Optimal measurement combinations as controlled variables

Vidar Alstad [1], Sigurd Skogestad [*] and Eduardo S. Hori

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

Abstract

This paper deals with the optimal selection of linear measurement combinations as controlled variables, $\mathbf{c} = \mathbf{H}\mathbf{y}$. The objective is to achieve "self-optimizing control", which is when fixing the controlled variables $\mathbf{c}$ indirectly gives near-optimal steady-state operation with a small loss. The nullspace method of Alstad and Skogestad [1] focuses on minimizing the loss caused by disturbances. We here provide an explicit expression for $\mathbf{H}$ for the case where the objective is to minimize the combined effect of disturbances and measurement errors. In addition, we extend the nullspace method to cases with extra measurements by using the extra degrees of freedom to minimize the loss caused by measurement errors. Finally, the results are interpreted more generally as deriving linear invariants for quadratic optimization problems.

1 Introduction

Optimizing control is an old research topic, and one method for ensuring optimal operation in chemical processes is real-time optimization (RTO)[10]. Using RTO, the optimal values (setpoints) for the controlled variables $\mathbf{c}$ are recomputed online based on online measurements and a model of the process, see Figure 1. In most RTO-applications, a steady-state model is used for the reconciliation (parameter/disturbance estimation) and optimization steps [20, 19], however dynamic versions of the RTO-framework have also been reported [7]. However, the cost of installing and maintaining RTO systems can be large. In addition, the system can be sensitive to uncertainty.

[*] Author to whom correspondence should be addressed: e-mail: skoge@chemeng.ntnu.no, Tel: +47 73 59 41 54, Fax.: +47 73 59 40 80

[1] Current affiliation: Norsk Hydro ASA, Oil & Energy Research Centre, N-3908 Porsgrunn, Norway

A completely different approach to optimizing control is to focus on selecting the right variables $\mathbf{c}$ to control, which is the idea of "self-optimizing control"[13]. The objective is to search for combinations of measurements, for example, linear combinations $\mathbf{c} = \mathbf{H}\mathbf{y}$, which when controlled, will (indirectly) keep the process close to the optimum operating conditions despite disturbances and measurement errors. The need for a RTO layer to compute new optimal setpoints $\mathbf{c}_s$ can then be reduced, or in many cases even eliminated. Thus, the implementation is trivial and the maintenance requirements are minimized. The idea in this paper is extend this approach, by providing explicit formulas for the optimal matrix $\mathbf{H}$.

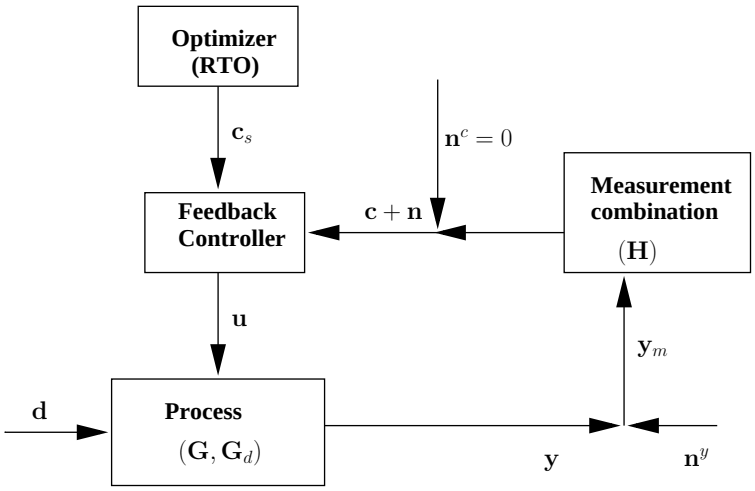

Fig. 1. Feedback implementation of optimal operation with separate layers for optimization (RTO) and control.

The issue of selecting $\mathbf{H}$ can also be viewed as a "squaring down" problem, as illustrated in Figure 2. The number of output variables that can be independently controlled (n_c) is equal to the number of independent inputs (n_u), but in most cases the number of available measurements (n_y) is larger, that is, $n_y > n_u$. The issue is then to select the nonsquare matrix $\mathbf{H}$ such that the map (transfer function) $\mathbf{G} = \mathbf{H}\mathbf{G}^y$ from $\mathbf{u}$ to $\mathbf{c}$ is square, see Figure 2. However, selecting $\mathbf{H}$ such that $\mathbf{G}$ is square is not the only issue. More importantly, as mentioned above, control of $\mathbf{c}$ should (directly or indirectly) result in "acceptable operation" of the system.

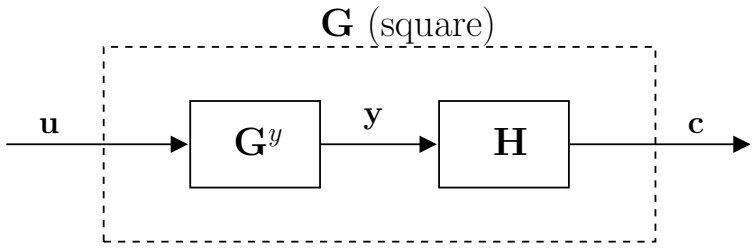

Fig. 2. Combining measurements $\mathbf{y}$ to get controlled variables $\mathbf{c}$ (linear case).

To quantify "acceptable operation" we introduce a scalar cost function J which should be minimized for optimal operation. In this paper, we assume that the (economic) cost mainly depends on the (quasi) steady-state behavior, which is a good assumption for most continuous plants in the process industry.

Self-optimizing control [13] can now be defined. It is when a constant set-point policy ($\mathbf{c}_s$ constant) yields acceptable loss, $L = J(\mathbf{u}, \mathbf{d}) - J^{\mathrm{opt}}(\mathbf{d})$, in spite of the presence of uncertainty, which is here assumed to be represented by (1) *external disturbances* $\mathbf{d}$ and (2) *implementation errors* $\mathbf{n} \triangleq \mathbf{c}_s - \mathbf{c}$, see Figure 1.

The implementation error $\mathbf{n}$ has two sources, (1) the steady-state control error $\mathbf{n}^c$ and (2) the measurement error $\mathbf{n}^y$; and for linear measurement combinations $\mathbf{n} = \mathbf{n}^c + \mathbf{H}\mathbf{n}^y$. In Figure 1, the control error $\mathbf{n}^c$ is shown as an exogenous signal, although in reality it is determined by the controller. In any case, we assume here that all controllers have integral action, so we can neglect the steady-state control error, i.e. $\mathbf{n}^c = 0$. The implementation error $\mathbf{n}$ is then given by the measurement error, i.e. $\mathbf{n} = \mathbf{H}\mathbf{n}^y$.

Ideas related to self-optimizing control have been presented repeatedly in the process control literature, but the first quantitative treatment was that of Morari et al. [11]. Skogestad [13] defined the problem more carefully, linked it to previous work, and was the first to include also the implementation error. He mainly considered the case where single measurements are used as controlled variables, that is, $\mathbf{H}$ is a selection matrix where each row has a single 1 and the rest 0's. Halvorsen et al. [3] considered the approximate "maximum gain method" and also proposed an "exact local method" for the optimal measurement combination $\mathbf{H}$. They proposed to obtain $\mathbf{H}$ numerically by solving $\min_{\mathbf{H}} \bar{\sigma}[\mathbf{M}(\mathbf{H})]$, but did not say anything about the properties of this optimization problem. Kariwala [8] proposed an iterative method involving singular value and eigenvalue decomposition. Hori et al. [5] considered indirect control, which can be formulated as a subproblem of the extended null space method presented in this paper. Additional related work is presented in [15, 16, 17] on measurement-based optimization to enforce the necessary condition of optimality under uncertainty, with application to batch processes. Bonvin et al. [2] extends these ideas and focus on steady-state optimal systems, where a clear distinction is made between enforcing active constraints and requiring the sensitivity of the objective to be zero.

This paper is an extension of the nullspace method of [1], where it was found that, in the absence of implementation errors (i.e., $\mathbf{n} = 0$), it is possible to have zero loss with respect to disturbances, provided the the number of (independent) measurements (n_y) is at least equal to the number of (independent) inputs (n_u) and disturbances (n_d), i.e., $n_y \geq n_u + n_d$. It is then optimal to select $\mathbf{H}$ such that $\mathbf{HF} = 0$, where $\mathbf{F} = \mathrm{d}\mathbf{y}^{opt}/\mathrm{d}\mathbf{d}^T$ is the optimal sensitivity

with respect to disturbances $\mathbf{d}$ [1]. Note that it is not possible to have zero loss with respect to implementation errors, because each new measurement adds a "disturbance" through its associated measurement error, $\mathbf{n}^y$.

In this paper, we include the implementation error and provide the following new results:

(1) Optimal $\mathbf{H}$ for combined disturbances and implementation errors (Section 3).
(2) Optimal $\mathbf{H}$ for disturbances using possible extra measurements to minimize the effect of implementation error (Extended null space method, Section 4).

2 Background

The material in this section is based on [3], unless otherwise stated. The most important notation is given in Table 1.

Table 1
Notation.

$\mathbf{u}$	-	vector of n_u unconstrained inputs (degrees of freedom)
$\mathbf{d}$	-	vector of n_d disturbances
$\mathbf{y}$	-	vector of n_y selected measurements used in forming $\mathbf{c}$
$\mathbf{c}$	-	vector of selected controlled variables (to be identified) with dimension $n_c = n_u$
$\mathbf{n}^y$	-	measurement error associated with $\mathbf{y}$
$\mathbf{n}^c$	-	control error associated with $\mathbf{c}$ (this paper: $\mathbf{n}^c = 0$)
$\mathbf{n}$	-	implementation error associated with $\mathbf{c}$; $\mathbf{n} = \mathbf{n}^c + \mathbf{H}\mathbf{n}^y = 0$)

The objective is to achieve optimal steady-state operation, where the degrees of freedom $\mathbf{u}$ are selected such that the scalar cost function $J(\mathbf{u}, \mathbf{d})$ is minimized for any given disturbance $\mathbf{d}$. Parameter variations may also be included as disturbances. We assume that any optimally "active constraints" have been implemented, so that $\mathbf{u}$ includes only the remaining unconstrained steady-state degrees of freedom. The "reduced space" optimization problem then becomes

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

The objective of this work is to find a set of controlled variables $\mathbf{c}$, or more

specifically an optimal measurement combination $\mathbf{c} = \mathbf{H}\mathbf{y}$, such that a constant setpoint policy (where $\mathbf{u}$ is adjusted to keep $\mathbf{c}$ constant; see Figure 1) yields optimal operation (1), at least locally.

With a given $\mathbf{d}$, solving eq. (1) for $\mathbf{u}$ gives $J^{opt}(\mathbf{d})$, $\mathbf{u}^{opt}(\mathbf{d})$ and $\mathbf{y}^{opt}(\mathbf{d})$. In practice it is not possible to have $\mathbf{u} = \mathbf{u}^{opt}(\mathbf{d})$, for example, because of implementations errors and changing disturbances. The resulting loss (L) is defined as the difference between the cost J, when using a non-optimal input $\mathbf{u}$, and $J^{opt}(\mathbf{d})$ [14]:

$$L = J(\mathbf{u}, \mathbf{d}) - J^{opt}(\mathbf{d}) \tag{2}$$

The local second-order accurate Taylor series expansion of the cost function around the nominal point $(\mathbf{u}^*, \mathbf{d}^*)$ can be written

$$J(\mathbf{u}, \mathbf{d}) = J(\mathbf{u}^*, \mathbf{d}^*) + \begin{bmatrix} \mathbf{J}_u & \mathbf{J}_d \end{bmatrix}^T \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix}^T \begin{bmatrix} \mathbf{J}_{uu} & \mathbf{J}_{ud} \\ \mathbf{J}_{ud}^T & \mathbf{J}_{dd} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} \tag{3}$$

where $\Delta\mathbf{u} = (\mathbf{u} - \mathbf{u}^*)$ and $\Delta\mathbf{d} = (\mathbf{d} - \mathbf{d}^*)$. For a given disturbance ($\Delta\mathbf{d} = 0$), the second-order accurate expansion of the loss function around the optimum ($\mathbf{J}_u = 0$) then becomes

$$L = \frac{1}{2}(\mathbf{u} - \mathbf{u}^{opt})^T \mathbf{J}_{uu} (\mathbf{u} - \mathbf{u}^{opt}) = \frac{1}{2}\mathbf{z}^T\mathbf{z} \tag{4}$$

where

$$\mathbf{z} \triangleq \mathbf{J}_{uu}^{1/2}(\mathbf{u} - \mathbf{u}^{opt}) \tag{5}$$

In this paper, we consider a constant setpoint policy where the controlled variables are linear combinations of the measurements [2]

$$\Delta\mathbf{c} = \mathbf{H}\Delta\mathbf{y} \tag{6}$$

We assume that $n_c = n_u$, that is, the number of (independent) controlled variables $\mathbf{c}$ is equal to the number of (independent) steady-state degrees of freedom ("inputs") $\mathbf{u}$. The constant setpoint policy implies that $\mathbf{u}$ is adjusted to give $\mathbf{c}_s = \mathbf{c} + \mathbf{n}$ where $\mathbf{n}$ is the implementation error for $\mathbf{c}$ (see Figure 1). As mentioned in the introduction, we assume that the implementation error is caused be the measurement error, i.e. $\mathbf{n} = \mathbf{H}\mathbf{n}^y$. We now want to express the loss error $\mathbf{z}$ in terms $\mathbf{d}$ and $\mathbf{n}^y$ when we use a constant setpoint policy, but first some additional notation is needed.

[2] We use Δ to denote deviation variables. Often, the Δ is omitted and we write, for example, $\mathbf{c} = \mathbf{H}\mathbf{y}$.

The linearized (local) model in terms of deviation variables is written

$$\Delta \mathbf{y} = \mathbf{G}^y \Delta \mathbf{u} + \mathbf{G}_d^y \Delta \mathbf{d} = \tilde{\mathbf{G}}^y \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{d} \end{bmatrix} \tag{7}$$

$$\Delta \mathbf{c} = \mathbf{G} \Delta \mathbf{u} + \mathbf{G}_d \Delta \mathbf{d} \tag{8}$$

where

$$\tilde{\mathbf{G}}^y = \begin{bmatrix} \mathbf{G}^y & \mathbf{G}_d^y \end{bmatrix} \tag{9}$$

is the augmented plant. From eqs. (6), (7) and (8) we get

$$\mathbf{G} = \mathbf{H}\mathbf{G}^y \quad \text{and} \quad \mathbf{G}_d = \mathbf{H}\mathbf{G}_d^y \tag{10}$$

The magnitudes of the disturbances $\mathbf{d}$ and measurement errors $\mathbf{n}^y$ are quantified by the diagonal scaling matrices $\mathbf{W}_d$ and $\mathbf{W}_{n^y}$, respectively. More precisely, we write

$$\Delta \mathbf{d} = \mathbf{W}_d \mathbf{d}' \tag{11}$$

$$\mathbf{n}^y = \mathbf{W}_{n^y} \mathbf{n}^{y\prime} \tag{12}$$

where we assume that $\mathbf{d}'$ and $\mathbf{n}^{y\prime}$ are any vectors satisfying

$$\| \begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^{y\prime} \end{bmatrix} \|_2 \leq 1 \tag{13}$$

A justification for using the combined 2-norm in eq. (13) is given in the discussion section of Halvorsen et al. [3].

The non-linear functions $\mathbf{u}^{opt}(\mathbf{d})$ and $\mathbf{y}^{opt}(\mathbf{d})$ are also linearized, and it can be shown that [3]

$$\Delta \mathbf{u}^{opt} = -\mathbf{J}_{uu}^{-1} \mathbf{J}_{ud} \Delta \mathbf{d} \tag{14}$$

$$\Delta \mathbf{y}^{opt} = \underbrace{-(\mathbf{G}^y \mathbf{J}_{uu}^{-1} \mathbf{J}_{ud} - \mathbf{G}_d^y)}_{\mathbf{F}} \Delta \mathbf{d} \tag{15}$$

where we have introduced the optimal sensitivity matrix $\mathbf{F}$ for the measurements. In terms of the controlled variables $\mathbf{c}$ we then have

$$(\mathbf{u} - \mathbf{u}^{opt}) = \mathbf{G}^{-1}(\mathbf{c} - \mathbf{c}^{opt}) = \mathbf{G}^{-1}(\Delta \mathbf{c} - \Delta \mathbf{c}^{opt}) \tag{16}$$

$$\Delta \mathbf{c}^{opt} = \mathbf{H} \Delta \mathbf{y}^{opt} = \mathbf{H}\mathbf{F} \Delta \mathbf{d} \tag{17}$$

$$\Delta \mathbf{c} = \Delta \mathbf{c}_s - \mathbf{n} = -\mathbf{n} = -\mathbf{H}\mathbf{n}^y \tag{18}$$

where we in the last equation have assumed a constant setpoint policy ($\Delta \mathbf{c}_s = 0$). Upon introducing the magnitudes of $\Delta \mathbf{d}$ and $\mathbf{n}^y$ from eqs. (11) and (12) we then get for the constant setpoint policy:

$$\mathbf{z} = \mathbf{M}_d \mathbf{d}' + \mathbf{M}_{n^y} \mathbf{n}^{y'} \tag{19}$$

where

$$\mathbf{M}_d = -\mathbf{J}_{uu}^{1/2}(\mathbf{H}\mathbf{G}^y)^{-1}\mathbf{H}\mathbf{F}\mathbf{W}_d \tag{20}$$

$$\mathbf{M}_{n^y} = -\mathbf{J}_{uu}^{1/2}(\mathbf{H}\mathbf{G}^y)^{-1}\mathbf{H}\mathbf{W}_{n^y} \tag{21}$$

Introducing

$$\mathbf{M} \triangleq [\mathbf{M_d}\ \mathbf{M_{n^y}}] \tag{22}$$

gives $\mathbf{z} = \mathbf{M}\begin{bmatrix}\mathbf{d}' \\ \mathbf{n}^{y\prime}\end{bmatrix}$, which is the desired expression for the loss error. A nonzero value for $\mathbf{z}$ gives a loss $L = \frac{1}{2}\|\mathbf{z}\|_2$ (4), and the worst-case loss for the expected disturbances and noise in (13) is then[3]

$$L_{wc} = \max_{\|\begin{smallmatrix}\mathbf{d}' \\ \mathbf{n}^{y\prime}\end{smallmatrix}\|_2 \le 1} L = \frac{1}{2}\left(\bar{\sigma}[\mathbf{M}]\right)^2 \tag{23}$$

where the last equality follows from the definition of the singular value $\bar{\sigma}$ and the assumption about the normalized disturbances and measurement errors being 2-norm bounded, see eq. (13). Thus, to minimize the worst-case loss we need to minimize $\bar{\sigma}(\mathbf{M})$ with respect to $\mathbf{H}$. This is the "exact local method" in Halvorsen et al. [3], and note that we have expressed $\mathbf{M}_d$ in (20) in terms of the easily available optimal sensitivity matrix $\mathbf{F}$.

3 Explicit formula for optimal H for combined disturbances and measurement errors

From (23), the optimal measurement combination is obtained by solving the problem ("exact local method")

$$\mathbf{H} = \arg\min_{\mathbf{H}} \bar{\sigma}(\mathbf{M}) \tag{24}$$

It may seem that this optimization problem is non-trivial as $\mathbf{M}$ depends non-linearly on $\mathbf{H}$, as shown in (20)-(22). Halvorsen et al. [3] proposed a numerical solution and Kariwala [8] provides an iterative solution for the optimal $\mathbf{H}$ involving the singular value and eigenvalue decompositions. However, (24) is in fact easy to solve , as shown in the following. We start by introducing

$$\mathbf{M}_n \triangleq \mathbf{J}_{uu}^{1/2}(\mathbf{H}\mathbf{G}^y)^{-1} = \mathbf{J}_{uu}^{1/2}\mathbf{G}^{-1} \tag{25}$$

which may be viewed as the effect of $\mathbf{n}$ on the loss variables $\mathbf{z}$. We then have

$$\mathbf{M} = [\mathbf{M_d}\ \mathbf{M_{n^y}}] = -\mathbf{M}_n\mathbf{H}[\mathbf{F}\mathbf{W}_d\ \ \mathbf{W}_{n^y}] \tag{26}$$

Next, we use the fact that the solution of eq. (24) is not unique, so that if $\mathbf{H}$ is an optimal solution, then another optimal solution is $\mathbf{H}_1 = \mathbf{D}\mathbf{H}$, where $\mathbf{D}$ is a non-singular matrix of dimension $n_u \times n_u$. For example, this follows because $\mathbf{M}_d$ and $\mathbf{M}_{n^y}$ in (20) and (21) are unaffected by the choice of $\mathbf{D}$. One implication is that $\mathbf{G} = \mathbf{H}\mathbf{G}^y$ may be chosen freely (which also is clear from Figure 2 since we may add an output block after $\mathbf{H}$ which allows $\mathbf{G}$ to be selected freely). Alternatively, and this is used here, it follows from (25) that $\mathbf{M}_n$ may be selected freely. However, the fact that $\mathbf{M}_n$ may be selected freely, does not mean that one can, for example, simply set $\mathbf{M}_n = \mathbf{I}$ in (26) and then minimize $\bar{\sigma}(\mathbf{M})$ with $\mathbf{M} = \mathbf{H}[\mathbf{F}\mathbf{W}_d \ \mathbf{W}_{n^y}]$. Rather, one needs to minimize $\bar{\sigma}(\mathbf{M})$ subject to the constraint $\mathbf{M}_n = \mathbf{I}$. Introducing

$$\tilde{\mathbf{F}} \triangleq [\mathbf{F}\mathbf{W}_d \ \mathbf{W}_{n^y}] \tag{27}$$

the optimization problem (24) can then be stated as

$$\mathbf{H} = \arg\min_{\mathbf{H}} \bar{\sigma}(\mathbf{H}\tilde{\mathbf{F}}) \ \text{ subject to } \ \mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2} \tag{28}$$

This is fairly easy to solve numerically because of the linearity in $\mathbf{H}$ in both the matrix $\mathbf{H}\tilde{\mathbf{F}}$ and in the equality constraints. In fact, an explicit a solution may be found, as shown below.

Choice of norm. The optimization problems (24) and (28) involve the singular value (induced 2-norm) of $\mathbf{M}$, $\bar{\sigma}(\mathbf{M})$, which represents the worst-case effect of combined 2-norm bounded disturbances and measurement errors on the loss. A closely related problem is to minimize the Frobenius norm (Euclidean or 2-norm) of $\mathbf{M}$, $\|\mathbf{M}\|_F = \sqrt{\sum_{i,j} |m_{ij}|^2}$, which represents some "average" effect of combined disturbances and measurement errors on the loss. Actually, which norm to use is more a matter of preference or mathematical convenience than of "correctness". First, the difference in minimizing the two norms is generally minor; the main difference is that minimizing $\bar{\sigma}(\mathbf{M})$ usually puts more focus on minimizing the largest elements. Second, as discussed below, it appears that for this particular problem, we have a kind of "super-optimality", where the choice of $\mathbf{H}$ that minimizes $\|M\|_F$, also minimizes $\bar{\sigma}(\mathbf{M})$ [9].

Scalar case. For the scalar case ($\mathbf{c}$ is a scalar and $\mathbf{n}_u = \mathbf{n}_c = 1$), $\mathbf{M}$ and $\mathbf{H}^T$ are (column) vectors, $\bar{\sigma}(\mathbf{M}) = \|\mathbf{M}\|_F$, and an analytic solution to (28) is easily derived. The optimization problem (28) becomes

$$\min_{\mathbf{H}^T} \|\tilde{\mathbf{F}}^T\mathbf{H}^T\|_2 \ \text{ subject to } \ \mathbf{G}^{yT}\mathbf{H}^T = \mathbf{J}_{uu}^{1/2} \tag{29}$$

and from standard results for constrained quadratic optimization, the optimal solution is (see proof in Appendix)

$$\mathbf{H}^T = (\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\mathbf{G}^y(\mathbf{G}^{yT}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\mathbf{G}^y)^{-1}\mathbf{J}_{uu}^{1/2} \tag{30}$$

where it is assumed that $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T$ has full rank.

General case. The explicit expression for $\mathbf{H}$ in (31) holds also for the general case, that is, for minimizing $\|\mathbf{M}\|_F$ for the case when $\mathbf{H}$ is a matrix. This can be proved by rewriting the general optimization problem (28) for the matrix case, into a vector problem by stacking the column of $\mathbf{H}^T$ into a long vector. In addition, Kariwala et al. [9] has shown, as already mentioned, that the matrix $\mathbf{H}$ that minimizes the Frobenius-norm of $\mathbf{M}$ also minimizes the singular value of $\mathbf{M}$ [9]. However, the reverse does not necessarily hold, that is, a solution that minimizes $\bar{\sigma}(\mathbf{M})$ does not necessarily minimize $\|\mathbf{H}\|_F$ [9], which is because the solution to the problem of minimizing $\bar{\sigma}(\mathbf{M})$ is not unique. Our findings can be summarized in the following Theorem (see the Appendix for details).

Theorem 1 *For combined disturbances and measurement errors, the optimal measurement combination problem in terms of the Frobenius-norm,* $\min_{\mathbf{H}} \|\mathbf{M}\|_F$ *with* $\mathbf{M}$ *given by (20)-(22), can be reformulated as* $\min_{\mathbf{H}} \|\mathbf{H}\tilde{\mathbf{F}}\|_F$ *subject to* $\mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2}$, *where* $\tilde{\mathbf{F}} = [\mathbf{F}\mathbf{W}_d \ \mathbf{W}_{n^y}]$. *Assuming* $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T$ *is full rank, we have the following explicit solution for the combination matrix* $\mathbf{H}$,

$$\mathbf{H}^T = (\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\mathbf{G}^y(\mathbf{G}^{yT}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\mathbf{G}^y)^{-1}\mathbf{J}_{uu}^{1/2} \tag{31}$$

This solution also minimizes the singular value of $\mathbf{M}$, $\bar{\sigma}(\mathbf{M})$, *that is, provides the solution to the "exact local method" in (24).*

Note that $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T = [\mathbf{F}\mathbf{W}_d \ \mathbf{W}_{n^y}][\mathbf{F}\mathbf{W}_d \ \mathbf{W}_{n^y}]^T$ in (31) needs to be full rank. This implies that (31) does not generally apply to the case with no measurement error, $\mathbf{W}_{n^y} = 0$, but otherwise the expression for $\mathbf{H}$ applies generally for any number n_y of measurements $\mathbf{y}$. One exception, when the expression for $\mathbf{H}$ in (31) applies also for $\mathbf{W}^y = 0$, is when $n_y \leq n_d$, because $\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T$ then remains full rank.

4 Extended nullspace method

The solution for $\mathbf{H}$ in (31) minimizes the loss with respect to combined disturbances and measurements errors. An alternative approach is to first minimize the loss with respect to disturbances, and then, if there are remaining degrees of freedom, minimize the loss with respect to measurement errors. One justification is that disturbances are the reason for introducing optimization and feedback in the first place. Another justification is that it may be easier later to reduce measurements errors than to reduce disturbances.

If we neglect the implementation error ($\mathbf{M}_{n^y} = 0$), then we see from (20) that $\mathbf{M}_d = 0$ (zero loss) is obtained by selecting selecting $\mathbf{H}$ such that

$$\mathbf{H}\mathbf{F} = 0 \tag{32}$$

This provides an alternative derivation of the nullspace method of [1]. It is always possible to find a non-trivial solution (i.e. $\mathbf{H} \neq 0$) $\mathbf{H}$ satisfying $\mathbf{HF} = 0$ provided the number of independent measurements (n_y) is greater than the number of independent inputs (n_u) and disturbances (n_d), i.e. $n_y \geq n_u + n_d$ [1]. One solution is to select $\mathbf{H}$ as the nullspace of $\mathbf{F}^T$:[1]

$$\mathbf{H} = \mathcal{N}(\mathbf{F}^T) \tag{33}$$

The main disadvantage with the nullspace method is that we have no control of the loss caused by measurement errors as given by the matrix $\mathbf{M}_n^y$. In this section, we study this in more detail, by deriving an explicit expression for $\mathbf{H}$ that allows us to compute the resulting $\mathbf{M}_{n^y}$. The explicit expression for $\mathbf{H}$ allows us to extend the nullspace method to cases with extra or too few measurements, i.e., to cases when $n_y \neq n_u + n_d$.

4.1 Explicit expression for $\mathbf{H}$ *for original null space method*

From the expansion of the loss function we have, see eqs. (5) and (14)

$$\mathbf{z} = \overbrace{\begin{bmatrix} \mathbf{J}_{uu}^{1/2} & \mathbf{J}_{uu}^{1/2}\mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \end{bmatrix}}^{\tilde{\mathbf{J}}} \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} \tag{34}$$

We assume that $\mathbf{H}$ is selected to have zero disturbance loss, which is possible if $n_y \geq n_u + n_d$. Then from (19) and (26), $\mathbf{z} = -\mathbf{M}_n\mathbf{H}\mathbf{n}^y$. With the controlled variables $\mathbf{c} = \mathbf{H}\mathbf{y}$ fixed at constant setpoints ($\Delta\mathbf{c} = \Delta\mathbf{c}_s = 0$) we then have $\Delta\mathbf{y} = -\mathbf{n}^y$, and get

$$\mathbf{z} = -\mathbf{M}_n\mathbf{H}\mathbf{n}^y = \mathbf{M}_n\mathbf{H}\Delta\mathbf{y} = \mathbf{M}_n\mathbf{H}\tilde{\mathbf{G}}^y \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} \tag{35}$$

where $\tilde{\mathbf{G}}^y = \begin{bmatrix} \mathbf{G}^y & \mathbf{G}_d^y \end{bmatrix}$ is the augmented plant. Comparing eqs. (34) and (35) yields

$$\mathbf{M}_n\mathbf{H}\tilde{\mathbf{G}}^y = \tilde{\mathbf{J}} \tag{36}$$

where $\tilde{\mathbf{J}}$ is defined in (34). We then have the following explicit expression for $\mathbf{H}$ for the case where $n_y = n_u + n_d$ such that $\tilde{\mathbf{G}}^y$ is invertible

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}[\tilde{\mathbf{G}}^y]^{-1} \tag{37}$$

This expression gives $\mathbf{H}$ for a case with zero disturbance sensitivity ($\mathbf{M}_d = 0$), and thus gives the same result as (33). Note that $\mathbf{M}_n$ can be regarded as a "free" parameter (e.g. we may set $\mathbf{M}_n = \mathbf{I}$, see Remark 2 below).

The explicit solution for $\mathbf{H}$ in (37) forms the basis for extending the nullspace method to cases where we have extra measurements ($n_y > n_u + n_d$) or too few measurements ($n_y < n_u + n_d$).

Assume that we have n_u independent unconstrained free variables $\mathbf{u}$, n_d disturbances $\mathbf{d}$, n_y measurements $\mathbf{y}$, and we want to obtain $n_c = n_u$ independent controlled variables $\mathbf{c}$ that are linear combinations of the measurements, $\mathbf{c} = \mathbf{H}\mathbf{y}$. From the results in Section 2, the loss imposed by a constant setpoint policy is $L = \frac{1}{2}\mathbf{z}^T\mathbf{z}$ where $\mathbf{z} = \mathbf{M}_d\mathbf{d}' + \mathbf{M}_{n^y}\mathbf{n}^{y'}$. Define $\mathbf{E}$ as the error in satisfying eq. (36):

$$\mathbf{E} = \mathbf{M}_n\mathbf{H}\tilde{\mathbf{G}}^y - \tilde{\mathbf{J}} \tag{38}$$

We want to derive a relationship between $\mathbf{E}$ and $\mathbf{M}_d$. From (15) and (9) the optimal sensitivity can be written

$$\mathbf{F} = -\tilde{\mathbf{G}}^y \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \tag{39}$$

which combined with (26) gives

$$\mathbf{M}_d = \mathbf{M}_n\mathbf{H}\tilde{\mathbf{G}}^y \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \mathbf{W}_d = (\mathbf{E} + \tilde{\mathbf{J}}) \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \mathbf{W}_d$$

Here $\tilde{\mathbf{J}} \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} = 0$ which gives

$$\mathbf{M}_d = \mathbf{E} \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \mathbf{W}_d \tag{40}$$

Note that the disturbance sensitivity is zero ($\mathbf{M}_d = 0$) if and only if $\mathbf{E} = 0$.

Let $\|\mathbf{E}\|_F = \sqrt{\sum_{i,j} e_{ij}^2}$ denote the Frobenius (Euclidean) norm of a matrix, and let $\dagger$ denote the pseudo-inverse of a matrix. Then we have the following theorem:

Theorem 2 Explicit expression for H in extended nullspace method. *Selecting*

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}(\mathbf{W}_{n^y}^{-1}\tilde{\mathbf{G}}^y)^\dagger\mathbf{W}_{n^y}^{-1} \tag{41}$$

minimizes $\|\mathbf{E}\|_F$, and in addition minimizes the noise sensitivity $\|\mathbf{M}_{n^y}\|_F$ among all solutions that minimize $\|\mathbf{E}\|_F$.

Proof: Rewrite the definition (38) for $\mathbf{E}$ as

$$\mathbf{E} = \underbrace{\mathbf{M}_n\mathbf{H}\mathbf{W}_{n^y}}_{-\mathbf{M}_{n^y}}\mathbf{W}_{n^y}^{-1}\tilde{\mathbf{G}}^y - \tilde{\mathbf{J}} \tag{42}$$

From the theory of linear algebra [18], the solution for $-\mathbf{M}_{n^y}$ that minimizes $\|\mathbf{E}\|_F$, and in addition minimizes $\|\mathbf{M}_{n^y}\|_F$ among all solutions that minimize $\|\mathbf{E}\|_F$, is given by $-\mathbf{M}_{n^y} = \tilde{\mathbf{J}}(\mathbf{W}_{n^y}^{-1}\tilde{\mathbf{G}}^y)^\dagger$, which gives (41). To see this, note that minimizing $\|\mathbf{E}\|_F$ is equivalent to finding the least square solution $\mathbf{X} = \mathbf{B}\mathbf{A}^\dagger$ to $\mathbf{X}\mathbf{A} = \mathbf{B}$, where $\mathbf{X} = -\mathbf{M}_{n^y}$, $\mathbf{A} = \mathbf{W}_{n^y}^{-1}\tilde{\mathbf{G}}^y$ and $\mathbf{B} = \tilde{\mathbf{J}}$. □

Remark 1 If we have "enough" measurements ($n_y \geq n_u + n_d$) then the choice for $\mathbf{H}$ in eq. (41) gives $\mathbf{E} = 0$ and thus $\mathbf{M}_d = 0$. However, for the case with "too few" measurements the above choice for $\mathbf{H}$ minimizes $\|\mathbf{E}\|_F$, whereas we really want to minimize $\|\mathbf{M}_d\|_F$. Nevertheless, since $\|\mathbf{M}_d\|_F \leq \|\mathbf{E}\|_F \cdot \| \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \mathbf{W}_d\|_F$, we see that minimizing $\|\mathbf{E}\|_F$ will result in a small value of $\|\mathbf{M}_d\|_F$.

Remark 2 The matrix $\mathbf{H}$ is non-unique and the matrix $\mathbf{M}_n$ in (41) can be viewed as a parameter that can be selected freely. For example, one may select $\mathbf{M}_n = \mathbf{I}$, or one may select $\mathbf{M_n}$ to get a decoupled response from $\mathbf{u}$ to $\mathbf{c}$, i.e. $\mathbf{G} = \mathbf{H}\mathbf{G}^y = I$. However, note that $\mathbf{M}_n\mathbf{H}$, which from eq (34) gives the measurement noise sensitivity, will not be affected as it is given by (36) and (41).

Remark 3 It is appropriate at this point to make a comment about the pseudo-inverse $\mathbf{A}^\dagger$ of a matrix. In general, we can write the least-square solution of $\mathbf{X}\mathbf{A} = \mathbf{B}$ as $\mathbf{X} = \mathbf{B}\mathbf{A}^\dagger$ where the following are true:

(1) $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is the left inverse for the case when $\mathbf{A}$ has full column rank (we have extra measurements). In this case, there are an infinite number of solutions and we seek the solution that minimizes $\|\mathbf{X}\|_F$.

(2) $\mathbf{A}^\dagger = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$ is the right inverse for the case when $\mathbf{A}$ has row column rank (we have too few measurements). In this case there is no solution and we seek the solution that minimizes the Frobenius norm of $\mathbf{E} = \mathbf{B} - \mathbf{X}\mathbf{A}$ (regular least squares).

(3) In the general case with extra measurements, but where some are dependent, $\mathbf{A}$ has neither full column or row rank, and the singular value decomposition may be used to compute the pseudo-inverse $\mathbf{A}^\dagger$.

4.3 Special cases of Theorem 2

We have some important special cases of the Theorem 2:

4.3.1 "Just enough" measurements (original nullspace method)

When $n_y = n_u + n_d$, the measurements and disturbances are independent, so $\tilde{\mathbf{G}}^y$ is invertible and (41) becomes

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{-1} \tag{43}$$

as derived earlier in (37). This choice gives $\mathbf{M}_d = 0$ (zero disturbance loss) and from (26) the resulting effect of the measurement noise is

$$\mathbf{M}_{n^y} = \tilde{\mathbf{J}}[\tilde{\mathbf{G}}^y]^{-1}\mathbf{W}_{n^y} \tag{44}$$

Note that we in this case have no degrees of freedom left for affecting the matrix $\mathbf{M}_{n^y}$.

4.3.2 Extra measurements: Select "just enough" subset

If we have extra measurements ($n_y > n_u + n_d$), then one possibility is to select a "just-enough" subset (such that we get $n_y = n_u + n_d$) before forming $\mathbf{c}$ and then obtain $\mathbf{H}$ from (43) to achieve zero disturbance loss ($\mathbf{M}_d = 0$). The degrees of freedom in selecting the measurement subset can then be used to minimize the loss with respect to the measurement noise, that is, to minimize the norm of $\mathbf{M}_{n^y}$ in (44). The worst-case loss caused by measurement noise is

$$\mathbf{L}_{wc} = \max_{\|\mathbf{n}'^y\|_2 \le 1} \mathbf{L} = \tfrac{1}{2}\bar{\sigma}(\mathbf{M}_{n^y})^2 = \tfrac{1}{2}\bar{\sigma}(\tilde{\mathbf{J}}(\tilde{\mathbf{G}^y})^{-1}\mathbf{W}_{n^y})^2 \le \tfrac{1}{2}\left(\bar{\sigma}(\tilde{\mathbf{J}})\underline{\sigma}(\tilde{\mathbf{G}^y})\bar{\sigma}(\mathbf{W}_{n^y})\right)^2 \tag{45}$$

The selection of measurements does not affect the matrix $\tilde{\mathbf{J}}$, since it from (33) depends only on the Hessian matrices $\mathbf{J}_{uu}$ and $\mathbf{J}_{ud}$. However, the selection of measurements affects the matrix $\tilde{\mathbf{G}}^y$. Thus, in order to minimize the effect of the implementation error, we propose the following two rules:

(1) **Optimal**: In order to minimize the worst-case loss, select measurements such that $\bar{\sigma}(\mathbf{M}_{n^y}) = \bar{\sigma}\left(\tilde{\mathbf{J}}[\tilde{\mathbf{G}}^y]^{-1}\mathbf{W}_{n^y}\right)$ is minimized.

(2) **Sub-optimal**: Assume that the measurements have been scaled with respect the measurement error such that $\mathbf{W}_{n^y} = \mathbf{I}$. From the inequality in eq. (45), it then follows that the effect of the measurement error $\mathbf{n}^y$ will be small when $\underline{\sigma}(\tilde{\mathbf{G}}^y)$ (the minimum singular value of $\tilde{\mathbf{G}}^y$) is large. Thus, it is reasonable to select measurements $\mathbf{y}$ such that $\underline{\sigma}(\tilde{\mathbf{G}}^y)$ is maximized.

Since the optimal rule needs information on the Hessian matrix of the cost function J, the sub-optimal selection rule of maximizing $\underline{\sigma}(\tilde{\mathbf{G}}^y)$ is simpler in practice. This sub-optimal rule was used successfully in [1] to select measurements from 60 candidates for a Petlyuk distillation case study.

4.3.3 Extra measurements: Use all

For the case with extra measurements ($n_y > n_u + n_d$), we may alternatively use all the measurements when forming $\mathbf{c}$ and obtain $\mathbf{H}$ from (41) in Theorem 2. This gives the solution that minimizes the implementation (measurement error) loss subject to having zero disturbance loss ($\mathbf{M}_d = 0$). More precisely, when $n_y > n_u + n_d$ and the measurements and disturbances are independent, the choice for $\mathbf{H}$ in (41), where $\dagger$ denotes the left inverse, minimizes $\|\mathbf{M}_{n^y}\|_F$ (Euclidean norm) among all solutions with $\mathbf{M}_d = 0$. Note that we need to include the noise weight before taking the pseudo inverse in (41).

4.3.4 "Too few" measurements

If there are many disturbances, then we may have too few measurements to get $\mathbf{M}_d = 0$. For the case when both the measurements and disturbances are independent, we have "too few" measurements when $n_y < n_u + n_d$. In this case, the optimal $\mathbf{H}$ given in (41) in Theorem 2 is not affected by the noise weight, and (41) becomes

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^\dagger \tag{46}$$

where $\dagger$ denotes the right inverse and $\mathbf{M}_n$ is, as before, free to choose. However, this explicit expression for $\mathbf{H}$ minimizes $\|\mathbf{E}\|_F$, whereas, as noted in Remark 1, we really want to minimize $\|\mathbf{M}_d\|_F$. Minimizing $\|\mathbf{M}_d\|_F$ is equivalent to solving the following optimization problem

$$\mathbf{H} = \arg\min_{\mathbf{H}} \|\mathbf{H}\mathbf{F}\mathbf{W}_d\|_2 \;\text{ subject to }\; \mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2} \tag{47}$$

For this case, with few measurements and no consideration of measurement error, we have not been able to derive an explicit expression for $\mathbf{H}$, similar to (31) in Theorem 1. However, for practical applications, eq. (46) is most likely acceptable, at least provided we scale the system (i.e. $\mathbf{G}_d^y$) such that $\mathbf{W}_d = I$. There may also be cases where we do have enough measurements, but we nevertheless want to use "too few" measurements to simplify implementation. In this case, we have that $\mathbf{M}_{n^y} = \tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^\dagger \mathbf{W}_{n^y}$ and to minimize $\bar{\sigma}(\mathbf{M}_{n^y})$ (the effect of measurement error), we may first select the set of measurements that maximizes $\underline{\sigma}(\mathbf{M}_{n^y})$, and then select $\mathbf{H}$ according to (46). Also, note that if we have sufficiently few measurements, i.e. $n_y < n_d$, then (31) applies with $\mathbf{W}_{n^y} = 0$ (see comment following Theorem 1).

5 Example

As a simple example, consider a scalar problem with $\mathbf{n}_u = 1$ and $\mathbf{n}_d = 1$ [3]. The cost function to be minimized is

$$J = (u - d)^2 \tag{48}$$

where the nominal disturbance is $d^* = 0$. Assume that the following four measurements are available:

$$y_1 = 0.1(u - d) \qquad y_2 = 20u \qquad y_3 = 10u - 5d \qquad y_4 = u$$

We assume that the system is scaled such that $|d| \leq 1$ and $|n_i| \leq 1$, i.e.,

$$\mathbf{W}_d = 1, \quad \mathbf{W}_{n^y} = I \tag{49}$$

and we want to find the optimal measurements or combinations to control at constant setpoints.

Solution. From (48), it is clear that $J_{opt}(d) = 0 \; \forall \; d$ and the optimal input is $u^{opt}(d) = d$. We find $\mathbf{J}_{uu} = 2$ and $\mathbf{J}_{ud} = -2$ and

$$\mathbf{G}^{yT} = \begin{bmatrix} 0.1 & 20 & 10 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{G}_d^{yT} = \begin{bmatrix} -0.1 & 0 & -5 & 0 \end{bmatrix} \tag{50}$$

The optimal sensitivity matrix $\mathbf{F}$ is obtained from (15) or (39). This gives $\mathbf{F} = [0 \; 20 \; 5 \; 1]^T$.

5.1 Single measurement candidates

Let us first consider the use of individual measurements as controlled variables ($c = y_i$, $i = 1, 2, 3, 4$). The losses $L_{wc} = \frac{1}{2}\bar{\sigma}(\mathbf{M})^2$ are

$$L_{wc}^1 = 100 \qquad L_{wc}^2 = 1.0025 \qquad L_{wc}^3 = 0.26 \qquad L_{wc}^4 = 2 \tag{51}$$

Measurement y_1 has $\Delta y_1^{opt} = 0$, so it happens to have zero disturbance loss ($M_d = 0$). However, this measurement is sensitive to noise (as can be seen from the small gain in $\mathbf{G}^y$) and we see that this choice actually has the largest loss $L_{wc}^1 = 100$. y_3 is the best single measurement candidate. This illustrates the importance of taking into account the implementation error (measurement noise).

Consider combining two measurements, $c = \mathbf{H}y = h_1 y_i + h_2 y_j$. Let us first consider combinations that give zero disturbance loss $\mathbf{M}_d = 0$, which is possible since $n_u + n_d = n_y = 2$. The "null space" combination ($\mathbf{H} = (h_1 \ h_2)$) is most easily obtained using (33). For example, for measurements $(2, 3)$, $\mathbf{F} = [20 \ 5]^T$ and

$$\mathbf{H} = [h_1 \ h_2] = \mathcal{N}([20 \ 5]) = [-0.2425 \ 0.9701] \tag{52}$$

The controlled variable is then $c = -0.2425 y_2 + 0.9701 y_3$. The same result is obtained from (41).

The results with the nullspace method for all six possible combinations are given in Table 2. The table gives the worst-case loss L_{wc} caused by the measurement error. We have $L_{wc} = \frac{1}{2}\bar{\sigma}(\mathbf{M})^2$, where, since $M_d = 0$, $\mathbf{M} = \mathbf{M}_{n^y} = \tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{-1}\mathbf{W}_{n^y}$. To compare, we also show in Table 2 $\underline{\sigma}(\tilde{\mathbf{G}}^y)$, which according to the "sub-optimal rule for selecting measurements" should be maximized in order to minimize the implementation error. We note that for this example that maximizing $\underline{\sigma}(\tilde{\mathbf{G}}^y)$ gives the same (correct) ranking as minimizing L_{wc}.

From Table 2, we see that combinations involving measurement y_1 are all sensitive to noise. Combination $(i, j) = (2, 3)$ is the best, followed by $(3, 4)$, while $(1, 2)$, $(1, 4)$ and $(1, 3)$ have the same noise sensitivity when they are combined using the nullspace method. The reason is that $\mathcal{N}(F^T) = [1 \quad 0]$, so that only measurement y_1 is used. Combination $(2, 4)$ yields infinite noise sensitivity to noise with the nullspace method, since $\tilde{\mathbf{G}}^y$ is singular.

Table 2
Combinations of two measurements, $c = h_1 y_i + h_2 y_j$, with zero disturbance loss ($\mathbf{M}_d = 0$) and resulting loss L_{wc} caused by measurement error.

		H from (41)			
y_i	y_j	h_1	h_2	L_{wc}	$\underline{\sigma}(\tilde{\mathbf{G}}^y)$
2	3	−0.2425	0.9701	0.0425	4.449
3	4	−0.1961	0.9806	1.04	0.446
1	2	−1	0	100	0.1
1	4	−1	0	100	0.0995
1	3	−1	0	100	0.0447
2	4	−0.0499	0.988	∞	0

Next, consider the optimal combination of two measurements for disturbances and measurements error ("exact local method"). The results are summarized in Table 3. Again, we find that the best combination is $(2, 3)$ with a loss

Table 3
Combinations of two measurements, $c = h_1 y_i + h_2 y_j$, with minimum loss L_{wc} for combined disturbances and measurement error.

		H from (31)		
y_i	y_j	h_1	h_2	L_{wc}
2	3	−0.2312	0.9729	0.0406
3	4	−0.8296	0.5584	0.198
1	3	0.2293	0.9733	0.2351
1	2	0.8753	0.4836	0.8969
2	4	0.0499	0.9988	0.9050
1	4	0.1869	0.9824	1.8670

$L_{wc}^{23} = 0.0406$. This gives $\mathbf{M}_d = -0.0635$ so, as expected, the disturbance loss is non-zero. For this combination, the result is very similar to the extended nullspace method which gave $L_{wc}^{23} = 0.0425$ and $\mathbf{M}_d = 0$. However, for the other five two-measurement combinations, the differences are much larger as the use of (31) gives a significantly lower sensitivity to measurements error, see Table 3. However, note that L_{wc} for those five cases is only slightly better than using a single measurement, see (51).

5.3 Measurement combinations: Use all four measurements

Consider again first the case when we want zero disturbance loss ($\mathbf{M}_d = 0$), and eq. (41) in the extended nullspace method gives (after normalizing the 2-norm of **H** to 1):

$$\mathbf{H}^{all} = \begin{bmatrix} 0.0206 & -0.2419 & 0.9700 & -0.0121 \end{bmatrix} \tag{53}$$

which gives $\mathbf{G} = 4.852$ and $\mathbf{M}_n = 0.2915$. The loss contribution from the disturbance and the noise are $\mathbf{M}_d = 0$ and $\mathbf{M}_{n^y} = [-0.0060\ 0.0705\ -0.2827\ 0.0035]$, respectively. The corresponding loss is $L_{wc} = \bar{\sigma}^2[\mathbf{M}_d\ \mathbf{M}_{n^y}]/2 = 0.04248$.

To compare, the optimal combination ("exact local method") with respect to *combined* disturbances and measurement noise, obtained from (31) is (after normalizing the 2-norm of **H** to 1) [3]

$$\mathbf{H}^{opt} = \begin{bmatrix} 0.0208 & -0.2317 & 0.9725 & -0.0116 \end{bmatrix} \tag{54}$$

which gives $\mathbf{G} = 5.082$ and $\mathbf{M}_n = 0.2783$. The loss contribution from the disturbance and the noise are $\mathbf{M}_d = -0.0606$ and $\mathbf{M}_{n^y} = [-0.0057\ 0.0645\ -$

0.2706 0.0032], respectively. The resulting loss is $L_{wc} = 0.0405$, which is very similar to the extended null-space method, and only marginally improved compared to using only two measurements ($L_{wc}^{23} = 0.0406$). The reduction in loss is small compared to using only two measurements ($L_{wc}^{23} = 0.0406$).

In summary, the simple two-step nullspace method, where one first selects a "just enough" set of measurements by maximizing $\underline{\sigma}(\tilde{\mathbf{G}}^y)$, and then obtains $\mathbf{H}$ from the null-space method, using either eqs. (43) or (33), works well for the example.

6 Example 2: Control of refrigeration cycle

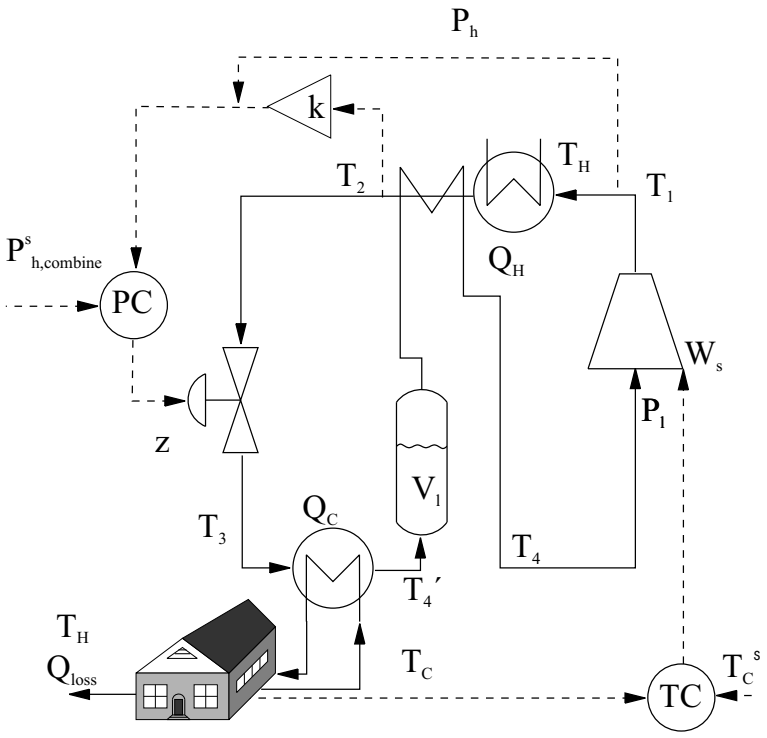

Fig. 3. Proposed control structure for the refrigeration cycle.

For a more physically motivated example, we consider the optimal operation of a CO_2 refrigeration cycle [6], for example, it could be the air condition (AC) unit for a house. The cycle has one unconstrained degree of freedom ($n_u = 1$), which may be viewed as the high pressure ($y_1 = p_h$) in the cycle. Ideally, p_h should be kept at its optimal value by varying the free input (u); which is the choke valve position. However, simply keeping a constant setpoint $p_{h,s}$ is far from optimal because of disturbances. Three disturbances ($n_d = 3$) are considered: The outside temperature T_H, the inside temperature T_C (e.g. because of a setpoint change) and the heat transfer rate UA. As a start, control of single variables was considered, because this is the simplest and is

the preferred choice if such a variable can be found. The best single controlled variable (with a large scaled gain and a small loss) was found to be the mass holdup ($c = M_c$) in the condenser [6], but it is very difficult to measure in practice. Therefore, a combination of measurements needs to be controlled. Theoretically, we need to combine at least 4 measurements ($n_u + n_d = 4$) to get zero loss, independent of the disturbances (i.e., to get $\mathbf{M}_d = 0$). However, to simplify the implementation we prefer to use fewer measurements. Two measurements which are easy to measure and have a reasonably large scaled gain [6], are the high pressure ($y_1 = p_h$) and the temperature before the choke valve ($y_2 = T_h$). It is not possible to get $\mathbf{M}_d = 0$ with only two measurement, so instead $\mathbf{H} = \begin{bmatrix} h_1 & h_2 \end{bmatrix}$ was obtained numerically by minimizing $\|\mathbf{M}_d\|_F$; see eq. (46) (actually, matrix $\mathbf{H}$ could have been obtained from (31) with $\mathbf{W}_{n^y} = 0$, since we have $n_y = 2 < n_d = 3$; see comment following Theorem 1). This gives a controlled variable $c = h_1 y_1 + h_2 y_2 = h_1 p_h + h_2 T_h$. To get a more physical variable, we select $h_1 = 1$, which gives a controlled variable in the units of pressure. We find [6]

$$c = p_{h,combined} = p_h + k(T_h - 25.5^o\mathrm{C})$$

where $k = h_2/h_1 = -8.53$ bar $/^o$C and the (constant) setpoint for $c = p_{h,combined}$ is 97.6 bar, which is the nominally optimal value for the high pressure. Controlling c may be viewed as controlling the high pressure, but with a temperature-corrected setpoint. A more detailed analysis using a nonlinear model shows that this combination gives very small losses for all disturbances and measurement errors [6].

Other case studies. The results of this paper, and in particular (31) in Theorem 1 have also been applied successfully to a distillation case study where the issue is to select temperature combinations [4].

7 Discussion

7.1 Local method

The above derivations are local, since we assume a linear process and a second-order objective function in the inputs and the disturbances. Thus, the proposed controlled variables are only globally optimal for the case with a linear model and a quadratic objective. In general, we should always, for a final validation, check the losses for the proposed structures using a non-linear model of the process.

The following reformulations of the results in this paper may be useful for extending them to other applications and comparing them with other results.

First, we give a reformulation of the original nullspace method [1].

Theorem 3 Linear invariants for quadratic optimization problem. *Consider an unconstrained quadratic optimization problem in the variables* $\mathbf{u}$ *(input vector of length* n_u*) and* $\mathbf{d}$ *(disturbance vector of length* n_d*)*

$$\min_{\mathbf{u}} \mathbf{J}(\mathbf{u},\mathbf{d}) = \begin{bmatrix}\mathbf{u} & \mathbf{d}\end{bmatrix} \begin{bmatrix}\mathbf{J}_{uu} & \mathbf{J}_{ud} \\ \mathbf{J}_{ud}^T & \mathbf{J}_{dd}\end{bmatrix} \begin{bmatrix}\mathbf{u} \\ \mathbf{d}\end{bmatrix} \tag{55}$$

In addition, there are "measurement" variables $\mathbf{y} = \mathbf{G}^y\mathbf{u} + \mathbf{G}_d^y\mathbf{d}$.

If there exists $n_y \geq n_u + n_d$ *independent measurements (where "independent" means that the matrix* $\tilde{\mathbf{G}}^y = \begin{bmatrix}\mathbf{G}^y & \mathbf{G}_d^y\end{bmatrix}$ *has full rank), then the optimal solution to (55) has the property that there exists* $n_c = n_u$ *linear variable combinations (constraints)* $\mathbf{c} = \mathbf{H}\mathbf{y}$ *that are invariant to the disturbances d. Here,* $\mathbf{H}$ *may be obtained from the nullspace method using (33) (where the optimal sensitivity* $\mathbf{F}$ *may be obtained from (15)) or from the explicit expression (37).*

Next, the result on the worst-case loss [3] and the explicit expression for the "exact local method" in Theorem 1 can be reformulated as follows.

Theorem 4 Loss by introducing linear constraint for noisy quadratic optimization problem. *Consider the unconstrained quadratic optimization problem in Theorem 3:*

$$\min_{\mathbf{u}} \mathbf{J}(\mathbf{u},\mathbf{d}) = \begin{bmatrix}\mathbf{u} & \mathbf{d}\end{bmatrix} \begin{bmatrix}\mathbf{J}_{uu} & \mathbf{J}_{ud} \\ \mathbf{J}_{ud}^T & \mathbf{J}_{dd}\end{bmatrix} \begin{bmatrix}\mathbf{u} \\ \mathbf{d}\end{bmatrix}$$

and a set of noisy measurements $\mathbf{y}_m = \mathbf{y} + \mathbf{n}^y$. *Assume that* $n_c = n_u$ *constraints* $\mathbf{c} = \mathbf{H}\mathbf{y}_m = \mathbf{c}_s$ *are added to the problem, which will result in a non-optimal solution with loss* $L = J(u,d) - J_{\text{opt}}(d)$. *Consider disturbances* $\mathbf{d}$ *and noise* $\mathbf{n}^y$ *with magnitudes*

$$\mathbf{d} = \mathbf{W}_d\mathbf{d}'; \quad \mathbf{n}^y = \mathbf{H}\mathbf{W}_{n^y}^y = \mathbf{W}_{n^y}\mathbf{n}^y; \quad \| \begin{bmatrix}\mathbf{d}' \\ \mathbf{n}^{y\prime}\end{bmatrix} \|_2 \leq 1$$

Then for a given $\mathbf{H}$, *the worst-case loss is* $L_{wc} = \bar{\sigma}(\mathbf{M})^2/2$, *where* $\mathbf{M}$ *is given*

in (20)-(22), and the optimal $\mathbf{H}$ *that minimizes* $\bar{\sigma}(\mathbf{M})$ *is given by (31) in Theorem 1. This optimal* $\mathbf{H}$ *lso minimizes* $\|\mathbf{M}\|_F$.

Note from Theorem 3 that if there are a sufficient number ($n_y \geq n_u + n_d$) of noisy-free measurements, then we can obtain zero loss in Theorem 4.

7.3 Relationship to indirect control

Indirect control is when we want to find a set of controlled variables $\mathbf{c} = \mathbf{H}\mathbf{y}$ such that the primary variables $\mathbf{y}_1$ are indirectly kept at constant setpoints. The case of indirect control is discussed in more detail by Hori et al. [5] and the results are a special case of the results for the nullspace method presented in this paper if we select

$$J = \frac{1}{2}\|\mathbf{y}_1 - \mathbf{y}_1^s\|_2 = \frac{1}{2}[\mathbf{y}_1 - \mathbf{y}_1^s]^T[\mathbf{y}_1 - \mathbf{y}_1^s] \tag{56}$$

To show this, write

$$\Delta\mathbf{y}_1 = \mathbf{G}_1\Delta\mathbf{u} + \mathbf{G}_{d1}\Delta\mathbf{d} = \tilde{\mathbf{G}}_1 \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} \tag{57}$$

and assume $\mathbf{n}_{y_1} = \mathbf{n}_u$, so $\mathbf{G}_1$ is a square matrix. We find that

$$\mathbf{J}_{uu} = \mathbf{G}_1^T\mathbf{G}_1 \tag{58}$$

$$\mathbf{J}_{ud} = \mathbf{G}_1^T\mathbf{G}_{d1} \tag{59}$$

Consider the case with $\mathbf{n}_y = \mathbf{n}_u + \mathbf{n}_d$, where we can achieve perfect indirect control with respect to disturbances. Substituting (58) and (59) into the explicit expression (37) for the nullspace method gives

$$\mathbf{H} = \mathbf{P}_{c,0}^{-1}\tilde{\mathbf{G}}_1(\tilde{\mathbf{G}}^y)^{-1} \tag{60}$$

where we have introduced the new "free" parameter $\mathbf{P}_{c,0} = \mathbf{G}_1\mathbf{J}_{uu}^{-1/2}\mathbf{M}_n = \mathbf{G}_1\mathbf{G}^{-1}$. This is identical to the results of Hori et al. [5].

8 Conclusion

Explicit expressions have been derived for the optimal linear measurement combination $\mathbf{c} = \mathbf{H}\mathbf{y}$.

The null space method [1] for selecting linear measurement combinations $\mathbf{c} = \mathbf{H}\mathbf{y}$ has been extended to the general case with extra measurements, $n_y > n_u + n_d$, see eq. (41) in Theorem 2. The idea of the extended nullspace method is to first focus on minimizing the steady-state loss caused by disturbances, and then, if there are remaining degrees of freedom, minimize the effect of measurement errors.

Alternatively, one may minimize the effect of combined disturbances and measurements errors, which is the "exact local method" of Halvorsen et al. [3]. In this paper, we have derived an explicit solution for $\mathbf{H}$ for this problem, see eq. (31) in Theorem 1. This expression applies to any number of measurements, including $n_y < n_u + n_d$.

To simplify, one often uses only a subset of the available measurements when obtaining the combination $\mathbf{c} = \mathbf{H}\mathbf{y}$. A simple rule, which can aims at minimizing the effect of measurement errors, is to select measurements to maximize $\underline{\sigma}(\tilde{\mathbf{G}}^y)$ or even better, to minimize $\bar{\sigma}(\tilde{\mathbf{J}}(\tilde{G}^y)^{-1}\mathbf{W}_{n^y})$. Here, $\tilde{G}^y$ is the steady-state gain matrix from the inputs and disturbances to the selected measurements.

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] D. Bonvin, G. Francois, and B. Srinivasan. Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty. *J. Proc. Control*, 15(6):701–712, 2005.

[3] 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.

[4] E. Hori and S. Skogestad. Selection of controlled variables: Maximum gain rule and combination of measurements. 2008. Accepted for publication in Ind. Eng. Chem. Res.

[5] E. Hori, S. Skogestad, and V. Alstad. Perfect steady state indirect control. *Ind. Eng. Chem. Res.*, 44(4):291–309, 2005.

[6] J. B. Jensen and S. Skogestad. Optimal operation of simple refrigeration cycles. Part II: Selection of controlled variables. *Comp. Chem. Eng.*, 31: 1590–1601, 2007.

[7] J.V. Kadam, W. Marquardt, M. Schlegel, T. Backx, O.H. Bosgra, P.-J. Brouwer, G. Dünnebier, D.V. Hessem, and A. Tiagounovand S.D. Wolf. Towards integrated dynamic real-time optimization and control of industrial processes. *FOCAPO 2003. 4th Int. Conf. of Computer-Aided Process*

Operations, Proceedings of the Conference held at Coral Springs, Florida, January 12-15, 2003, pages 593–596, 2003.

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

[9] V. Kariwala, Y. Cao, and S. Janardhanan. Local self-optimizing control with average loss minimization. 2007. Accepted for publication in Ind. Eng. Chem. Res.

[10] T.E. Marlin and A.N. Hrymak. Real-time optimization of continous processes. *American Institute of Chemical Engineering Symposium Series - Fifth International Conference on Chemical Process Control*, 93:85–112, 1997.

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

[12] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, 1999.

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

[14] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control - 2nd Edition*. John Wiley & Sons, 2005.

[15] B. Srinivasan, CJ. Primus, D. Bonvin, and N. L. Ricker. Run-to-run optimization via control of generalized constraints. *Control Eng. Pract.*, 9(8):911–919, 2001.

[16] B. Srinivasan, D. Bonvin, and E. Visser. Dynamic optimization of batch processes - I. characterization of the nominal solution. *Computers and chemical engineering*, 27(1):1–26, 2003.

[17] B. Srinivasan, D. Bonvin, and E. Visser. Dynamic optimization of batch processes - III. role of measurements in handling uncertainty. *Computers and chemical engineering*, 27(1):27–44, 2003.

[18] G. Strang. *Linear algebra and its applications*. Harcourt Brace & Company, 3 edition, 1988.

[19] Y. Zhang, D. Nader, and J.F. Forbes. Results analysis for trust constrained real-lime optimization. *J. Proc. Control*, 11:329–341, 2001.

[20] Y. Zhang, D. Monder, and J.F. Forbes. Real-time optimization under parametric uncertainty: A probability constrained approach. *J. Proc. Control*, 12(3):373–389, 2002.

A Analytical solution for the exact local method

A.1 Scalar case

The minimization problem for the scalar case in (29) can be rewritten as:

$$\min_{\mathbf{x}} \|\tilde{\mathbf{F}}^T\mathbf{x}\|_2 = \min_{\mathbf{x}} \mathbf{x}^T\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\mathbf{x} \text{ subject to } \mathbf{G}^{yT}\mathbf{x} = \mathbf{J}_{uu}^{1/2} \tag{A.1}$$

where we have introduced $\mathbf{x} \triangleq \mathbf{H}^T$, which is a column-vector in the scalar case.

The solution of this problem must satisfy the following KKT-conditions (e.g. [12, p. 444]):

$$\begin{bmatrix} \tilde{\mathbf{F}}\tilde{\mathbf{F}}^T & -\mathbf{G}^y \\ \mathbf{G}^{yT} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{J}_{uu}^{1/2} \end{bmatrix} \tag{A.2}$$

To find the optimal $\mathbf{x}$, we must invert the KKT-matrix and from the Schur complement of the inverse of a partitioned matrix (e.g. [14, p. 516]), we obtain that the optimum $\mathbf{x}$ is

$$\mathbf{x} = \mathbf{H}^T = (\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\mathbf{G}^y(\mathbf{G}^{yT}(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T)^{-1}\mathbf{G}^y)^{-1}\mathbf{J}_{uu}^{1/2} \tag{A.3}$$

Comment: In the scalar case, $\mathbf{G}^y$ is a (column) vector and $\mathbf{J}_{uu}^{1/2}$ is a scalar. However, the expression and proof also applies if $\mathbf{G}^y$ were a matrix and $\mathbf{J}_{uu}^{1/2}$ were a vector. This fact is important for the extension to the multivariable case.

A.2 Extension to multivariable case

To show that the solution for the scalar case also applies to the multivariable case, we first transform the multivariable case into a scalar problem. In this proof, we consider a system with 2 controlled variables ($n_u = 2$), but it can easily be extended to any dimension.

The optimization problem is $\min_{\mathbf{H}^T} \|\mathbf{H}\tilde{\mathbf{F}}\|_F$ subject to $\mathbf{H}\mathbf{G}^y = \mathbf{J}_{uu}^{1/2}$, where we introduce $\mathbf{X} = \mathbf{H}^T$. The matrices $\mathbf{X}$ and $\mathbf{J}_{uu}^{1/2}$ are split into vectors

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{J}_{uu}^{1/2} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \end{bmatrix} \tag{A.4}$$

We further introduce the long vectors

$$\mathbf{x}_n = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{J}_n = \begin{bmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{bmatrix} \tag{A.5}$$

and the large matrices

$$\mathbf{G}_n^{yT} = \begin{bmatrix} \mathbf{G}^{yT} & 0 \\ 0 & \mathbf{G}^{yT} \end{bmatrix}, \quad \tilde{\mathbf{F}}_n = \begin{bmatrix} \tilde{\mathbf{F}} & 0 \\ 0 & \tilde{\mathbf{F}} \end{bmatrix} \tag{A.6}$$

Then, $\mathbf{H}\tilde{\mathbf{F}} = \mathbf{X}^T\tilde{\mathbf{F}} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \end{bmatrix} \tilde{\mathbf{F}} = \begin{bmatrix} \mathbf{x}_1^T\tilde{\mathbf{F}} \\ \mathbf{x}_2^T\tilde{\mathbf{F}} \end{bmatrix}$ and for the 2-norm, the following applies

$$\|\mathbf{H}\tilde{\mathbf{F}}\|_F = \| \begin{bmatrix} \mathbf{x}_1^T\tilde{\mathbf{F}} \\ \mathbf{x}_2^T\tilde{\mathbf{F}} \end{bmatrix} \|_F = \| \begin{bmatrix} \mathbf{x}_1^T\tilde{\mathbf{F}} & \mathbf{x}_2^T\tilde{\mathbf{F}} \end{bmatrix} \|_F = \|\mathbf{x}_n^T\tilde{\mathbf{F}}_n\|_F \tag{A.7}$$

where it is noted that $\|\cdot\|_F = \|\cdot\|_2$ for a vector.

The constraints $\mathbf{G}^{yT}\mathbf{X} = \mathbf{J}_{uu}^{1/2}$ become

$$\begin{bmatrix} \mathbf{G}^{yT}\mathbf{x}_1 & \mathbf{G}^{yT}\mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \end{bmatrix} \tag{A.8}$$

or $\mathbf{G}^{yT}\mathbf{x}_1 = \mathbf{J}_1$ and $\mathbf{G}^{yT}\mathbf{x}_2 = \mathbf{J}_2$, which can be rewritten as

$$\begin{bmatrix} \mathbf{G}^{yT}\mathbf{x}_1 \\ \mathbf{G}^{yT}\mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{bmatrix} \text{ or } \mathbf{G}_n^{yT}\mathbf{x}_n = \mathbf{J}_n \tag{A.9}$$

(A.7) and (A.9) is a vector optimization problem of the form in (A.1) and from (A.3) the solution is

$$\mathbf{x}_n = (\tilde{\mathbf{F}}_n\tilde{\mathbf{F}}_n^{\ T})^{-1}\mathbf{G}_n^y(\mathbf{G}_n^{yT}(\tilde{\mathbf{F}}_n\tilde{\mathbf{F}}_n^{\ T})^{-1}\mathbf{G}_n^y)^{-1}\mathbf{J}_n \tag{A.10}$$

We now need to "unpack" this to find the optimal $\mathbf{H}^T = \mathbf{X}$. Substituting the values and rearranging (A.10)

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \left(\begin{bmatrix} \tilde{\mathbf{F}} & 0 \\ 0 & \tilde{\mathbf{F}} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{F}} & 0 \\ 0 & \tilde{\mathbf{F}} \end{bmatrix}^T \right)^{-1} \begin{bmatrix} \mathbf{G}^y & 0 \\ 0 & \mathbf{G}^y \end{bmatrix} \left(\begin{bmatrix} \mathbf{G}^y & 0 \\ 0 & \mathbf{G}^y \end{bmatrix}^T \left(\begin{bmatrix} \tilde{\mathbf{F}} & 0 \\ 0 & \tilde{\mathbf{F}} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{F}} & 0 \\ 0 & \tilde{\mathbf{F}} \end{bmatrix}^T \right)^{-1} \begin{bmatrix} \mathbf{G}^y & 0 \\ 0 & \mathbf{G}^y \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{bmatrix}$$

and we see that

$$\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix} = \begin{bmatrix}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\left(\mathbf{G}^{yT}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\right)^{-1}\mathbf{J}_1\\ \left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\left(\mathbf{G}^{yT}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\right)^{-1}\mathbf{J}_2\end{bmatrix} \tag{A.11}$$

and finally,

$$\begin{aligned}\mathbf{H}^T = \mathbf{X} = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2\end{bmatrix} = \left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\left(\mathbf{G}^{yT}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\right)^{-1}\begin{bmatrix}\mathbf{J}_1 & \mathbf{J}_2\end{bmatrix} =\\ = \left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\left(\mathbf{G}^{yT}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right)^{-1}\mathbf{G}^y\right)^{-1}\mathbf{J}_{uu}^{1/2}\end{aligned} \tag{A.12}$$

This proves that the solution for the scalar case also applies for the multivariable.

Figure 1:Feedback implementation of optimal operation with separate layers for optimization (RTO) and control.
Figure 2: Combining measurements $\mathbf{y}$ to get controlled variables $\mathbf{c}$ (linear case).
Figure 3: Proposed control structure for the refrigeration cycle.
Table 1: Notation.
Table 2: Combinations of two measurements, $c = h_1 y_i + h_2 y_j$, with zero disturbance loss ($\mathbf{M}_d = 0$) and resulting loss L_{wc} caused by measurement error.
Table 3: Combinations of two measurements, $c = h_1 y_i + h_2 y_j$, with minimum loss L_{wc} for combined disturbances and measurement error.