maximized. In addition, T_Rea is maximized as well whereas p_Rea and m_Rea are unaffected due to mass conservation and the assumption of constant feed pressure to the reactor. Hence, in our case G₀ = 0 and

\[ G^Y = G^Y_0 = 0 \]  
(33)

This special sensitivity occurs, if the outlet variables are equivalent to the cost function.

The optimal sensitivity matrices change however due the neglected dependency of d₀ on yₐ₀ and hence through changes in Gₐ and Jₐᵤ. This explains the changes in the selection matrices H₁, see (30) and (31).

4.3 Loss calculation

In order to evaluate the performance of both CV selections, H₁,₀ in (30) and H₁ in (31), the loss as defined in (2) with the cost function (26) and the (nonlinear) model including the recycle was calculated. The setpoints for the controller in the problem without recycle were given by the optimal setpoints without recycle. The comparison of both losses is shown in Figure 3. As can be seen from the red curves in Figure 3, there is a loss even at the nominal point. This is not necessarily caused by a poor H₀ matrix, but by a non-optimal operating point.

Hence, the setpoint for the SOC variables should be adjusted to the new nominal optimum in which the recycle is considered. The new loss calculations are shown in Figure 4. It is interesting to note, that the differences are surprisingly small. For an inlet pressure disturbance and mass flow disturbance, the loss is smaller for H, whereas the loss is higher for H than for H₀ for an inlet pressure and ammonia mass fraction disturbance.

Both H₀ and H use the same weighting matrices (9). As the reactor inlet mass flow \( \dot{m}_{\text{Feed}} \) is varying between 42 kg/s and 84 kg/s for a flowrate disturbance, we can directly see the incorrect weighting of the inlet mass flow. Changing the value of the mass flow disturbance in the weighting matrix to 20 kg/s results in new controlled variables (H₀,₂)

\[ c_{1,0,2} = -0.181 \quad c_{2,0,2} = -0.053 \quad c_{3,0,2} = 0.971 \]  
(34)

which are more similar to (31). The corresponding loss is depicted in Figure 5. We can directly see that the difference in the loss is marginal, especially for \( \dot{m}_{\text{Feed}} \), which had the largest loss in Figure 4. This is not surprising as the the optimal selection matrix H₀,₂ (34) is close to H (31).

4.4 Discussion

It has to be highlighted that in this specific case study, it was possible to define a cost function in the system without recycle which corresponds to the cost function in the system with recycle. This is not necessarily the case for all submodels of recycle systems. If one would consider the case of a detailed separation section, the aim would be to minimize the cooling costs for a given feed. This feed would also represent some of the disturbances to the model. An unconstrained optimal solution would be given by no cooling and hence no separation. Therefore, separation requirements are needed, either on the separated product or through assigning cost values to all connection streams. Hence, the optimal point would be based on these separation requirements. On the other hand, the
Hence, the setpoint for the SOC variables should be $m_i$ in the problem without recycle were given by the optimal the recycle was calculated. The setpoints for the controller $G_{ud}$ are the ones with recycle. Furthermore, $J$ claim. The plant gains $G$ to the new optimal value allows the verification of this assumption of constant feed pressure to the reactor. Hence, $H$.

In order to evaluate the performance of both CV selections, $w_{d}$ differential (19). In the case of the ammonia reactor with $H$ for a scalar multiplier. This can be explained by the total weighting matrix to 20 kg/s results in new controlled. Changing the value of the mass flow disturbance in the direct see the incorrect weighting of the inlet mass flow. $H$ and $H$ ammonia mass fraction disturbance.

Fig. 5. Loss as a function of the disturbance for both cases ($H_{0,2}$ and $H$). The setpoints for local selection matrices $H_{0,2}$ are adjusted to optimal setpoints of the global recycle system and the weighting matrix $W_d$ changed.

From the definition of the loss in (2), it is obvious that there is a constant loss at the nominal operation point if the setpoint for the self-optimizing variables is not adjusted. Recall that the starting point of this investigation is that it is however too complicated to optimize the overall model, and hence, to calculate the true optimal setpoint. Therefore, a model-free approach, e.g. extremum-seeking control or necessary conditions of optimality tracking (François et al., 2005), should be used on top of self-optimizing control for calculating the optimal setpoint.

5. CONCLUSION

The dependency of considered disturbances on the input (and measurements) changes the optimal selection matrix in the application of self-optimizing control. This is the case even if the actual values of the disturbances, and hence, the feed to the submodel are unchanged.

The loss is in the investigated case study similar if the setpoints to the controllers and the disturbance weighting matrix $W_d$ are adjusted. This cannot be generalized and is depending on the neglected dependencies.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support provided by YARA International ASA.

REFERENCES