Self-Optimizing Control of a Large-Scale Plant: The Tennessee Eastman Process

Truls Larsson, Kristin Hestetun, Espen Hovland, and Sigurd Skogestad*

Department of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

This paper addresses the selection of controlled variables, that is, “what should we control”. The concept of self-optimizing control provides a systematic tool for this, and we show how it can be applied to the Tennessee Eastman process, which has a very large number of candidate variables. In this paper, we present a systematic procedure for reducing the number of alternatives. One step is to eliminate variables that, if they had constant setpoints, would result in large losses or infeasibility when there were disturbances (with the remaining degrees of freedom reoptimized). The following controlled variables are recommended for this process: optimally constrained variables, including reactor level (minimum), reactor pressure (maximum), compressor recycle valve (closed), stripper steam valve (closed), and agitator speed (maximum); and unconstrained variables with good self-optimizing properties, including reactor temperature, composition of C in purge, and recycle flow or compressor work. The feasibility of this choice is confirmed by simulations. A common suggestion is to control the composition of inerts. However, this seems to be a poor choice for this process because disturbances or implementation error can cause infeasibility.

1. Introduction

This paper addresses the selection of controlled variables for the Tennessee Eastman process. However, the main objective of the paper is to demonstrate how to select controlled variables for a large-scale process, so the paper should be of interest also for readers without a prior knowledge of or interest in the Tennessee Eastman process.

We base the variable selection on the concept of self-optimizing control using steady-state models and steady-state economics. “Self-optimizing control” is when an acceptable (economic) loss can be achieved using constant setpoints for the controlled variables, without the need to reoptimize when disturbances occur. The constant-setpoint policy is simple, but it will not be optimal (and thus have a positive loss) as a result of the following two factors: (1) disturbances, i.e., changes in (independent) variables and parameters that cause the optimal setpoints to change, and (2) implementation errors, i.e., differences between the setpoints and the actual values of the controlled variables (e.g., because of measurement errors or poor control).

The effect of these factors (the loss) depends on the choice of controlled variables, and the objective is to find a set of controlled variables for which the loss is acceptable.

Downs and Vogel introduced the Tennessee Eastman challenge problem at an AIChE meeting in 1990. The purpose was to supply the academic community with a problem that contained many of the challenges that people in industry meet. The process has eight components, including four reactants (A, C, D, and E), two products (G and H), an inert (B), and a byproduct (F).

The reactions are:

\[ A(g) + C(g) + D(g) \rightarrow G(l) \text{(product)} \]
\[ A(g) + C(g) + E(g) \rightarrow H(l) \text{(product)} \]
\[ A(g) + E(g) \rightarrow F(l) \text{(byproduct)} \]
\[ 3D(g) \rightarrow 2F(l) \text{(byproduct)} \]

The process has four feed streams (A, D, E, and C), one product stream, and one purge stream. Almost all of the inert (B) enters in the largest feed (C), which actually contains almost 50% of component A. The process has five major units: a reactor, a product condenser, a vapor-liquid separator, a recycle compressor, and a product stripper (see Figure 1). There are 41 measurements and 12 manipulated variables. Detailed data are given in Downs and Vogel. We have based our simulation on the model available at the home page of Ricker. We study here the optimal operation of the base case (mode 1) with a given 50/50 product ratio between components G and H and a given production rate.

This plant has been studied by many authors, and it has been important for the development of plantwide control as a field. Many authors have used it to demonstrate their procedures for the design of control systems. We consider here the selection of controlled variables.

McAvoy and Ye proposed to control reactor temperature, reactor pressure, recycle flow rate, compressor work, concentration of B (inert) in purge, and concentration of E in product flow. Ricker and Lee tested this strategy, finding that the compressor power loop often saturated during transients.

Lyman and Georgakis recommended a control structure involving control of the variables reactor temperature; reactor level, recycle flow rate; agitation rate; compositions of A, D, and E in the reactor feed;
composition of B (inert) in the purge; and composition of E in the product. Even though they considered the operating cost for the control structure, their structure is not economically optimal because some variables that should be kept at their constrained values (such as the compressor recycle valve) are used as manipulated variables.

Ricker\textsuperscript{10} considered the steady-state optimal operation of the plant. In all cases, he found that it is optimal to have the maximum reactor pressure, minimum reactor level, maximum agitator speed, and minimum steam valve opening. Furthermore, in most cases, it is optimal to use the minimum compressor recycle valve opening. Ricker\textsuperscript{10} notes that the controlled variables “must be carefully chosen; arbitrary use of feedback control loops should be avoided”.

Ricker and Lee\textsuperscript{8} used nonlinear model predictive control (NMPC) and compared their approach with the multiloop (decentralized) strategy of McAvoy and Ye,\textsuperscript{7} which they find performs adequately for many scenarios, although they suggest that compressor power should not be controlled. For these simpler cases, the NMPC strategy improves performance, but the difference seems too small to justify the NMPC design effort. On the other hand, for the more difficult cases, the decentralized approach would require multiple overrides to handle all conditions, and nonlinear model predictive control might be preferred.

In another study, Ricker\textsuperscript{11} considered decentralized control and concluded that there is little, if any, advantage in using NMPC on this application. His approach is similar to the one in this paper. First, he chooses to control the variables that optimally should be at their constraints (“active constraint control”). Second, he excludes variables for which the economic optimal value varies significantly. He ends up controlling the compressor recycle valve position (at minimum), the steam valve position (at minimum), the reactor level (at minimum), the reactor temperature, the composition of C in the reactor feed, and the composition of A in the reactor feed. He notes that it is important to determine appropriate setpoint values for the latter three unconstrained variables.

Luyben et al.\textsuperscript{12} set the agitation rate and the recycle valve at their constrained, values and chose to control the reactor pressure, reactor level, separator temperature, stripper temperature, ratio between E and D feed rates, and compositions of A and B (inert) in the purge.

Tyreus\textsuperscript{13} uses a thermodynamic approach to select controlled variables. This can provide useful guidelines but cannot, in general, provide the optimal solution as thermodynamics is independent of cost data. He sets the agitation on full speed, closes the steam valve and the compressor recycle valve, and chooses to control reactor temperature, reactor pressure, reactor level, and compositions of A in the reactor feed and B (inert) in the purge flow.

To summarize, most authors do not control all of the variables that are constrained at the optimum and therefore do not operate optimally in the nominal case. Most control reactor pressure, reactor level, reactor temperature, and composition of B (inert). It is common to control the stripper temperature, separator temperature, and composition of C and/or A in the reactor feed. The main objective of this paper is to search systematically for a set of controlled variables, if such a set exists, that results in self-optimizing control for the Tennessee Eastman process. In particular, the issue is to find a good choice for the three unconstrained variables.

2. Self-Optimizing Control

We give here an introduction to self-optimizing control and refer the reader to the paper of Skogestad\textsuperscript{9} for more
details, including a discussion of the related literature (two more recent references are refs 14 and 15).

Many people do realize that the selection of controlled variables is actually an issue but ask the question: “Why are we controlling hundreds of temperatures, pressures, and compositions in a chemical plant, when there is no specification on most of these variables? Is it just because we can measure them or is there some deeper reason?”

The main basis for control is that the plant has many degrees of freedom that need to be specified during operation, and the “deeper” reason for selecting a particular set of controlled variables is that it provides “self-optimization” when there are disturbances or other changes in the operating point.

The basic idea of self-optimizing control was formulated about 20 years ago by Morari et al.,1 who wrote that “In attempting to synthesize a feedback optimizing control structure, our main objective is to translate the economic objectives into process control objectives. In other words, we want to find a function c of the process variables which when held constant, leads automatically to the optimal adjustments of the manipulated variables, and with it, the optimal operating conditions. […] This means that by keeping the function cu(d) at the setpoint cs through the use of the manipulated variables u, for various disturbances d, it follows uniquely that the process is operating at the optimal steady state.” The ideas of Morari et al.1 were further developed by Skogestad,2–4 who also considered the implementation and studied several problem cases with an unconstrained optimum. Self-optimizing control is defined as follows:

Self-optimizing control is when we can achieve an acceptable loss with constant setpoint values for the controlled variables (without the need to reoptimize when disturbances occur).

To quantify the loss we must define a scalar economic cost function J, for example, of the form

\[ J(u,d) = \int_0^T \phi(u,d) \, dt \]  

(1)

where the independent variables include the degrees of freedom u and the disturbances d. The cost J should be minimized with respect to u subject to satisfying given constraints, including product specifications (e.g., minimum purity), manipulated variable constraints (e.g., nonzero flow), other operational limitations (e.g., maximum temperature), and model equations (equality constraints).

We assume here that the optimization problem is feasible, that is, the constraints are not in conflict such that a solution exists (otherwise, the problem needs to be reformulated). In theory, all of the various objectives for plant operation should be included in the cost J, and minimization of J should then result in the optimal tradeoff between these generally conflicting objectives. In this paper, we assume that the economic performance is primarily determined by steady-state considerations and that the integration in eq 1 can be replaced by time-averaging over the various steady states. The effect of the dynamic control performance can be partly included in the economic analysis by introducing a control error term as part of the implementation error. To achieve truly optimal operation, we would need a perfect model, we would need to measure all disturbances, and we would need to solve the resulting dynamic optimization problem on-line. This is unrealistic, and the question is whether it is possible to find a simpler implementation that still operates satisfactorily. The term self-optimizing control is used when satisfactory economic operation can be achieved with the use of constant setpoints for the controlled variables, that is, without the need for reoptimization when there are disturbances.

To quantify this more precisely, we define the (economic) loss L as the difference between the actual value of the cost function and the truly optimal value, i.e.

\[ L = J(u,d) - J_{\text{opt}}(d) \]

Truly optimal operation corresponds to \( L = 0 \), but in general \( L > 0 \). A small value of the loss function L is desired as it implies that the plant is operating close to its optimum. Self-optimizing control is achieved if a constant-setpoint policy results in an acceptable loss L (without the need to reoptimize when disturbances occur). The main issue here is not to find the optimal setpoints, but rather to find the right variables to keep constant. The precise value of an “acceptable” loss must be selected on the basis of engineering and economic considerations.

The idea of self-optimizing control is illustrated in Figure 2. We see that a loss results when we keep a constant setpoint rather than reoptimizing when a disturbance occurs. For the case in Figure 2, it is better to keep the setpoint \( c_{\text{ss}} \) constant than to keep \( c_{\text{min}} \) constant.

An additional concern with the constant-setpoint strategy is that there is always a difference between the setpoint \( c_s \) and the actual value \( c \) because of implementation errors caused by measurement errors and imperfect control. To minimize the effect of the errors on the operating cost, the cost surface as a function of \( c \) should be as flat as possible. This is illustrated in Figure 3, where we distinguish between three cases in terms of actual implementation: (a) In the case of a constrained optimum, the optimal cost is achieved when one of the variables is at its maximum or minimum (the figure shows the case when the optimum is obtained for \( c = c_{\text{min}} \)). In this case, there is no loss imposed by keeping the variable constant at its “active” constraint. Implementation of an active constraint is usually easy, e.g., it is easy to keep a valve fully open or closed. (b) For the case of an unconstrained flat optimum, the cost
is insensitive to the value of the controlled variable $c$. Finally, for an unconstrained sharp optimum, the cost (operation) is sensitive to the actual value of the controlled variable $c$, and self-optimizing control is not possible. In this case, we would like to find another controlled variable $c$ for which the optimum is flatter.

Thus, the constrained case is easy, because we select as controlled variables the optimally constrained variables. However, it is not at all clear which variables to select in the unconstrained case. Skogestad recommends that a controlled variable $c$ suitable for constant-setpoint control (self-optimizing control) should have the following properties:

**Requirement 1.** The optimal value of $c$ should be insensitive to disturbances, i.e., $c_{\text{opt}}(d)$ depends only weakly on $d$.

**Requirement 2.** The value of $c$ should be sensitive to changes in the manipulated variable $u$, i.e., the gain $G = y/u$ should be large (equivalently, because $\partial^2 G / \partial c^2 = G^{-1} \partial^2 y / \partial u G^{-1}$, the optimum should be "flat" with respect to the variable $c$, i.e., $\partial^2 / \partial c^2$ should be small).

**Requirement 3.** For cases with two or more controlled variables, the selected variables in $c$ should not be closely correlated.

**Requirement 4.** The variable $c$ should be easy to measure and control.

The above requirements might be useful for identifying candidate variables, but the requirements are somewhat qualitative, and checking them might require quite a lot of computation.

By proper variable scaling, the three first requirements can be combined into the single approximate condition of maximizing the minimum singular value of the gain matrix $G$. This condition is computationally attractive, but because it only provides local information, it can be very misleading in some cases (e.g., see Figure 5 where the minimum occurs very close to infeasibility).

A better and more exact approach is therefore to evaluate the cost function for the expected set of disturbances and implementation errors. We apply here the stepwise procedure for self-optimizing control of Skogestad. The main steps are as follows:

**Step 1.** Degree of freedom analysis
**Step 2.** Definition of optimal operation (cost and constraints)
**Step 3.** Identification of important disturbances
**Step 4.** Optimization
**Step 5.** Identification of candidate controlled variables
**Step 6.** Evaluation of loss for alternative combinations of controlled variables (loss imposed by keeping constant setpoints when there are disturbances or implementation errors)
**Step 7.** Final evaluation and selection (including controllability analysis)

Skogestad applied this stepwise procedure to a reactor case and a distillation case, but in both cases, there was only one unconstrained degree of freedom, so the evaluation in step 6 was manageable. However, for the Tennessee Eastman process, there are three unconstrained degrees of freedom at the optimum and a very large number of candidate variables to select from, so it is necessary to apply some more effort in step 5 to reduce the number of alternatives. We present below some general criteria that are useful for eliminating controlled variables.

### 3. Problem Definition

**Step 1.** Degree of Freedom Analysis. The Tennessee Eastman process has 12 manipulated variables, 41 measurements, and 20 disturbances. An analysis (see Table 1) shows that, at steady state, two degrees of freedom are lost because there are two liquid levels with...
Step 2. Definition of Optimal Operation. Downs and Vogel\textsuperscript{5} specified the economic cost \( J \) ($/h) for the process, which is to be minimized. In simple terms

\[
J = (\text{loss of raw materials in purge and products}) + (\text{steam costs}) + (\text{compression costs}) \quad (2)
\]

The first term, related to loss of unreacted raw materials, dominates the cost. All of the manipulated variables have associated constraints, and there are also “output” constraints, including equality constraints on the product quality and product rate.

Step 3. Identification of Important Disturbances. A closer analysis reveals that disturbances 3, 4, 5, and 7 have no steady-state effect on the economics provided that we make appropriate use of the available manipulated variables. For example, disturbance 4 (a step in the reactor cooling water inlet temperature) is easily counteracted by increasing the reactor cooling water flow rate; thus, this disturbance will have no impact on the economics provided that we adjust the cooling rate. Similar arguments can be made for disturbances 3, 5, and 7, provided that we manipulate the reactor coolant flow, separator cooling water flow and the C feed rate. Disturbance 6 (loss of feed A) is considered to be so serious that it should be handled by overrides; therefore, it is not included in this study.

This leaves only the following three disturbances: disturbance 1, change in A/C ratio in the C feedstream; disturbance 2, change in fraction of B (inert) in the C feedstream; and throughput disturbances: Change in production rate by \( \pm 15\% \).

Step 4. Optimization. Ricker\textsuperscript{10} solved the optimization problem using the above cost function, and he provides a good explanation of what happens at the optimum. At the optimum, there are five active constraints (see Table 1), and they need to be controlled to achieve optimal operation (at least nominally). We minimized the cost function \( J \) with respect to the three remaining unconstrained degrees of freedom and obtained the same optimal values as given by Ricker.\textsuperscript{10} The optimal (minimum) operation cost is 114.323 $/h in the nominal case, 111.620 $/h for disturbance 1, and 169.852 $/h for disturbance 2. A continuation method\textsuperscript{16} was used to solve the optimization problem and to generate the cost function surfaces.

The three unconstrained degrees of freedom need to be set during operation, and the main issue, addressed in the next section, is which three variables we should select as controlled variables in a constant-setpoint policy such that we achieve acceptable economic loss (self-optimizing control). We define an acceptable loss as 6 $/h when summed over the four disturbances just mentioned.

4. Candidate Controlled Variables

We are now at Step 5: Identification of Candidate Controlled Variables. This step is the main focus of this paper.

Let us initially not make the assumption that we will satisfy specifications or use active constraint control. We then have 12 degrees of freedom, and we want to select 12 controlled variables that are to be controlled at constant setpoints. We can choose from 41 measurements and 12 manipulated variables, so there are 53 candidate variables. Even in the simplest case, where we do not consider variable combinations (such as differences, ratios, and so on), there are

\[
\begin{align*}
3 & \times 53 & = 159 \\
2 & \times 52 & = 104 \\
1 & \times 50 & = 50 \\
0 & \times 49 & = 0 \\
-1 & \times 48 & = -48 \\
-2 & \times 47 & = -94 \\
-3 & \times 46 & = -138 \\
-4 & \times 45 & = -180 \\
-5 & \times 44 & = -220 \\
-6 & \times 43 & = -264 \\
-7 & \times 42 & = -306 \\
-8 & \times 41 & = -348 \\
-9 & \times 40 & = -390 \\
-10 & \times 39 & = -432 \\
-11 & \times 38 & = -474 \\
-12 & \times 37 & = -516 \\
\end{align*}
\]

possible combinations. It is clearly impossible to evaluate the loss with respect to disturbances and implementation errors for all of these combinations.

The following criteria are proposed to reduce the number of alternatives. Most of them are rather obvious, but nevertheless, we find them useful.

(1) Eliminate variables with no effect on the economics (including variables with no steady-state effect). The value of these variables can be arbitrarily selected, which reduces the number of degrees of freedom and thus the number of controlled variables to be selected. We must, of course, also eliminate the corresponding variables from further consideration as candidate controlled variables.

(2) The variables directly associated with equality constraints should be controlled. (Again, this reduces the number of controlled variables to be selected, and we must also eliminate the corresponding variables from further consideration.)

(3) We choose to control the active constraints. (Again, this reduces the number of controlled variables to be selected, and we must also eliminate the corresponding variables from further consideration.)

(4) Eliminate/group closely related variables

(5) Use process insight to eliminate additional variables

(6) Eliminate single variables that, if they had constant setpoints, would yield infeasibility or large losses when there were disturbances or implementation errors (with the remaining degrees of freedom reoptimized).

(7) Eliminate combinations (pairs, triplets, etc.) of variables that yield infeasibility or large loss

(8) Use local analysis to eliminate variables or variable combinations that result in a small minimum singular value of the appropriately scaled gain matrix G (not used in this paper).

After this, we enter into the final evaluation for the remaining combinations of variables, in which we (1) evaluate disturbance losses and (2) evaluate implementation losses. We now apply these criteria to our case study.

4.1. Eliminate Variables with No Effect on the Economics. There are two variables with no steady-state effect, namely, stripper level and separator level, and their values have no effect on the steady-state economics. This reduces the number of degrees of freedom, and thus controlled variables to be selected, from 12 to 10. The corresponding variables must also be eliminated from further consideration; this eliminates two measurements.

Of course, we need to measure and control the two levels to obtain stable operation, but we do not, at this point, need to make a decision about which manipulated variable we will actually use for this as it does not affect the steady-state cost.

4.2. Equality Constraints. The two equality constraints must be satisfied, which reduces the number of controlled variables to be selected from 10 to 8. The
directly related variables can be eliminated from further consideration.

The stripper liquid flow (product rate) is directly correlated with the production rate (which is specified). This eliminates one manipulated variable and one directly related measurement.

The separator liquid flow is also strongly correlated with the production rate and should not be kept constant (eliminates one manipulated variable).

The ratio G/H in the product is specified to be 1, and because the product contains mostly G and H, this means that there should be about 50% G and 50% H in the product (specified). This eliminates the two corresponding related measurements (% G in the product and % H in the product).

Together, the two equality constraints specify the amounts of products G and H. From stoichiometry, one can then conclude that none of the four feed streams (A, D, E, and C) should be kept constant. However, mainly for illustration, we retain these variables as degrees of freedom for now, but we show that they are indeed eliminated on the basis of feasibility and cost considerations.

Note that we do not need, at this point, to make a decision about which manipulated variables we will actually use to satisfy the equality constraints as this does not affect the steady-state analysis.

4.3. Active Constraint Control. As mentioned, there are five active constraints, and this reduces the number of controlled variables to be selected from eight to three.

Again, the directly related variables should be eliminated from further consideration. We know that three of the constraints are related to manipulated variables (compressor recycle valve, stripper steam valve, agitator speed); this eliminates these three manipulated variables and also one directly related measurement (stripper steam) from further consideration. Two of the constraints are related to outputs (reactor level and pressure), which eliminates another two measurements.

We started with 41 measurements and 12 manipulated variables, from which we wanted to select 12 controlled variables. We are now left with 33 measurements and 7 manipulated variables, from which we want to select 3 unconstrained controlled variables. This gives 9880 possible combinations, which is still much too large.

4.4. Eliminate/Group Closely Related Variables. The controlled variables should be independent (requirement 3).

Six of the remaining manipulated variables are measured (feeds A, D, E, and C; stripper liquid flow; and purge flow), that is, there is a one-to-one correlation with a measurement (eliminates six measurements).

The purge and recycle streams have the same composition, and because the recycle stream makes up about 2/3 of the reactor feed, it follows that there are only small differences between controlling the purge and reactor feed compositions. We therefore eliminate reactor feed composition (eliminates six measurements).

Note that the choice of which variables to keep and which to eliminate was more or less arbitrary, but because the variables are closely related, it does not matter very much in the further analysis. The main idea is to keep one variable in each group of related variables.

4.5. Process Insight: Eliminate Further Candidates. Based on an understanding of the process, some additional variables can be excluded from the set of possible candidates for control.

The pressure drops should be as small as possible; thus, with constant (maximum) reactor pressure, the pressures in the separator and stripper should be allowed to float (eliminates two measurements).

The condenser and reactor cooling water flow rates should not be held constant, as that would imply a loss for disturbances 4 and 5 (eliminates two manipulated variables). For the same reason, we should not keep the reactor and separator cooling water outlet temperatures constant (eliminates two measurements).

4.6. Eliminate Single Variables That Yield Infeasibility or Large Loss. The idea is to keep a single candidate variable constant at its nominally optimal value, and evaluate the loss for (1) various disturbances and (2) expected implementation error for this variable for the “best” case with the remaining degrees of freedom reoptimized. If the loss is large (or even worse, if no feasible solution is found), then this variable can be eliminated from further consideration.

Infeasibility. Keeping any one of the following four manipulated variables constant results in infeasible operation for disturbance 2 (inert feed fraction): D feed flow, E feed flow, C feed flow (stream 4), and purge flow (see Table 2). This result is independent of how the two remaining degrees of freedom are used, as is further illustrated in Figure 4, where we see that the nominally optimal purge rate results in infeasible operation for disturbance 2. We also see from Figure 4 that a small negative implementation error in the purge rate will yield infeasibility.

Loss. We have now 1 manipulated variable (A feed flow) and 17 measurements left. Table 3 shows the loss (deviation above optimal value) for fixing each one of these 18 variables at a time, and reoptimizing with respect to the two remaining degrees of freedom. The
losses with constant A feed flow and constant reactor feed rate are totally unacceptable for disturbance 1 (eliminates one manipulated variable and one measurement). In fact, we could have eliminated these earlier based on their close relationship to the product rate equality constraint. The remaining 15 measurements yield reasonable losses. However, we have decided to eliminate variables with a loss larger than 6 $/h when summed for the three disturbances. This eliminates the following five measurements: separator temperature, stripper temperature, B (inert) in purge, G in purge, and H in purge.

4.7. Eliminate Pairs of Constant Variables with Infeasibility or Large Loss. We are now left with 11 candidate measurements, that is, (11!-9!)/3!-2! = 165 possible combinations of three variables.

The next natural step is to proceed with keeping pairs of variables constant and evaluate the loss with the remaining degree of freedom reoptimized. However, there are 55 combinations of pairs, so this in itself would be a very large effort. We therefore choose to skip this step in the procedure.

5. Selection of Controlled Variables

We are now at Step 6: Evaluation of Loss for Alternative Combinations of Controlled Variables. This is done by computing the loss imposed by keeping constant setpoints when there are disturbances or implementation errors.

As mentioned, we are now left with 165 possible combinations of three variables. An initial screening based on computing losses\(^1\) indicates that one of the three controlled variables should be reactor temperature, which is the only remaining temperature among the candidate variables. Furthermore, reactor temperature is proposed by most authors, and it is normally easy to control, so we will now only consider combinations that include reactor temperature.

A further evaluation shows that we should eliminate % F (byproduct) in the purge as a candidate variable, because the optimum is either very "sharp" in this variable, or optimal operation is achieved close to its maximum achievable value (see a typical plot in Figure 5). In either case, operation will be very sensitive to the implementation error for this variable.

5.1. Evaluation of Disturbance Losses. The losses for the remaining 9-8/2 = 36 possible combinations of two variables are shown in Table 4. Note that the recycle flow is the flow from the compressor, which is returned to the reactor, and it should not be confused with the compressor recycle valve, which we have chosen to keep closed. Not surprisingly, keeping both the recycle flow and compressor work constant results in infeasibility. This is as expected, because from process insight, these two variables are closely correlated (and we should probably have eliminated one of them earlier).

We note that constant % F in the product in all cases results in a large loss or infeasibility for disturbance 2. This, combined with the above finding that we should not control % F in the purge, leads to the conclusion that it is not favorable to control the composition of byproduct (F) for this process.

The following four combinations of variables have a summed loss of less than 6 $/h:

- Case I: reactor temperature, recycle flow, and C in purge (loss 3.8 $/h).
- Case II: reactor temperature, compressor work, and C in purge (loss 3.9 $/h).
- Case III: reactor temperature, C in purge, and E in purge (loss 5.1 $/h).
- Case IV: reactor temperature, C in purge, and D in purge (loss 5.6 $/h).

The choice of Rickr,\(^1\) with reactor temperature, A in purge, and C in purge, is somewhat less favorable, with a summed loss of 9.8 $/h.

5.2. Evaluation of Implementation Losses. In addition to disturbances, there will always be an implementation error related to each controlled variable, that is, a difference between its setpoint and its actual value, e.g., because of measurement error or poor control. In Figure 5, we plot the cost as a function of the three controlled variables for best case I (the plots for case II are nearly identical and are not shown). We see that the optimum is flat over a large range of the three controlled variables, and we conclude that implementation error will not cause a problem.

5.3. Summary. In conclusion, control of reactor temperature, C in purge, and recycle flow and control of reactor temperature, C in purge, and compressor
work (cases I and II, respectively) result in small losses for disturbances, have flat optima (and are thus insensitive to implementation error), and therefore represent good candidates for self-optimizing control.

6. Implementation and Analysis of Controllability

We are now at Step 7: Final Evaluation and Selection. The analysis up to now has been based purely on steady-state economics, and we have said nothing about implementation of the proposed controlled variables. Obviously, this is also an important consideration, as one choice of controlled variables might result in a system that is easy to control, whereas another might result in serious control problems, for example, caused by unstable (RHP) zeros (the multivariable extension of inverse response behavior).

The truly optimal approach would be to solve the entire problem as one big optimization problem, taking into account both economics and control. However, this is intractable for most real problems, and the approach taken in this paper is therefore preferred. Here, we first identify candidate sets of controlled variables with acceptable steady-state economics. We then check the controllability of the best alternative (case I in our case).

If it is acceptable, then we have found a viable solution. If it is not acceptable, then we check the remaining candidates. If none of these turns out to be controllable, then we relax our requirements on the steady-state economics and consider more candidates.

A procedure for controllability analysis is given on page 246 in Skogestad and Postlethwaite.² It is based on first obtaining a linearizing model and then determining whether the disturbances in questions can be rejected with the available inputs taking into account the presence of RHP zeros, etc. However, we were not able to obtain a linearized model of the effect of the disturbances for the Tennessee Eastman process. Partly for this reason and partly because most engineers are more convinced by closed-loop simulations, we use the “simulation approach” here to evaluate the controllability.

In the simulation approach, we propose a particular control structure, tune the controllers, and show with simulations that control is acceptable. Note that model uncertainty should generally be included in these simulations. If we use decentralized control (as we do), then model uncertainty is generally less critical, as a decentralized controller does not really make much use of the plant model. If we can find a particular tuning with acceptable control, then we can conclude that the plant is controllable, at least for the disturbance and uncertainty scenario considered. However, the simulation approach generally suffers from the problem that it depends on the particular tunings and disturbances used in the simulations, and this can make it difficult to draw definite conclusions.

7. Simulation of the Proposed Control Structure

We consider here control of the plant using the controlled variables with the best steady-state economics (case I). We propose a decentralized control system and show by simulations that acceptable control is indeed possible. Note that Simulink files are avail-

---

Table 4. Loss ($/h$) with All Three Degrees of Freedom Fixed

<table>
<thead>
<tr>
<th>case</th>
<th>fixed variables</th>
<th>disturbance 1 throughput</th>
<th>disturbance 2 throughput</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>recycle flow A in purge</td>
<td>1.2</td>
<td>1.3</td>
<td>0.6</td>
</tr>
<tr>
<td>I</td>
<td>recycle flow D in prod</td>
<td>0.2</td>
<td>2.6</td>
<td>4.8</td>
</tr>
<tr>
<td>II</td>
<td>comp work A in purge</td>
<td>126.0</td>
<td>8.0</td>
<td>135.3</td>
</tr>
<tr>
<td>II</td>
<td>comp work E in prod</td>
<td>1.8</td>
<td>8.0</td>
<td>25.6</td>
</tr>
<tr>
<td>II</td>
<td>recycle flow E in prod</td>
<td>1.2</td>
<td>3.2</td>
<td>10.3</td>
</tr>
<tr>
<td>Ricker</td>
<td>A in purge C in purge</td>
<td>0.0</td>
<td>2.4</td>
<td>1.4</td>
</tr>
<tr>
<td>IV</td>
<td>C in purge D in prod</td>
<td>0.0</td>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>III</td>
<td>E in purge E in prod</td>
<td>0.0</td>
<td>2.4</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>comp work F in prod</td>
<td>2.0</td>
<td>32.8</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>comp work F in prod</td>
<td>0.2</td>
<td>2.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

² Reactor temperature is fixed in all cases.
7.1. Decentralized Control Structure. Our first attempt was to design a decentralized control structure following the procedure of Larsson and Skogestad\textsuperscript{15} (the heading numbers below refers to Table 1 in that paper). The resulting structure and PI tunings are given in Table 5.

1. Controlled Variables. Let us summarize the above results (case I). From the degree of freedom analysis in Table 1, we know that there are 12 manipulated variables. However, two of these degrees of freedom are consumed to control liquid levels with no steady-state effect, namely, (1) separator level and (2) stripper level. Furthermore, there are two equality constraints, (3) the production rate (given) and (4) the ratio between G and H in the product (given). For optimal operation, there are five active constraints. Three are related to manipulated inputs, namely, (5) compressor recycle valve (closed), (6) stripper steam valve (closed), and (7) agitator speed (maximum), whereas two are related to outputs, (8) reactor level (maximum) and (9) reactor pressure (minimum).

There are then three remaining unconstrained degrees of freedom that can be used to optimize the operation. On the basis of the above steady-state economic analysis, we propose that the following three variables be controlled at their nominally optimal setpoints (case I) to achieve self-optimizing control: (1) reactor temperature, (11) %C in purge, and (12) recycle flow.

Three of the above "controlled variables" are manipulated inputs (5, 6, and 7), which need no further consideration in terms of control. However, a strategy for controlling the remaining nine variables must be found, and in terms of decentralized control, this implies that each controlled variable must be paired with one of the remaining nine manipulated inputs, which are separator liquid flow, stripper liquid product flow, C feed flow, D feed flow, E feed flow, purge flow, reactor cooling water flow, A feed flow, and condenser cooling water flow.

In addition, we might need to close some loops for stabilization or to improve local disturbance rejection, but the setpoints for these "inner" loops can be used as degrees of freedom, so this does not effect the above steady-state analysis, although it might affect selection of the control structure (pairing of variables).

2. Production Rate. Where should the throughput be set? This is a very important choice, as it determines the structure of the remaining inventory control system.

In our case, the most obvious choice is to use the stripper liquid flow, which is the product stream. However, this stream is most likely needed for stabilizing the stripper level, which has no steady-state effect. The production rate must therefore be set upstream, and in the decentralized scheme, we will use largest of the four feed streams, which in this case is the C feed flow.

Another rather obvious choice for adjusting the production rate is to use the total feed rate, i.e., the sum of the four feed streams. However, this does not give a decentralized control scheme and will be considered later in the improved control structure.

3a. Regulatory Control Layer: Stabilization. There are two integrating liquid levels that can be stabilized as follows (here, the symbol $\rightarrow$ means "is paired with" or, more precisely, "is controlled by")

(1) separator level $\rightarrow$ separator liquid flow
(2) stripper level $\rightarrow$ stripper liquid product flow
(10) cooling water outlet temperature $\rightarrow$ cooling water flow

Note that this loop introduces the setpoint for the reactor cooling water outlet temperature as a new
degree of freedom, i.e., it replaces the cooling water flow as a manipulated variable when seen from the layer above.

3b. Regulatory Control Layer: Local Disturbance Rejection. In general, we use extra local measurements in inner cascades to improve local disturbance rejection, for example, the use of flow controllers based on measurement of the flow is very common. However, as a first try, we will not use any inner cascades here (except for the temperature controller $10'$, which we introduced for stabilization of the reactor).

4. Supervisory Control Layer: Decentralized Control. We still have seven variables (3, 4, and 8–12 in the above list) that need to be controlled at given setpoints using the seven remaining manipulated variables.

As a first attempt, we use decentralized control and base the pairings on a relative gain array (RGA) analysis of the stabilized $7 \times 7$ system. The main rules for the RGA analysis are (e.g., 2): (1) Avoid pairing on negative RGA elements at steady state. (2) Prefer pairing on RGA elements close to 1 (with the other elements close to 0) at the bandwidth frequency.

The best pairings according to these rules for loops 3, 4, and 8–12 are as given in Table 5. Note that the RGA analysis recommends that production rate should be controlled using the C feed flow, which is the largest.

### Table 5. Decentralized Control Structure

<table>
<thead>
<tr>
<th>loop</th>
<th>controlled variable</th>
<th>manipulated variable</th>
<th>gain $K_c$</th>
<th>integral time $T_i$ (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>separator level</td>
<td>separator liquid flow</td>
<td>-2.5</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>stripper level</td>
<td>stripper liquid product flow</td>
<td>-0.5</td>
<td>300</td>
</tr>
<tr>
<td>3</td>
<td>production rate</td>
<td>C feed flow</td>
<td>0.005</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>product ratio G/H</td>
<td>D feed flow</td>
<td>2</td>
<td>150</td>
</tr>
<tr>
<td>5</td>
<td>reactor level</td>
<td>E feed flow</td>
<td>5</td>
<td>1200</td>
</tr>
<tr>
<td>6</td>
<td>reactor pressure</td>
<td>purge flow</td>
<td>-1</td>
<td>90</td>
</tr>
<tr>
<td>7</td>
<td>reactor temperature</td>
<td>setpoint cooling water outlet temp</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>cooling water outlet temperature</td>
<td>reactor cooling water flow</td>
<td>-10</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>% C in purge</td>
<td>A feed flow</td>
<td>-5</td>
<td>500</td>
</tr>
<tr>
<td>10</td>
<td>recycle flow</td>
<td>condenser cooling water flow</td>
<td>-3</td>
<td>200</td>
</tr>
</tbody>
</table>

Figure 7. Decentralized control structure. $t = 0–22$ h, disturbance 1 in A/C ratio in C feed; $t = 35–45$ h, ramp down of setpoint in % G in the product; $t = 50–60$ h, ramp up of setpoint in % G in the product; $t = 65–70$ h, distractions 12 and 15 in the reactor cooling water; $t = 70–75$ h, disturbance 8 in A, B, and C feed composition; $t = 80–95$ h, ramp up of setpoint for the production rate.
feed stream. With these pairings, the magnitudes of the paired RGA elements at frequency 0.5 rad/h (corresponding to a closed-loop response time of about 2 h) range from about 0.29 (loop 11) to 1.6 (loop 9), indicating that there are significant interactions. This is also expected on the basis of physical insight. For example, if we change the production rate (setpoint loop 3), then we will need to change all of the flows, introducing interactions into most of the other loops.

Dynamic simulations show that this simple decentralized control system performs acceptably for most of the disturbances given by Downs and Vogel, including setpoint changes in the production rate and in the G/H ratio. The PI controllers were initially tuned individually using the Ziegler–Nichols method and were then implemented and retuned sequentially, starting with the fast loops. A typical response for a combination of disturbances is shown in Figure 7. We note that a disturbance in the A/C ratio in the C feed stream (disturbance 1) increases the %C in the purge stream, which, through loop 11, increases the A feed flow. This causes increased reactor pressure, but it remains inside the limit of 2895 kPa.

### Table 6. Improved Control Structure with Some Decoupling

<table>
<thead>
<tr>
<th>loop</th>
<th>controlled variable</th>
<th>manipulated variable</th>
<th>gain $K_c$</th>
<th>integral time $T_1$ (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>separator level</td>
<td>separator liquid flow</td>
<td>−0.001</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>stripper level</td>
<td>stripper liquid product flow</td>
<td>−0.0002</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>production rate</td>
<td>total feed flow</td>
<td>3.2</td>
<td>120</td>
</tr>
<tr>
<td>4</td>
<td>product ratio G/H</td>
<td>D/E feed flow ratio</td>
<td>−0.032b</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>compressor recycle valve: closed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>stripper steam valve: closed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>agitator speed: maximum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>reactor level</td>
<td>setpoint sep temperature (loop 8')</td>
<td>0.8</td>
<td>60</td>
</tr>
<tr>
<td>8' (casc)</td>
<td>separator temperature</td>
<td>condenser cooling water flow</td>
<td>−4</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>reactor pressure</td>
<td>purge flow</td>
<td>−0.0001</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>reactor temperature</td>
<td>reactor cooling water flow</td>
<td>−8</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>%C in purge</td>
<td>C feed flow</td>
<td>0.0009</td>
<td>562</td>
</tr>
<tr>
<td>12</td>
<td>recycle flow</td>
<td>A feed flow</td>
<td>0.00125</td>
<td>120</td>
</tr>
</tbody>
</table>

*a In addition: inner cascade flow controllers with pure integral action on all flows. b Decoupling: tunings are with implementation strategy of Ricker.*

**Figure 8.** Improved control structure with some decoupling (same disturbances and setpoint changes as in Figure 7).
some disturbances because inner flow controllers are not included. Our "improved" control structure with some decoupling is given in Table 6. It is very similar to that of Ricker\textsuperscript{11} but we had to make some modifications because he controls A in the purge instead of in the recycle flow. The main idea is to introduce physical decoupling by (1) using the total feed flow to control the production rate and (2) using the D/E feed flow ratio to control the product ratio G/H (as is reasonable from stoichiometry). In addition, flow controllers are implemented to improve local disturbance rejection.

The dynamic response with this control structure is significantly better, as illustrated by the dynamic simulations presented in Figure 8. As before, the disturbance in the A/C ratio in the C feed stream increases %C in the purge stream, which, through loop 11, reduces the C feed flow. However, the total feed flow is constant (loop 3), so all of the other flows are increased to compensate for this, and there is almost no interaction into the other loop. For example, the reactor pressure remains almost constant.

The responses to other disturbances are also very good, and simulations (not included) show that they are generally similar to or slightly better than the responses of Ricker\textsuperscript{11} which was the basis for our improved structure.

In conclusion, the dynamic simulations show that the set of controlled variables with the best self-optimizing properties in terms of steady-state economics (case I) is controllable. It is possible to achieve acceptable control with a simple decentralized control structure, but performance is improved markedly by introducing some simple "decoupling" elements such as use of the total feed flow as a manipulated variable.

8. Discussion

8.1. Should Inert Composition Be Controlled? A common suggestion is that it is necessary to control the inert composition (in our case, mole fraction of component B) in order to (indirectly) control the inventory of inert components.\textsuperscript{7,9,12,13} However, recall that we eliminated the composition of B in the purge at an early stage because it gave a rather large loss for disturbance 2 (see Table 3). Moreover, and more seriously, we generally find that the shape of the economic objective function as a function of inert composition is very unfavorable, either with a sharp minimum or with the optimum value close to infeasibility. A typical example of the latter is shown in Figure 9. In conclusion, we do not recommend that inert composition be controlled. The inventory of inert in the system is, in our case, indirectly controlled, for example, by controlling the reactor pressure and recycle flow.

8.2. Combinations of Variables. In this paper, we have presented a number of criteria for reducing the number of alternative groupings of variables to be controlled. Note that the number of alternatives would have been much larger if we had also considered combinations of variables, such as sums, differences, ratios, and so on.

However, note that combinations between already selected variables do not need to be considered as they do not affect the economic loss analysis presented in this paper. For example, assume that we have selected the two temperatures \(T_1\) and \(T_2\) as controlled variables. Then, the steady-state disturbance effect will be the same if we choose to keep \(T_1\) and \(T_2\) constant, or if we choose to keep, say, \(T_1 - T_2\) and \(T_2\) constant. However, the effect of the implementation error might differ if we have direct measurements of the combined variables, and this might offer some advantage, for example, if we have an accurate measurement of the difference \(T_1 - T_2\).

Similarly, manipulated variable combinations and inner cascade controllers do not affect the analysis. For example, the use of the total feed rate as a manipulated variable in the improved control scheme has a dynamic effect but no steady-state effect, because the controlled variables that we keep constant at steady state are the same for the two schemes.

9. Conclusion

In this study of the Tennessee Eastman process, we have focused on the selection of the controlled variables using the concept of self-optimizing control, that is, achieving acceptable loss with constant setpoints in the face of disturbances and implementation errors. The conclusion is that, in addition to the constrained variables, we should control reactor temperature, composition of C in the purge, and recycle flow or compressor work.

A very common suggestion is that it is necessary to control the inert composition. However, this choice can lead to serious operational problems, as demonstrated by Figure 9, and in a more careful evaluation, we did not find any favorable combination that included the inert composition.

Note that a systematic approach such as that taken in this paper can result in a control scheme that is not an obvious choice even for a trained control engineer. For example, during the review of the first version of this paper, in which we had not included dynamic simulations to confirm the feasibility of the proposed controlled variables, one reviewer wrote that "it may be impossible to design such a plantwide scheme", whereas another wrote that "this reviewer is convinced that the final control scheme suggested in this paper will not work either in simulation or in practice". Nevertheless, in this paper, we have confirmed using dynamic simulation the feasibility of the proposed control structure.

Literature Cited

\textsuperscript{(1)} Morari, M.; Stephanopoulos, G.; Arkun, Y. Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the problem. Process decomposition and the

Received for review June 19, 2000 Revised manuscript received June 26, 2001 Accepted July 30, 2001