Local Self-Optimizing Control with Average Loss Minimization

Vinay Kariwala[§,*], Yi Cao[‡] and S. Janardhanan[§,†]

[§] Division of Chemical & Biomolecular Engineering,

Nanyang Technological University, Singapore 637459

[‡]School of Engineering, Cranfield University, Cranfield, Bedford MK43 0AL, UK

Abstract

Self-optimizing control provides nearly optimal operation of process systems in face of varying disturbances, where the inputs are adjusted to hold selected controlled variables (c) at constant setpoints. It is possible to have better self-optimizing properties by controlling linear combinations of measurements ($c = Hy$) than by controlling individual measurements. Previous work[10] focussed on selecting combination matrix H to minimize worst-case loss, that arises due to the use of sub-optimal self-optimizing control policy. In this paper, we present a method for finding combination matrix H that minimizes average loss for local self-optimizing control. It is further shown that the combination matrix that minimizes average loss is super-optimal in the sense that it also minimizes worst-case loss simultaneously. The usefulness of the results is demonstrated using an evaporator case study.

Keywords: Control structure design, Controlled variables, Optimal operation, Self-optimizing control.

[*]Corresponding author: Tel: +65-6316-8746; Fax: +47-6794-7553 ; E-mail:vinay@ntu.edu.sg

[†]Present address: Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi - 110016, India

1 Introduction

Controlled variables (CVs) are the variables which are maintained at constant setpoint values by a control system. Appropriate selection of CVs represents a critical decision during the design of a control system, which not only affects the control system performance but also influences the overall plant operation. Traditionally, CVs have been selected based on intuition and process knowledge. In early 1980s, Morari and co-workers studied the implication of CV selection on optimal plant operation.[13] Recently, this work has been extended by Skogestad,[15] who coined the term self-optimizing control. In self-optimizing control, instead of trying to achieve the best possible performance, a small trade-off is made between the performance and the simplicity of the approach. Here, a set of appropriate CVs is found which, if held constant, provides acceptable loss in comparison with the use of an online optimizer to achieve optimal operation.

CVs are usually selected as a subset of available measurements, but it is possible to have better self-optimizing properties by controlling linear combinations of measurements.[1,7,10] In either case, the choice of CVs based on the general non-linear formulation of self-optimizing control requires solving large-dimensional non-convex optimization problems.[15] To quickly pre-screen alternatives, local methods are used. The first local approach to select CVs within the self-optimizing control framework is the approximate minimum singular value or the maximum gain rule.[16] Halvorsen *et al.*[7] provided a rigorous analysis of the worst-case loss and presented a nonlinear optimization approach to find the locally optimal linear combinations of measurements that can be used as CVs. The nonlinear optimization approach can be time consuming and may also converge to local optima. Alstad and Skogestad[1] proposed the use of the computationally more efficient null space method to find measurement combinations. However, the method ignores implementation error, which arises in every real process due to the presence of measurement error and other uncertainties. Thus, the null space method is capable of providing only a sub-optimal solution.

Recently, Kariwala[10] presented an explicit solution to the problem of finding the optimal measurement combinations which minimize local worst-case loss. The method proposed by Kariwala[10] is efficient and only requires the use of singular value and eigenvalue decomposition of certain matrices. In this paper, we improve the computational efficiency of this method further.

All aforementioned designs are based on worst-case loss minimization. The solution obtained by the use of available methods may be conservative as the worst-case may not occur frequently in practice. An alternative to worst-case loss minimization is to select CVs by minimizing the average loss, which represents the expected loss incurred over the long-term operation of the plant. The main contribution of this work is the extension of the available techniques for local self-optimizing control based on worst-case loss minimization to average loss minimization. We derive expressions for local average loss and subsequently use these expressions to explicitly characterize the combination matrices that minimize average loss. A natural question about the new design approach is that how the self-optimizing control performance will be if actual disturbances and implementation error are very different from the average-case scenario. To address this issue, we show that the combination matrices that minimize average loss are super-optimal in the sense that they also minimize worst-case loss simultaneously. Thus the use of average-case optimal combination matrices is always advantageous, as they not only minimize the average loss, but also ensure that the largest loss that can occur over all disturbance and implementation error scenarios is no larger than seen by using worst-case optimal combination matrices. The usefulness of the results is demonstrated using the evaporator case study,[14] which also highlights some of the outstanding issues related to selection of CVs in the self-optimizing control framework.

The rest of this paper is organized as follows: After summarizing self-optimizing control principles, the optimization problems to minimize the worst-case and average losses are defined in Section 2. Then, numerically efficient algorithms to calculate the minimal achievable losses and corresponding combination matrices are derived in Section 3. In Section 4, an evaporator case study is presented to demonstrate the

advantages of these new algorithms. Finally, the work is concluded in Section 5.

2 Self-optimizing control

In this section, we present the available expression for worst-case loss and derive expression for average loss for local self-optimizing control. We denote the inputs or manipulated variables and disturbances by u and d, respectively. Let us assume that the steady-state economics of the plant are characterized by the scalar objective functional $J(u, d)$. In presence of varying disturbances, the optimal operation policy is to update u according to d using an online optimizer, which provides the optimal value of the objective functional denoted as $J_{\text{opt}}(d)$.

An alternative and simpler strategy is to update u using a feedback controller, which manipulates u to hold the CVs c at their specified setpoints. For this strategy, let the value of the objective functional be $J_c(d, n)$. Note that when the feedback-based policy in used, u and thus J is also affected by the error n in implementing constant setpoint policy, which results due to measurement noise and other uncertainties. The simpler strategy results in a loss and self-optimizing control is said to occur, when the loss is acceptable.[15] Based on this concept, the CVs can be selected by comparing the worst-case or average losses of different alternatives, where

$$\text{Worst-case loss} = \max_{d\in\mathcal{D}} \max_{n\in\mathcal{N}} \left(J_{\text{opt}}(d) - J_c(d, n)\right) \tag{1}$$

$$\text{Average loss} = \frac{1}{|\mathcal{D}|\,|\mathcal{N}|} \int_{\mathcal{D}} \int_{\mathcal{N}} \left(J_{\text{opt}}(d) - J_c(d, n)\right) \tag{2}$$

where $\mathcal{D}$ and $\mathcal{N}$ represent the sets of allowable disturbances and implementation errors, respectively and $|\cdot|$ denotes the cardinality of a set. Using (1) or (2), the evaluation of loss for a general nonlinear process is difficult. To quickly pre-screen the alternatives, instead local methods are used, as discussed below.

2.1 Local methods

The local methods are based on the second order accurate Taylor series expansion of the loss function $(J_{\text{opt}}(d) - J_c(d,n))$ and a linear model of the process. It is also assumed that the set of active constraints for the process does not change with d and n. Cao[2] has considered the case, when the set of active constraints changes with the disturbances.

Let the linearized model of the process, obtained around the nominally optimal operating point, be given as

$$y = G^y\, u + G_d^y\, W_d\, d + W_n\, n \tag{3}$$

where y denote the process measurements and the diagonal matrices W_d and W_n contain the magnitudes of expected disturbances and implementation errors associated with the individual measurements, respectively. We have $y, n \in \mathbb{R}^{n_y}$, $u \in \mathbb{R}^{n_u}$ and $d \in \mathbb{R}^{n_d}$ with $n_y \geq n_u$. The CVs c are given as

$$c = H\, y = G\, u + G_d\, W_d\, d + n_c \tag{4}$$

where

$$G = H\, G^y, \quad G_d = H\, G_d^y \quad \text{and} \quad n_c = H\, W_n\, n \tag{5}$$

It is assumed that the dimension of c is the same as the dimension of u and $G = H\, G^y$ is invertible. The second assumption is necessary to ensure that the CVs can be maintained at the specified setpoints using a controller with integral action. Halvorsen *et al.*[7] have shown that for given d and n, the loss is given as

$$L = \frac{1}{2} \left\| \begin{bmatrix} M_d & M_n \end{bmatrix} \begin{bmatrix} d \\ \\ n \end{bmatrix} \right\|_2^2 \tag{6}$$

where

$$M_d = J_{uu}^{1/2}\left(J_{uu}^{-1}J_{ud} - G^{-1}G_d\right)W_d \tag{7}$$

$$M_n = J_{uu}^{1/2}G^{-1}W_n \tag{8}$$

Here J_{uu} and J_{ud} represent $\frac{\partial^2 J}{\partial u^2}$ and $\frac{\partial^2 J}{\partial u \partial d}$, evaluated at the nominally optimal operating point, respectively. Note that for a well-defined minimization problem, J_{uu} is always positive-definite and thus $J_{uu}^{1/2}$ is guaranteed to exist.

Remark 1 Though J is considered to be the economic objective function in this paper, the proposed results also hold for other scalar objective functions. For example, in the case of indirect control, the CVs are found such that J defined as the steady-state offset in primary Vs is minimized.[9]

To present the worst-case and average loss expressions, similar to Halvorsen *et al.*,[7] we first consider that d and n are constrained to satisfy

$$\left\| \begin{bmatrix} d \\ n \end{bmatrix} \right\|_2 \leq 1 \tag{9}$$

Some discussion on scenarios, where the allowable set of d and n differs from that given by (9), is provided in Section 2.3.

Theorem 1 Let M_d and M_n be defined by (7) and (8), respectively. The worst-case loss over the allowable set of disturbances and implementation errors defined by (9) is given as[7]

$$L_{\text{worst}} = \frac{1}{2}\bar{\sigma}^2\left(\begin{bmatrix} M_d & M_n \end{bmatrix}\right) \tag{10}$$

where $\bar{\sigma}(\cdot)$ is the maximum singular value.

The proof of Theorem 1 is given by Halvorsen *et al.*[7] As argued in Section 1, the minimization of worst-case loss can be conservative, as it may not occur frequently in practice. Often, it is more appropriate to minimize the average loss. In the next proposition, we derive an expression for average loss.

Proposition 1 Let M_d and M_n be defined by (7) and (8), respectively. The average loss over the allowable set of disturbances and implementation errors defined by (9) is given as

$$L_{\text{average}} = \frac{1}{6\,(n_y + n_d)} \left\| \begin{bmatrix} M_d & M_n \end{bmatrix} \right\|_F^2 \tag{11}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Proof: Let $M = \begin{bmatrix} M_d & M_n \end{bmatrix}$ and $\tilde{d} = \begin{bmatrix} d \\ n \end{bmatrix}$. Based on (6), the average loss is given as

$$\begin{aligned} L_{\text{average}} &= \frac{1}{2} E\left[\text{tr}\left(M\,\tilde{d}\,\tilde{d}^T\,M^T\right)\right] \\ &= \frac{1}{2} E\left[\text{tr}\left(M^T\,M\,\tilde{d}\,\tilde{d}^T\right)\right] \\ &= \frac{1}{2} \text{tr}\left(M^T\,M\,E\left[\tilde{d}\,\tilde{d}^T\right]\right) \end{aligned}$$

where $\text{tr}(\cdot)$ denotes the trace of the matrix and $E[\cdot]$ is the expectation operator. To find $E\left[\tilde{d}\,\tilde{d}^T\right]$, note that the set (9) can be represented as $\|\tilde{d}\|_2 = \alpha$, where α is uniformly distributed over the range $0 \le \alpha \le 1$. Since each element of $\tilde{d} \in \mathbb{R}^{n_y+n_d}$ is independently and identically distributed, $E[\tilde{d}_i \tilde{d}_j] = \delta_{ij} E[\alpha^2/(n_y+n_d)]$, where $\tilde{d}_i$ is the ith element of $\tilde{d}$ and δ_{ij} is the Kronecker delta. Thus,

$$E\left[\tilde{d}\,\tilde{d}^T\right] = \frac{E\left[\alpha^2\right]}{n_d + n_y} I = \frac{\int_0^1 \alpha^2 d\alpha}{n_d + n_y} I = \frac{1}{3(n_d + n_y)} I$$

which implies that the average loss is given by (11). ■

For specified CVs, the computation of local worst-case and average losses using (10) and (11), respectively, requires computation of matrix norms only. In comparison, the computation of exact losses using (1)-(2) is much more involved. Though the expressions for local losses are guaranteed to be accurate only in a small neighborhood of the nominal operating point, they are useful for quick pre-screening of alternatives.

2.2 Selection of controlled variables

Note that the losses in (10) and (11) depend on H and the CVs are selected by minimizing the losses with respect to H. Next, we briefly discuss different approaches for selecting H.

Individual measurements. When individual measurements are selected as CVs, the optimization problem involves finding the best n_u measurements such that the loss in (10) or (11) is minimized. In this case, the elements of H are restricted to be 0 or 1 and

$$HH^T = I, \text{ where } H \in \mathbb{R}^{n_u \times n_y}. \tag{12}$$

Under minor assumptions, the exact local method minimizing the worst-case loss in (10) can be simplified to provide the approximate minimum singular value (MSV) or maximum gain rule.[16] To select the optimal subset of measurements based on the MSV rule, branch and bound based search methods have been proposed by Cao *et al.*[4] and, Kariwala and Skogestad.[11] These methods avoid enumeration of all possible alternatives. Finding efficient search methods for selection of CVs based on (10) and (11) is currently under research.

Measurement combinations. Instead of using individual measurements, it is possible to use combinations of measurements as CVs. In this case, the integer restriction of H is relaxed but the condition $\text{rank}(H) = n_u$ is still imposed to ensure invertibility of $H\,G^y$. Halvorsen *et al.*[7] used non-linear optimization for finding H, which can be very time consuming, and more importantly can converge to local optima. As an alternative, Alstad and Skogestad[1] proposed the use of the null space method. In this method, the implementation error is ignored and H is selected such that

$$H\,\left(G^y\,J_{uu}^{-1}\,J_{ud} - G_d^y\right) = 0 \tag{13}$$

or H is in the null space of $G_y\, J_{uu}^{-1}\, J_{ud} - G_d$. It can be verified that when (13) holds, $\bar{\sigma}(M_d) = 0$. Clearly, the assumption of ignoring the implementation error is limiting and can only provide a sub-optimal solution. Due to this assumption, in some cases, the local loss obtained by controlling measurement combinations found using null space can be higher than the case where individual measurements are controlled, *e.g.* for the binary distillation column discussed by Hori and Skogestad.[9] In addition, for (13) to hold, it is necessary that $n_y \geq n_u + n_d$. When less than $(n_u + n_d)$ measurements are available, the null space method cannot be applied. Recently, Kariwala[10] has presented an explicit solution to the problem of minimizing local worst-case loss. In this paper, we primarily aim at extending the results of Kariwala[10] to average loss minimization.

2.3 Allowable region of disturbances and implementation errors

[Table 1 about here.]

In the derivation of Theorem 1 and Proposition 1, similar to Halvorsen *et al.*,[7] we have assumed that the allowable set of d and n is given by (9). The set in (9), however, implies correlation between different disturbances and implementation errors. In practice, the variation of individual disturbance and implementation error is independent of each other. This scenario can be appropriately represented by defining the allowable set as

$$|d_i| \leq 1, |n_j| \leq 1 \quad i = 1, 2, \cdots, n_d; j = 1, 2, \cdots, n_y \tag{14}$$

or $\left\| \begin{bmatrix} d^T & n^T \end{bmatrix}^T \right\|_\infty \leq 1$, where $\|\cdot\|_\infty$ denotes maximum absolute value of the elements of the vector. In this case, by following the proof of Proposition 1, it can be shown that the average loss is given as $(1/6)\| \begin{bmatrix} M_d & M_n \end{bmatrix} \|_F^2$. Another possibility is to allow d and n to be distributed normally, *i.e.*

$$d_i \sim N(0,1), n_j \sim N(0,1) \quad i = 1, 2, \cdots, n_d; j = 1, 2, \cdots, n_y \tag{15}$$

where $N(0,1)$ denotes a normal distribution with zero mean and unity standard deviation. In this case, W_d and W_n contain the standard deviations of d and n, respectively. Here, d_i^2 and n_j^2 follow χ^2 distribution and $E[d_i^2] = E[n_j^2] = 1$. Thus, following the proof of Proposition 1, the expression for average loss becomes $(1/2)\| \begin{bmatrix} M_d & M_n \end{bmatrix} \|_F^2$. In summary, the average loss for commonly used allowable regions for d and n changes only by constant factors (see Table 1) and selection of H is not affected by allowable region for d and n.

On the other hand, the problem of computing local worst loss is ill-defined, when d and n are distributed normally. This happens as in the worst-case d and n can be arbitrary large implying $L_{\text{worst}} = \infty$. When the allowable set of d and n is given by (14), derivation of explicit expression for worst-case loss reduces to solving a combinatorial optimization problem.[3] Note that the set $\left\| \begin{bmatrix} d^T & n^T \end{bmatrix}^T \right\|_2 \leq \sqrt{n_y + n_d}$ includes the set (14). Thus, the worst-case loss over allowable region of d and n defined by (14) is upper bounded by $0.5(n_y+n_d)\bar{\sigma}^2 \left(\begin{bmatrix} M_d & M_n \end{bmatrix} \right)$. In this case, selection of H by minimizing (10) is equivalent to minimizing an upper bound on worst-case loss.

3 Optimal Measurement Combination

In this section, we present explicit solutions to the problems of finding optimal measurement combinations that can be used as CVs. The cases of worst-case and average loss minimization are dealt with in turn. In the following discussion, as a shorthand notation, we denote

$$Y = \begin{bmatrix} (G^y J_{uu}^{-1} J_{ud} - G_d^y) W_d & W_n \end{bmatrix} \tag{16}$$

3.1 Worst-Case Loss Minimization

The various available methods that can be used to find locally optimal or sub-optimal measurement combinations for worst-case loss minimization are discussed in Section 2.2. Among the available methods, the solution presented by Kariwala[10] is optimal and efficient, which is recalled below.

Theorem 2 Let $\lambda_1, \lambda_2, \cdots, \lambda_{n_y}$ be the eigenvalues of $\left(\gamma^2 G^y J_{uu}^{-1} (G^y)^T - Y Y^T\right)$ arranged in decreasing order. Then, the minimal loss is given as

$$L = 0.5\,\gamma^2 \tag{17}$$

where $\gamma > 0$ is the smallest scalar satisfying

$$\lambda_{n_u}\left(\gamma^2 G^y J_{uu}^{-1} (G^y)^T - Y Y^T\right) = 0 \tag{18}$$

Let $\nu_1, \nu_2, \cdots, \nu_{n_u}, \cdots, \nu_{n_y}$ be the mutually orthogonal eigenvectors of $\left(\gamma^2 G^y J_{uu}^{-1} (G^y)^T - Y Y^T\right)$ such that (18) holds. Then, the H matrix can be chosen as

$$H = C \begin{bmatrix} \nu_1 & \nu_2 & \cdots & \nu_{n_u} \end{bmatrix}^T \tag{19}$$

where $C \in \mathbb{R}^{n_u \times n_u}$ is any non-singular matrix.

In Kariwala,[10] singular value decomposition is used for computing γ that satisfies (18). In the following discussion, we present an alternate method, which is computationally more efficient and is expressed directly in terms of G^y, J_{uu} and Y.

Lemma 1 For $A \in \mathbb{R}^{m \times n}$, $m \leq n$, the largest m eigenvalues of $A\,A^T - I$ and $A^T\,A - I$ are the same.

Proof: Let the singular value decomposition of A be $A = U\Sigma V^T$. As U and V are unitary matrices, $A\,A^T - I = U\Sigma^2 U^T - I = U(\Sigma^2 - I)U^T$. Similarly, it can be shown that $A^T\,A - I = V(\Sigma^2 - I)V^T$.

Considering U and V to be similarity transformation matrices, it follows that the largest m eigenvalues of $A\,A^T - I$ and $A^T\,A - I$ are the same. ■

Proposition 2 The γ that solves (18) is given as

$$\gamma = \overline{\lambda}^{1/2} \left(J_{uu}^{-0.5} \, (G^y)^T \, (Y\,Y^T)^{-1} \, G^y \, J_{uu}^{-0.5} \right)^{-1} \tag{20}$$

Proof: Let R be an upper-triangular matrix that satisfies $Y\,Y^T = R^T\,R$ (Cholesky factor). Using congruence transformation (see *e.g.* Skogestad and Postlethwaite[16]), it follows that the eigenvalues of $\left(\gamma^2\,G^y\,J_{uu}^{-1}\,(G^y)^T - Y\,Y^T\right)$ are the same as the eigenvalues of $\left(\gamma^2\,R^{-T}\,G^y\,J_{uu}^{-1}\,(G^y)^T\,R^{-1} - I\right)$. Define $Q = R^{-T}\,G^y\,J_{uu}^{-0.5}$. Based on Lemma 1, the first n_u eigenvalues of $\left(\gamma^2\,Q\,Q^T - I\right)$ and $\left(\gamma^2\,Q^T\,Q - I\right)$ are same, when arranged in decreasing order. Then,

$$\gamma^2\,Q^T\,Q - I \succeq 0$$

$$\Leftrightarrow \quad \left(Q^T\,Q\right)^{-1} \preceq \gamma^2\,I$$

Thus, the smallest γ that solves (18) is given as $\bar{\lambda}^{1/2}\left(\left(Q^T\,Q\right)^{-1}\right)$. Now, the expression in (20) follows by recognizing that $Q^T\,Q = J_{uu}^{-0.5}\,(G^y)^T\,(Y\,Y^T)^{-1}\,G^y\,J_{uu}^{-0.5}$. ■

The explicit expression for γ in (20) is used to prove an important property of the combination matrices in Section 3.3. Furthermore, as shown in Section 4, controlling the combinations of a subset of measurements can often provide similar loss as obtained by controlling combinations of all available measurements. To find a suitable subset of measurements, the loss needs to be evaluated several times. Here, the use of the explicit expression for γ in (20) is advantageous, which can be used to compute loss about 10 times faster than the singular value decomposition based method presented by Kariwala.[10]

3.2 Average Loss Minimization

The following lemma establishes the basis for finding the optimal combination of measurements, when average loss is minimized.

Lemma 2 The matrix H minimizing the loss in (11) can be found by solving

$$\min_{H,X} \frac{1}{6\,(n_d + n_y)} \operatorname{tr}(X) \tag{21}$$

$$\text{s.t.} \quad H\,\left(G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}(G^y)^T - Y\,Y^T\right)\,H^T \succeq 0 \tag{22}$$

$$X \succeq 0 \tag{23}$$

$$\operatorname{rank}(H) = n_u \tag{24}$$

Proof: We note that

$$\begin{aligned}\left[\begin{array}{cc} M_d & M_n \end{array}\right] &= J_{uu}^{1/2}\,G^{-1}\left[\begin{array}{cc} (G\,J_{uu}^{-1}\,J_{ud} - G_d) & H\,W_n \end{array}\right] \\ &= J_{uu}^{1/2}\,(H\,G^y)^{-1}\,H\,Y\end{aligned}$$

The Frobenius norm of $\left[\begin{array}{cc} M_d & M_n \end{array}\right]$ can be minimized by minimizing $\operatorname{tr}(X)$, where X satisfies

$$J_{uu}^{1/2}\,(H\,G^y)^{-1}\,(H\,Y)\,(H\,Y)^T\,(H\,G^y)^{-T}\,J_{uu}^{1/2} \preceq X \tag{25}$$

$$\Leftrightarrow \quad (H\,G^y)^{-1}\,(H\,Y)\,(H\,Y)^T\,(H\,G^y)^{-T} \preceq J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}$$

$$\Leftrightarrow \quad H\,Y\,Y^T\,H^T \preceq H\,G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}\,(G^y)^T\,H^T$$

$$\Leftrightarrow \quad H\,\left(G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}(G^y)^T - Y\,Y^T\right)\,H^T \succeq 0$$

As the matrix on the left hand side of (25) is positive semi-definite, we require that $X \succeq 0$. We further note that H that solves the optimization problem does not necessarily renders $H\,G^y$ invertible. Hence, the rank constraint in (24) also need to be imposed on H. ■

Lemma 2 shows that the optimal measurement combinations for average loss minimization can be found by solving an optimization problem involving matrix inequalities. The matrix inequality (22), however, is quadratic in H and thus is difficult to solve. To derive the solution to this problem, we first present the following Lemma.

Lemma 3 For $A \in \mathbb{R}^{n\times n}$, there exists $B \in \mathbb{R}^{m\times n}$, $m \leq n$, such that $B\,A\,B^T \succeq 0$ if and only if $\lambda_m(A) \geq 0$, where λ_m denotes the m-th largest eigenvalue of A.

The proof of Lemma 3 is embedded in the proof of Proposition 1 in Kariwala[10] and is not repeated here for the sake of brevity. Based on Lemma 3, we present the optimal measurement combinations for average less minimization in the next proposition.

Proposition 3 For the optimization problem (21)-(24), it is optimal to select X as

$$X = \left(J_{uu}^{-0.5}\,(G^y)^T\,(Y\,Y^T)^{-1}\,G^y\,J_{uu}^{-0.5}\right)^{-1} \tag{26}$$

and the minimal loss is given as

$$L_{\text{average}} = \frac{1}{6\,(n_d+n_y)}\sum_{i=1}^{n_u}\frac{1}{\lambda_i\left(J_{uu}^{-0.5}\,(G^y)^T\,(Y\,Y^T)^{-1}\,G^y\,J_{uu}^{-0.5}\right)} \tag{27}$$

Let $\nu_1, \nu_2, \cdots, \nu_{n_u}, \cdots, \nu_{n_y}$ be the mutually orthogonal eigenvectors of $\left(G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}(G^y)^T - Y\,Y^T\right)$ where X is given by (26). Then the optimal H matrix can be selected as

$$H = C\begin{bmatrix} \nu_1 & \nu_2 & \cdots & \nu_{n_u}\end{bmatrix}^T \tag{28}$$

where $C \in \mathbb{R}^{n_u\times n_u}$ is any non-singular matrix.

Proof: Based on Lemma 3, (22) holds if and only if, we select X such that

$$\lambda_{n_u}\left(G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}(G^y)^T - Y\,Y^T\right) \geq 0 \tag{29}$$

Now, let R be an upper-triangular matrix that satisfies $Y\,Y^T = R^T\,R$ (Cholesky factor). Then, similar to the proof of Proposition 2, the first n_u eigenvalues of $\left(G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}(G^y)^T - Y\,Y^T\right)$ are the same as the first n_u eigenvalues of $\left(R^{-T}\,G^y\,J_{uu}^{-0.5}\,X\,J_{uu}^{-0.5}(G^y)^T\,R^{-1} - I\right)$.

Define $Q = R^{-T}\,G^y\,J_{uu}^{-0.5}$. As $X \succeq 0$, $X^{0.5}$ exists. Based on Lemma 1, the first n_u eigenvalues of $\left(Q\,X\,Q^T - I\right)$ and $\left(X^{0.5}\,Q^T\,Q\,X^{0.5} - I\right)$ are the same implying that (29) holds if and only if

$$\begin{aligned} & \lambda_{n_u}\left(X^{0.5}\,Q^T\,Q\,X^{0.5} - I\right) \geq 0 \\ \Leftrightarrow \quad & X^{0.5}\,Q^T\,Q\,X^{0.5} \succeq I \\ \Leftrightarrow \quad & Q^T\,Q \succeq X^{-1} \\ \Leftrightarrow \quad & \left(Q^T\,Q\right)^{-1} \preceq X \end{aligned}$$

As we want to minimize $\mathrm{tr}(X)$, the optimal solution is to choose $X = \left(Q^T\,Q\right)^{-1}$, which gives X in (26). With this choice of X, the expression for average loss follows readily. The proof that (28) represents the set of optimal H is the same as the proof of a similar result in Kariwala[10] and is omitted here. ■

3.3 Worst-case Vs. Average Loss Minimization

In the previous two sections, we presented the optimal combination matrices that minimize worst-case and average losses for local self-optimizing control. When only one degree of freedom is available for self-optimizing control, $\bar{\sigma}\left(\begin{bmatrix} M_d & M_n \end{bmatrix}\right) = \|\begin{bmatrix} M_d & M_n \end{bmatrix}\|_F$ and the combination matrix obtained by minimizing either of the losses is same. When more than one degrees of freedom are available, the designer needs to decide whether to minimize the worst-case or average loss in order to find measurement combinations. Arguably, the minimization of average loss is better, as worst-case may not occur frequently in practice. The minimization of average loss, however, may not always be satisfactory, as average-case optimal design may perform very poorly for certain disturbance and implementation error scenarios. This dilemma is resolved by the following important property of average-case optimal combination matrices:

Proposition 4 Any H matrix that minimizes average loss in (11) also minimizes the worst-case loss in (10).

Proof: We first recall that the optimal combination matrix that minimizes worst-case loss can be obtained by solving the following matrix inequality problem[10]

$$\begin{aligned} & \min_{H} \gamma \\ \text{s.t.} \quad & H \left(\gamma^2 G^y J_{uu}^{-1} (G^y)^T - Y Y^T\right) H^T \succeq 0 \\ & \text{rank}(H) = n_u \end{aligned} \tag{30}$$

Now, let X be chosen as given by (26) and H be any non-singular matrix that satisfies (22). As $H\ G^y\ J_{uu}^{-0.5}$ has full rank, we have

$$\begin{aligned} & X \preceq \bar{\lambda}(X) \\ \Leftrightarrow \quad & (H\ G^y\ J_{uu}^{-0.5})\, X\, (H\ G^y\ J_{uu}^{-0.5})^T \preceq \bar{\lambda}(X)\, (H\ G^y\ J_{uu}^{-0.5})\, (H\ G^y\ J_{uu}^{-0.5})^T \\ \Leftrightarrow \quad & (H\ G^y\ J_{uu}^{-0.5})\, X\, (H\ G^y\ J_{uu}^{-0.5})^T - H\, Y\, Y^T\, H^T \preceq \bar{\lambda}(X)\, (H\ G^y\ J_{uu}^{-0.5})\, (H\ G^y\ J_{uu}^{-0.5})^T - H\, Y\, Y^T\, H^T \\ \Leftrightarrow \quad & H \left(G^y\ J_{uu}^{-0.5}\, X\, J_{uu}^{-0.5} (G^y)^T - Y\, Y^T\right) H^T \preceq H \left(\bar{\lambda}(X)\, G^y\ J_{uu}^{-1}\, (G^y)^T - Y\, Y^T\right) H^T \end{aligned}$$

As the matrix on the left hand side of the last inequality is positive semi-definite, we have

$$H \left(\bar{\lambda}(X)\, G^y\ J_{uu}^{-1}\, (G^y)^T - Y\, Y^T\right) H^T \succeq 0$$

Thus, the average-case optimal H satisfies (30) with $\gamma^2 = \bar{\lambda}(X)$. Based on (20), we note that $\bar{\lambda}^{0.5}(X)$ represents the smallest value of γ for which (30) holds and thus average-case optimal H also minimizes worst-case loss. ■

Based on Proposition 4, the matrix H in (28) is super-optimal. Note that the converse is not true, *i.e.* any matrix H that minimizes the worst-case loss does not necessarily minimize the average-case loss; see

Table 3 for examples. This behavior can be understood by noting the relationship between the optimization problems involving worst-case and average loss minimization. The worst-case loss depends on the largest singular value of the matrix $\begin{bmatrix} M_d & M_n \end{bmatrix}$, while the average loss depends on the Frobenius norm or sum of the squares of all the singular values (including the largest singular value) of the same matrix. It is understandable that although the worst-case optimal combination matrix minimizes the largest singular value, it does not necessarily minimize the contribution of the smaller singular values towards the average loss. However, Proposition 4 implies that at least one of solutions to the worst-case loss minimization, which is not included in the solution set (19), is super-optimal to both minimization problems. This leaves the determination of all worst-case optimal H matrices an interesting open problem. In the present context, however, the designer can simply choose the combination matrix by minimizing average loss and be assured that the worst-case loss will no larger than that can be obtained by using combination matrix found by direct minimization of worst-case loss.

4 Evaporator Case Study

[Figure 1 about here.]

Problem description. The optimal measurement combination design approach is applied to a slightly modified version (see Remark 2) of the evaporation process of Newell and Lee.[14] This is a "forced-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 2 and model equations are given in Appendix A.

The economic objective is to maximize the operational profit [$/h], formulated as a minimization problem of the negative profit (31). The first three terms of (31) are operational costs relating to steam, water and

pumping.[8,17] The fourth term is the raw material cost whilst the last term is the product value.

$$J = 600F_{100} + 0.6F_{200} + 1.009(F_2 + F_3) + 0.2F_1 - 4800F_2 \tag{31}$$

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

$$X_2 \geq 35 + 0.5\% \tag{32}$$

$$40 \text{ kPa} \leq P_2 \leq 80 \text{ kPa} \tag{33}$$

$$P_{100} \leq 400 \text{ kPa} \tag{34}$$

$$0 \text{ kg/min} \leq F_{200} \leq 400 \text{ kg/min} \tag{35}$$

$$0 \text{ kg/min} \leq F_1 \leq 20 \text{ kg/min} \tag{36}$$

$$0 \text{ kg/min} \leq F_3 \leq 100 \text{ kg/min} \tag{37}$$

Note that a 0.5% back-off has been enforced on X_2 to ensure that the variable remains feasible for all possible disturbances. The process model has three state variables, L_2, X_2 and P_2 with eight degrees of freedom. Three of them are disturbances, X_1, T_1 and T_{200}. The rest five degrees of freedom are manipulated variables, F_1, F_2, P_{100}, F_3 and F_{200}. The case with $X_1 = 5\%$, $T_1 = 40°C$ and $T_{200} = 25°C$ is taken as the nominal operating point. The allowable disturbance set corresponds to $\pm 5\%$ variation in X_1 and $\pm 20\%$ variation in T_1 and T_{200} of their nominal values.

Nominal operating point. The optimization problem in (31) with process constraints (32)-(37) is solved for the nominal disturbances. The minimum negative profit obtained is -582.233 [\$/h] and the corresponding values of process variables are shown in Table 2.

[Table 2 about here.]

Degrees of freedom analysis. At the optimal point, there are two active constraints, $X_2 = 35.5\%$ and $P_{100} = 400$ kPa. These two constraints remain active within the whole disturbance region. The reader is

referred to Cao[2] for physical explanation of these two active constraints. The 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, there are two degrees of freedom left for self-optimizing control. Without loss of generality, we select

$$u = \begin{bmatrix} F_{200} & F_1 \end{bmatrix}^T$$

Remark 2 In original problem,[14] F_1 is specified to be a disturbance. Then, there is only one degree of freedom available for self-optimizing control[2] and the optimal combination matrices for local worst-case and average loss minimization are same. The use of F_1 as an additional degree of freedom allows us to clearly demonstrate the advantages of average loss minimization. For this purpose, the prices of raw material and product are included in the cost function (31), where the prices are chosen such that the nominal operating point after modification is similar to the original problem.[14] We have also restricted the allowable range of X_1 to $\pm 5\%$ of nominal value, as compared to $\pm 20\%$ variation allowed in previous studies.[2,5] This is done to avoid scenarios, where the optimal operational policy requires shutting down the production ($F_1 = 0$) for low X_1 values for the modified problem.

Available measurements. For self-optimizing control, we consider that the 2 CVs are to be chosen as a subset or combination of the following available measurements:

$$y = \begin{bmatrix} P_2 & T_2 & T_3 & F_2 & F_{100} & T_{201} & F_3 & F_5 & F_{200} & F_1 \end{bmatrix}^T$$

Here, in addition to the derived variables Q_{100} and Q_{200}, F_4 is not included in the measurement set, as the cost of measuring the vapor flowrate online is usually high. The steam temperature T_{100} is also not considered, as it is a function of steam pressure P_{100}, which is held constant at 400 kPa (active constraint control). All the disturbances are considered to be unmeasured.

Local analysis. Using MATLAB® 2006b symbolic toolbox,[12] the following Hessian and gain matrices are obtained at the nominally optimal operating point:

$$J_{uu} = \begin{bmatrix} 0.006 & -0.133 \\ -0.133 & 16.737 \end{bmatrix}; J_{ud} = \begin{bmatrix} 0.023 & 0 & -0.001 \\ -158.373 & -1.161 & 1.484 \end{bmatrix}$$

$$G^y = \begin{bmatrix} -0.093 & 11.678 \\ -0.052 & 6.559 \\ -0.047 & 5.921 \\ 0 & 0.141 \\ -0.001 & 1.115 \\ -0.094 & 2.170 \\ -0.032 & 6.594 \\ 0 & 0.859 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}; G_d^y = \begin{bmatrix} -3.626 & 0 & 1.972 \\ -2.036 & 0 & 1.108 \\ -1.838 & 0 & 1 \\ 0.267 & 0 & 0 \\ -0.317 & -0.018 & 0.020 \\ -0.674 & 0 & 1 \\ -2.253 & -0.066 & 0.673 \\ -0.267 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Based on the allowable disturbance set, W_d is chosen as $W_d = \text{diag}(0.25, 8, 5)$. The implementation or measurement errors for the pressure and flow measurements are taken to be $\pm 2.5\%$ and $\pm 2\%$, respectively, of the nominal operating value. For temperature measurements, implementation error is considered as $\pm 1^\circ C$. These values are adapted from Govatsmark and Skogestad,[5] except that the implementation error for flowrates has been changed from an unrealistically high value of 10% to 2%. This leads to $W_n = \text{diag}(1.285,\ 1,\ 1,\ 0.027,\ 0.189,\ 1,\ 0.494,\ 0.163,\ 4.355,\ 0.189)$.

For this model, the best individual measurements were found to be

$$c_2 = \begin{bmatrix} F_3 & F_{200} \end{bmatrix}^T$$

for which the local worst-case and average losses are 56.713 [\$/h] and 3.808 [\$/h], respectively. In comparison, when the optimal combinations of all the 10 measurements is used, the local worst-case and average

losses decrease to 7.474 [$/h] and and 0.193 [$/h], respectively. This shows that the loss can be reduced substantially by controlling measurement combinations.

[Figure 2 about here.]

[Table 3 about here.]

In practice, use of combinations of all available measurements is often not necessary. The lowest worst-case and average losses obtained by combining best n_y out of 10 measurements, $n_y \leq 10$, are shown in Figure 2. It is clear that having combinations of 3 or 4 measurements gives a good trade off between complexity and incurred loss. Some individual measurements and measurement subsets of sizes 3 and 4 that can be combined to yield promising self-optimizing variables (found by enumeration) are shown in Table 3. The combinations of 3 measurements that give minimal worst-case and average losses are:

$$c_{3w} = \begin{bmatrix} 0.108\,F_2 + 0.689\,F_{100} - 0.717\,F_{200} \\ 0.988\,F_2 - 0.155\,F_{100} - 0.0001\,F_{200} \end{bmatrix}$$

$$c_{3a} = \begin{bmatrix} 0.157\,F_2 + 0.986\,F_{100} - 0.044\,F_{200} \\ 0.988\,F_2 - 0.158\,F_{100} \end{bmatrix}$$

Similarly, the combinations of 4 measurements that give minimal worst-case and average losses are:

$$c_{4w} = \begin{bmatrix} 0.020\,F_2 + 0.199\,F_{100} + 0.980\,T_{201} + 0.020F_3 \\ 0.994\,F_2 - 0.110\,F_{100} + 0.002\,T_{201} - 0.009F_3 \end{bmatrix}$$

$$c_{4a} = \begin{bmatrix} -0.999\,F_2 - 0.018\,T_{201} + 0.034\,F_3 - 0.001\,F_{200} \\ 0.038\,F_2 - 0.473\,T_{201} + 0.856\,F_3 - 0.203\,F_{200} \end{bmatrix}$$

Based on Table 3, the reader should note the following:

1. The average loss obtained by using average-case optimal H can be substantially lower than obtained by using worst-case optimal H, *e.g.* for 3-measurement subset $\{F_2, F_{100}, F_{200}\}$. The worst-case loss using H matrix obtained by minimizing either of worst-case or average loss is the same.

2. While for $n_y = 2$ and 3, the measurement subset that provide minimal worst-case loss is same as the measurement subset with minimal average loss, this is not true in general. For example, for $n_y = 4$, combining F_2, F_{100}, T_{201} and F_3 gives minimal worst-case loss. The average loss obtained by combining these measurements, however, is higher than that can be obtained by combining F_2, T_{201}, F_3 and F_{200}.

3. The set of best individual measurements $\{F_3, F_{200}\}$ is not a subset of optimal 3-measurement set $\{F_2, F_{100}, F_{200}\}$. A sequential approach to find the 3-measurement set that minimizes worst-case loss would lead to the solution $\{F_2, F_3, F_{200}\}$, which gives a much higher worst-case loss ($27.573\,[\$/h]$) as compared to the optimal solution ($11.636\,[\$/h]$). A branch and bound method is currently being developed to find the optimal measurement set efficiently.

Verification using nonlinear model. The five designs ($c_2, c_{3w}, c_{3a}, c_{4w}, c_{4a}$) are verified using nonlinear model. Here, a difficulty is that although linear analysis assumes that the set of active constraints do not change with disturbances, the constraints on P_2 become active for many disturbance and implementation error scenarios. Usually, the self-optimizing variables are held constant at their nominal operating value. To ensure that the constraints are satisfied for all allowable disturbances and implementation errors, one may use different setpoints.[6] An alternative is to use a cascade control strategy, where the variable liable to violate a constraint (P_2 in the present case) is controlled in the inner loop and the self-optimizing variable is controlled in the outer loop; see Cao[2] for details. In this paper, the latter strategy is used to maintain P_2 within the prescribed constraints with the setpoints being the nominal operating values. Using this method, the losses for different candidate self-optimizing variables are evaluated in two different ways:

1. By setting different disturbances and implementation errors to their lower or upper limits one at a time (results shown in Table 4); and

2. Using a set of 100 randomly generated disturbances and implementation errors, which are uniformly distributed over their allowable set (results shown in Table 5).

[Table 4 about here.]

From Table 4, we note that individual measurements (c_2) are most sensitive to changes in disturbance X_1, while most of the measurement combinations are most sensitive to the changes in implementation errors. This behavior can be analyzed based on local analysis by comparing the magnitudes of the elements of input singular vector corresponding to the largest singular value of the matrix $[M_d \quad M_n]$. For example, for c_2 the relevant input singular vector is $[0.977\ 0.209\ -0.010\ -0.026\ -0.007]^T$, where the large first element shows sensitivity to changes in X_1.

[Table 5 about here.]

Table 5 shows that use of measurement combinations as controlled variables can reduce the loss significantly, as compared to the control of the individual measurements. For 3-measurement combinations (c_{3w} and c_{3a}), use of average-case optimal H provides lower average loss than the use of worst-case optimal H, while the worst-case losses for the two cases are nearly the same. Contrary to local analysis, however, the losses obtained by controlling 4-measurement combinations (c_{4w} and c_{4a}) are worse than the corresponding losses seen with the use of the 3-measurement combinations, which is analyzed next.

Effect of modeling error due to linearization. We note that both of c_{4w} and c_{4a} use F_3. As found using Monte-Carlo simulations, among all candidate measurements the linear model for F_3 shows the largest modeling error due to linearization. Local analysis (see Table 3) shows that a promising

4-measurement set that does not contain F_3 is $\{F_2, F_{100}, F_5, F_{200}\}$, whose worst-case and average-case optimal combinations are given as:

$$\tilde{c}_{4w} = \begin{bmatrix} -0.996\,F_2 + 0.079\,F_{100} + 0.030\,F_5 + 0.002\,F_{200} \\ 0.070\,F_2 + 0.586\,F_{100} + 0.806\,F_5 - 0.055\,F_{200} \end{bmatrix}$$

$$\tilde{c}_{4a} = \begin{bmatrix} 0.117\,F_2 + 0.749\,F_{100} + 0.586\,F_5 - 0.285\,F_{200} \\ 0.992\,F_2 - 0.104\,F_{100} - 0.064\,F_5 \end{bmatrix}$$

The worst-case and average losses for $\tilde{c}_{4w}$ and $\tilde{c}_{4a}$ are smaller than the corresponding losses obtained using c_{3w} and c_{3a}, as expected from local analysis. This substantiates our earlier claim that the poor performance of c_{4w} and c_{4a} is due to the large modeling error incurred by F_3. Furthermore, although the best individual measurement set also contain F_3, control of next best individual measurements $\tilde{c}_2 = [P_2 \quad T_{201}]^T$ gives larger losses than control of c_2. We hypothesize that the computation of locally optimal combination matrices is more sensitive to model inaccuracies than the selection of best measurement subset. In this paper, modeling error is handled in an *ad hoc* fashion and systematic inclusion of the modeling error, arising due to linearization, in local analysis is an issue for future research.

Finally, we recommend the use of $\tilde{c}_{4a}$, which provides lowest losses among the different alternatives and has good self-optimizing properties. If further reduction in the complexity of control structure is desired, c_{3a} can also be used.

5 Conclusions and Future Work

An average loss minimization problem has been proposed to select controlled variables (CVs) based on self-optimizing control principles. Explicit solutions to the problems of finding measurement combinations based on both worst-case and average loss minimization have been derived. In comparison with previous work on worst-case loss minimization based CV selection approaches,[1,7,10] the solution obtained by

minimizing average loss is super-optimal in the sense that the optimal combination matrix based on the average loss minimization is also optimal for the case of worst-case loss minimization. A modified version of the evaporator process has been studied to demonstrate the theoretical results derived in the work. The case study shows that using measurement combinations as CVs can significantly reduce the operational loss, whilst the measurement combinations of a properly selected measurement subset can give similar loss as is achievable by combining full measurement set. The case study also indicates that modeling error due to linearization can have a significant impact on achievable self-optimizing control performance. How to properly deal with such errors for CV selection is an open problem worth for further investigation.

Acknowledgements

The authors gratefully acknowledge the financial support from Office of Finance, Nanyang Technological University, Singapore through grant no. M52120046.

References

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

[2] Y. Cao. Direct and indirect gradient control for static optimisation. *Internationl Journal of Automation and Computing*, 2(1):60–66, 2005.

[3] Y. Cao and D. Biss. An extension of singular value analysis for assessing manipulated variable constraints. *J. Process Control*, 6(1):37–48, 1996.

[4] Y. Cao, D. Rossiter, and D. H. Owens. Globally optimal control structure selection using branch and bound method. In *Proc. 5th International symposium on DYCOPS*, pages 183–188, Corfu, Greece, 1998.

[5] M. S. Govatsmark and S. Skogestad. Control structure selection for an evaporation process. In *Proc. European Symposium on Computer Aided Process Engineering 11*, pages 657–662, Kolding, Denmark, 2001.

[6] M. S. Govatsmark and S. Skogestad. Selection of controlled variables and robust setpoints. *Ind. Eng. Chem. Res.*, 44(7):2207–2217, 2005.

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

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

[9] E. S. Hori and S. Skogestad. Maximum gain rule for selecting controlled variables. In *Proc. 8th Interantional symposium on DYCOPS*, Cancun, Mexico, 2007.

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

[11] V. Kariwala and S. Skogestad. Branch and bound methods for control structure design. In *Proc. 16th ESCAPE and 9th International Symposium on PSE*, Garmisch-Partenkirchen, Germany, 2006.

[12] The MathWorks, 3 Apple Hill Drive, Natick, MA 01760-2098. *MATLAB® Symbolic Math Toolbox User's Guide*, 3.1.5 edition, September 2006.

[13] M. Morari, Y. Arkun, and G. Stephanopoulos. Studies in the synthesis of control structures for chemical process, Part I: Formulation of the problem. Process decomposition and the classification of the control tasks. Analysis of the optimizing control structures. *AIChE J.*, 26(2):220–232, 1980.

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

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

[16] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control: Analysis and Design.* John Wiley & sons, Chichester, UK, 1st edition, 1996.

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

A Model equations

$$\frac{dL_2}{dt} = \frac{F_1 - F_4 - F_2}{20} \tag{38}$$

$$\frac{dX_2}{dt} = \frac{F_1 X_1 - F_2 X_2}{20} \tag{39}$$

$$\frac{dP_2}{dt} = \frac{F_4 - F_5}{4} \tag{40}$$

$$T_2 = 0.5616 P_2 + 0.3126 X_2 + 48.43 \tag{41}$$

$$T_3 = 0.507 P_2 + 55.0 \tag{42}$$

$$F_4 = \frac{Q_{100} - 0.07 F_1 (T_2 - T_1)}{38.5} \tag{43}$$

$$T_{100} = 0.1538 P_{100} + 90.0 \tag{44}$$

$$Q_{100} = 0.16 (F_1 + F_3)(T_{100} - T_2) \tag{45}$$

$$F_{100} = Q_{100}/36.6 \tag{46}$$

$$Q_{200} = \frac{0.9576 F_{200} (T_3 - T_{200})}{0.14 F_{200} + 6.84} \tag{47}$$

$$T_{201} = T_{200} + \frac{13.68 (T_3 - T_{200})}{0.14 F_{200} + 6.84} \tag{48}$$

$$F_5 = \frac{Q_{200}}{38.5} \tag{49}$$

List of Figures

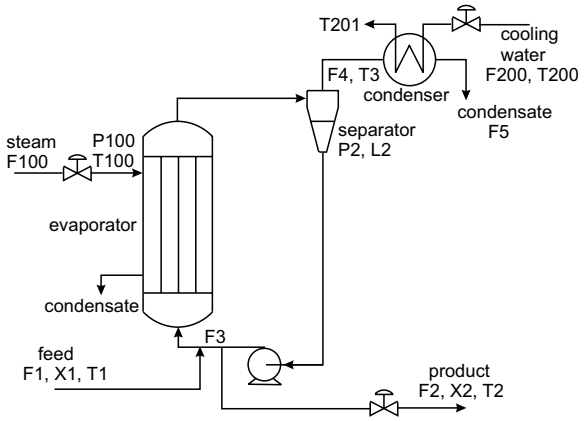

Figure 1: Evaporator System

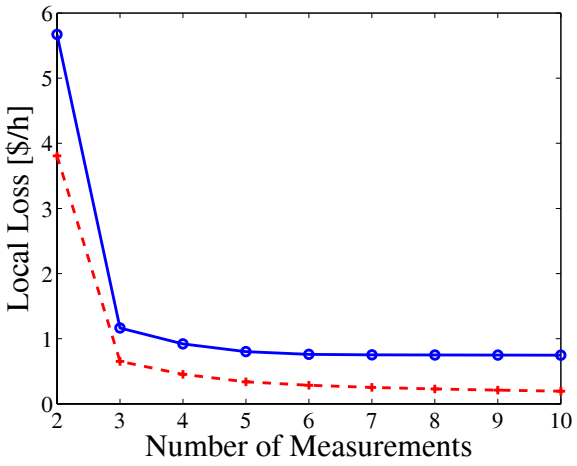

Figure 2: Local losses using combination of best measurements of given subset size. The solid and dotted lines show the variations of $L_{\text{worst}}/10$ and L_{average}, respectively.

List of Tables

Case	Allowable set for d, n	L_{worst}	L_{average}
1	$\left\|\begin{bmatrix} d^T & n^T \end{bmatrix}^T\right\|_2 \leq 1$	$\frac{1}{2}\bar{\sigma}^2\left(\begin{bmatrix} M_d & M_n \end{bmatrix}\right)$	$\frac{1}{6\,(n_y+n_d)}\|\begin{bmatrix} M_d & M_n \end{bmatrix}\|_F^2$
2	$\left\|\begin{bmatrix} d^T & n^T \end{bmatrix}^T\right\|_\infty \leq 1$	$\frac{1}{2}(n_y+n_d)\bar{\sigma}^2\left(\begin{bmatrix} M_d & M_n \end{bmatrix}\right)$	$\frac{1}{6}\|\begin{bmatrix} M_d & M_n \end{bmatrix}\|_F^2$
3	$\left\|\begin{bmatrix} d^T & n^T \end{bmatrix}^T\right\| \sim N(0, I_{n_d+n_y})$	∞	$\frac{1}{2}\|\begin{bmatrix} M_d & M_n \end{bmatrix}\|_F^2$

Table 1: Local worst-case and average losses for different allowable sets of disturbances and implementation errors; For case 2, the indicated expression is an upper bound on worst-case loss.

Var.	Description	Value	Units
F_1	Feed flowrate	9.469	kg/min
F_2	Product flowrate	1.334	kg/min
F_3	Circulating flowrate	24.721	kg/min
F_4	Vapor flowrate	8.135	kg/min
F_5	Condensate flowrate	8.135	kg/min
X_1	Feed composition	5.000	%
X_2	Product composition	35.500	%
T_1	Feed temperature	40.000	°C
T_2	Product temperature	88.400	°C
T_3	Vapor temperature	81.066	°C
L_2	Separator level	1.000	meter
P_2	Operating pressure	51.412	kPa
F_{100}	Steam flowrate	9.434	kg/min
T_{100}	Steam temperature	151.520	°C
P_{100}	Steam pressure	400.000	kPa
Q_{100}	Heat duty	345.292	kW
F_{200}	Cooling water flowrate	217.738	kg/min
T_{200}	Inlet temperature of cooling water	25.000	°C
T_{201}	Outlet temperature of cooling water	45.550	°C
Q_{200}	Condenser duty	313.210	kW

Table 2: Variables and Optimal Values

n_y	Measurements	Worst-case loss [\$/h]	Average loss [\$/h]	
			average-case optimal H	worst-case optimal H
2	F_3, F_{200}	56.713	3.808	3.808
	T_{201}, F_3	57.140	4.330	4.330
	P_2, T_{201}	57.862	4.388	4.388
	F_{100}, F_{200}	58.370	3.900	3.900
	P_2, F_{200}	58.386	3.964	3.964
3	F_2, F_{100}, F_{200}	11.636	0.652	1.238
	F_2, F_{100}, T_{201}	13.327	1.123	1.124
	F_2, T_{201}, F_3	16.619	1.121	1.143
	F_2, F_5, F_{200}	17.797	0.993	1.565
4	F_2, F_{100}, T_{201}, F_3	9.195	0.587	0.793
	F_2, T_{201}, F_3, F_{200}	9.427	0.453	0.701
	F_2, F_{100}, F_5, F_{200}	9.879	0.474	0.845
	F_2, F_{100}, F_3, F_{200}	10.547	0.507	0.799

Table 3: Promising candidates for self-optimizing control based on local analysis

CV	X_1^+	T_1^+	T_{200}^+	n_{c1}^+	n_{c2}^+	X_1^-	T_1^-	T_{200}^-	n_{c1}^-	n_{c2}^-
c_2	53.218	2.476	0.351	0.046	0.044	40.873	2.511	0.360	0.048	0.045
$\tilde{c}_2$	53.521	3.167	9.593	0.123	0.361	41.167	3.193	6.489	0.123	0.398
c_{3w}	10.310	0.712	0.298	0.056	7.912	11.605	0.623	0.302	0.055	19.814
c_{3a}	0.085	0.024	0.046	0.221	9.953	4.570	0.025	0.022	0.202	17.171
c_{4w}	4.071	0.041	8.808	0.404	10.312	4.221	0.008	6.240	0.362	11.505
c_{4a}	11.794	0.438	0.079	0.371	9.364	6.503	0.275	0.258	0.365	5.500
$\tilde{c}_{4w}$	7.450	0.510	0.400	0.075	7.936	10.576	0.581	0.360	0.072	22.653
$\tilde{c}_{4a}$	0.277	0.236	0.313	20.180	3.042	2.326	0.344	0.260	11.676	3.303

Table 4: Losses for candidate self-optimizing variables using nonlinear model. Different disturbances and implementation errors are changed to their lower (denoted by superscript $-$) or upper (denoted by superscript $+$) limits one at a time. For measurement combinations, the implementation error is calculated as $n_c = \sum_{i=1}^{n_y} |HW_n n|$.

CV	Average loss	Maximum loss	Standard deviation of loss
c_2	17.181	55.163	15.791
$\tilde{c}_2$	20.536	70.041	17.095
c_{3w}	6.792	25.883	5.451
c_{3a}	4.041	25.785	4.791
c_{4w}	6.825	22.338	5.433
c_{4a}	4.981	51.569	6.499
$\tilde{c}_{4w}$	5.322	19.550	4.393
$\tilde{c}_{4a}$	2.808	19.353	3.418

Table 5: Loss statistics for candidate self-optimizing variables using nonlinear model with 100 uniformly distributed disturbances and implementation errors (simultaneous changes)