Elsevier Editorial System(tm) for Journal of Process Control

Manuscript Draft

Manuscript Number: JPROCONT-D-07-00046

OK

Optimal

Title: Ex~~tended nullspace method for selecting~~ng measurement combinations ~~as~~ ~~controlled variab~~les for near-optimal steady-state operation

ok

Article Type: Research Paper

Section/Category:

Keywords: control structure selection, self-optimizing control, measurement selection

Corresponding Author: Prof. Sigurd Skogestad,

Corresponding Author's Institution: Norwegian University of Science and Technology

First Author: Sigurd Skogestad, Professor

Order of Authors: Sigurd Skogestad, Professor; Vidar Alstad, Ph.D.

Manuscript Region of Origin:

of Alstad and Skogestad (1997)

optimal

Abstract: This paper deals with the selection of linear measurement combinations as controlled variables, c = Hy. The objective is to achieve self-optimizing control where fixing the controlled variables c indirectly gives near-optimal steady-state operation with a small loss. The nullspace method focuses on minimizing the loss caused
by disturbances. The original nullspace method deals with the case where we have as many independent measurements y as inputs plus disturbances, and one may obtain zero disturbance loss, ~~at least locally. In~~ and in ~~this paper~~ this paper we provide an explicit expression for the combination matrix H which ~~allows us to extend the~~ extend ~~nullspace method~~ ~~extension~~ the nullspace method where we have to cases with extra measurements, where the extra degrees of freedom are used to minimize the loss caused by measurement errors, and to cases with with too few measurements, ~~where zero loss with respect to disturbances is impossible.~~

In addition, a new explicit expression for H is derived for the case where the objective is to minimize the combined effect of disturbances and measurement errors.

Latex source

A completely different approach to optimizing control is to focus on selecting the right variable c to control, which is the idea of "self-optimizing control" (~~Skogestad, 2000~~ [10]). The objective is to search for combinations of measurements c = Hy, which when controlled, will keep the process at its optimum operating condition despite disturbances and measurement errors. Thus, the implementation is trivial and the maintenance requirements are minimal. The idea in this paper is extend this approach, by providing explicit formulas for selecting the optimal H.

Extended nullspace method for selecting measurement combinations as controlled variables for optimal steady-state operation

and Eduardo S. Hori

Vidar Alstad [1] ~~and~~ Sigurd Skogestad *

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

(B1)

Abstract

This paper deals with the selection of linear measurement combinations as controlled variables, $\mathbf{c} = \mathbf{H}\mathbf{y}$. The objective is to achieve self-optimizing control where fixing the controlled variables $\mathbf{c}$ indirectly gives near-optimal steady-state operation with a small loss. The nullspace method focuses on minimizing the loss caused by disturbances. The original nullspace method deals with the case where we have as many independent measurements $\mathbf{y}$ as inputs plus disturbances, and one may obtain zero disturbance loss, at least locally. In this paper, we provide an explicit expression for the combination matrix $\mathbf{H}$ which allows us to extend the nullspace method to cases with extra measurements, where the extra degrees of freedom are used to minimize the loss caused by measurement errors, and to cases with with too few measurements, where zero loss with respect to disturbances is impossible.

move (A) here from next page

(B)

1 Introduction

n_c

The number of output variables that can be independently controlled is equal to the number of independent inputs (~~manipulated variables~~ n_u). ~~However~~, in most cases the number of available measurements (n_y) is larger than the number of independent inputs (~~n_u), and the issue is then to choose which variables $\mathbf{c}$ to control (such that $n_c = n_u$). This can be viewed as a "squaring~~ down" problem. ~~In the linear case we can write $\mathbf{y} = \mathbf{G}^y\mathbf{u}$ and $\mathbf{c} = \mathbf{H}\mathbf{y}$, see Figure~~ 1, , that is, $n_y > n_u$.

* 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

(C) The issue of selecting H can also be viewed as a "squaring down" problem as illustrated in Figure 1, ~~where~~ $y = G^y u$ ~~cut the mass~~ ~~since the~~

see Figure 1 , as mentioned above,

~~and~~ the issue is then to select the nonsquare matrix **H** such that the map (transfer function) $\mathbf{G} = \mathbf{H}\mathbf{G}^y$ from **u** to **c** is square. However, selecting **H** such that **G** is square is not the only issue. More importantly, control of **c** should (directly or indirectly) result in "acceptable operation" of the system.

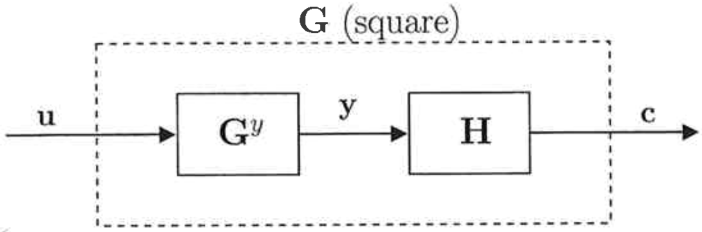

Fig. ~~1~~ 2. Combining measurements **y** to get controlled variables **c** (linear case)

To quantify "acceptable operation" we introduce a scalar cost function $J(\mathbf{u})$ which should be minimized for optimal operation, and "acceptable operation" then means that the loss is acceptable, that is, the actual cost is sufficiently close to the optimal. 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.

Optimizing control is an old research topic, and

A

One method for ensuring optimal operation in chemical processes is real-time optimization (RTO)[6]. Using RTO, the optimal values (setpoints) for the controlled variables **c** are recomputed online based on online measurements and a model of the process, see Figure 2. In most RTO applications a steady-state model is used for the parameter/disturbance estimation and the optimization steps [15, 16], however dynamic versions of the RTO-framework have also been reported in literature [5]. However, the cost of installing and maintaining such systems can be large. In addition, the system can be sensitive to uncertainty.

B1

The need for a RTO layer to compute new optimal setpoints $\mathbf{c}_s$ can then be reduced, or in ~~some~~ many cases even eliminated, ~~by selecting the right controlled variables c. This is the idea of~~ self-optimizing control [10] ~~which~~ is when a constant setpoint policy (is constant) yields acceptable ~~operation~~ loss in spite of the presence of uncertainty, which is here assumed to be represented by (1) *external disturbances* **d** and (2) *implementation errors* $\mathbf{n} \triangleq \mathbf{c}_s - \mathbf{c}$, see Figure 2.

The implementation error **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 2, 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 **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

A completely different approach to optimizing control is to search for combinations (from $\min_H \bar{\sigma}(M(H))$)

They proposed to obtain H numerically, but did not ~~give any information~~ ~~say how~~ say anything about the properties of this optimization problem. Kariwala (2007) proposed ~~a computationally that H could be~~ an iterative method involving a singular value ~~decomposition~~ and eigenvalue decomposition.

process control literature, but the first quantitative treatment was that of Morari et al. [7]. Skogestad [10] 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, **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 that may be used to obtain the optimal measurement combination **H**. ~~However, this method is also computationally unattractive and in addition somewhat difficult to use~~. Hori et al. [4] considered indirect control, which can be formulated as a subproblem of the null space method presented in this paper. Additional related work is presented in [13, 12, 11] on measurement-based optimization to enforce the necessary condition of optimality under uncertainty, ~~The ideas are illustrated~~ with application to ~~on~~ 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.

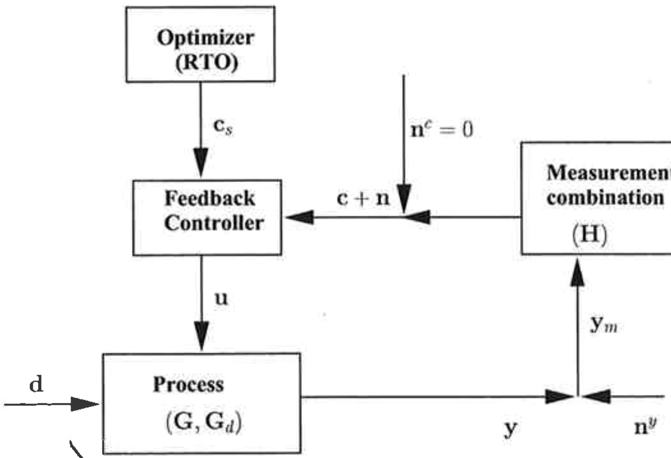

Fig. 2. Feedback implementation of optimal operation.

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) at least equals the number of (independent) inputs (n_u) plus disturbances (n_d), i.e., $n_y \geq n_u + n_d$. It is then optimal to select **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 **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, $\mathbf{n}^y$. The original nullspace method considered the case ~~when $n_y = n_u + n_d$. In this paper, we include the implementation error and extend the null space method to the following cases:~~

~~(1) Extra measurements ($n_y > n_u + n_d$):~~

and for any number of measurements

and the case when $n_y = n_u + n_d$

with no implementation (measurement) error. In this work we provide explicit formulas for H ~~for the~~ when measurement error is included, ~~both for~~ for the following distinct cases:

(1) Optimal H for combined disturbances and implementation errors (Section 3). (2) Optimal H for disturbances, ~~...~~ but min, and mini. considering implementation errors only if there are extra ~~additional~~ measurements (Extended nullspace method, Section 4).

(a) Use all available measurements
(b) Selecton of minimum number $(n_u + n_d)$ of measurements
(2) Too few measurements ($n_y < n_u + n_d$).

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 $\mathbf{y}$
$\mathbf{n}$	-	implementation error associated with $\mathbf{c}$; $\mathbf{n} = \mathbf{n}^c + \mathbf{H}\mathbf{n}^y = \mathbf{H}\mathbf{n}^y$)

n^c - control error associated with c (this paper: $n^c = 0$)

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 2) 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 $J^{opt}(\mathbf{d})$ [9]:

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

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 constaint 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 2). 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.

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

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

[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 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 assume

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

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

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

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

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

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

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

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

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:

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

where

$$\mathbf{M}_d = -\mathbf{J}_{uu}^{1/2}(\mathbf{HG}^y)^{-1}\mathbf{HFW}_d \tag{20}$$

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

Introducing

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

gives $\mathbf{z} = \mathbf{M}\begin{bmatrix} \mathbf{d}' \\ \mathbf{n}^{y\prime} \end{bmatrix}$. 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

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

where the last equality follows from the definition of the singular value $\bar{\sigma}$. Thus, to minimize the worst-case loss we need to minimize $\bar{\sigma}(\mathbf{M})$ with respect to $\mathbf{H}$. This is identical to the "exact local method" in Halvorsen et al. [3], except that $\mathbf{M}_d$ in (20) is expressed in terms of the easily available optimal sensitivity matrix $\mathbf{F}$.

Explicit formula for optimal H.

3 Obtaining ~~the optimal H numerically~~

for the "exact local method"

From (23), the optimal measurement combination is obtained by solving the problem

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

in fact ~~actually~~ solve (28)

~~This problem is fairly~~ easy to ~~solve numerically, 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 as the effect of $\mathbf{n}$ on the loss variables $\mathbf{z}$. We get

$$\mathbf{M}_d = -\mathbf{M}_n\mathbf{H}\mathbf{F}\mathbf{W}_d, \quad \mathbf{M}_{n^y} = -\mathbf{M}_n\mathbf{H}\mathbf{W}_{n^y} \tag{26}$$

or

$$\mathbf{M} = [\mathbf{M_d}\ \mathbf{M_{n^y}}] = -\mathbf{M}_n\mathbf{H}[\mathbf{F}\mathbf{W}_d\ \ \mathbf{W}_{n^y}] \tag{27}$$

Next, we use the fact that the solution is not unique, so that if $\mathbf{H}$ is an optimal solution to the problem (24), 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 1 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.

the matrix $\tilde{F}$ which we want to minimize and in the constraints

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 (27) 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 $\mathbf{M}_n = \mathbf{I}$. The optimization problem (24) can then be stated as

introducing $\tilde{F} \triangleq [FW_d\ W_{n^y}]$ (*)

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

explicit

and as we show an ~~explicit~~ solution is even available.

This is fairly easy to solve numerically ~~because~~ of the linearity in H in both ~~the objective function and constraints~~.

(21) (8) Choice of norm in here

and $n_u = n_c = 1$ $\quad \bar{\sigma}(M) = \|M\|_2$ and

Scalar case. For the scalar case ($\mathbf{c}$ is a scalar), $\mathbf{M}$ and $\mathbf{H}$ are vectors and an analytic solution ~~is available~~. ~~This follows since the singular value of a vector~~

to (28) is easily derived

~~It was~~ It may seem that this optimization problem is non-trivial as M depends nonlinearly on H as shown in (21)-(22). Kariwala (2007) ~~suggests~~ provides an iterative solution involving ~~a~~ singular value and eigenvalue decompositions. However, (24) is

(F) in here (see email)

The optimization problem becomes

~~is equal to the 2-norm, and we get a quadratic optimization problem subject~~
to ~~linear equality constraints,~~

invert the resulting KKT-matrix using

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

where H is a vector for the scalar case. Introducing $\mathbf{Q} \triangleq \tilde{\mathbf{F}}\tilde{\mathbf{F}}^T$, the

~~where $\mathbf{Q} \triangleq [\mathbf{F}\mathbf{W}_d \ \ \mathbf{W}_{n^y}][\mathbf{F}\mathbf{W}_d \ \ \mathbf{W}_{n^y}]^T$. The~~ optimal solution is (e.g. [8, p. 444] and the property of equality constrained QPs Schur complement) ~~to invert the KKT-matrix~~

$H^T =$

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

where it is assumed that ~~$\mathbf{G}^y$~~ G has full rank.

Choice of norm. The optimization problems in (24) and (28) involve the singular value of **M**, which represents the worst-case effect of combined disturbances and measurement errors on the loss. A closely related problem is to minimize the 2-norm (Euclidean or Frobenius norm), $\|\mathbf{M}\|_2 = \sqrt{\sum_{i,j} |m_{ij}|^2}$, 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". 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. In ~~the extended nullspace method presented below, the 2-norm is used for mathematical convenience.~~

Actually, as discussed below, it appears that for this particular problem we have a kind of "super-optimality" when the choice of H that minimizes $\|M\|_2$, also minimizes $\bar{\sigma}(M)$ (Kariwala, DYCOPS '07, pp. 259-264, Lemma 7)

E

6

E (more)

4 Extended nullspace method

The solution for H in (22) ~~in (24), (28) and (30)~~ minimize 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. ~~A third justification is that there exists a simple analytic solution, namely the nullspace method.~~

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 **H** such that

$$\mathbf{H}\mathbf{F} = 0 \tag{31}$$

This provides an alternative derivation of the nullspace method of [1]. It is always possible to find a non-trivial solution (i.e. $\mathbf{H} \neq 0$) **H** satisfying $\mathbf{H}\mathbf{F} = 0$ 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 **H** as the nullspace of $\mathbf{F}^T$:[1]

$$\mathbf{H} = \mathcal{N}(\mathbf{F}^T) \tag{32}$$

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 $\mathbf{M}_n^y$. One objective of this paper is to 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 H *for orginal null space method*

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

$$\mathbf{z} = \overbrace{\begin{bmatrix} \mathbf{J}_{uu}^{1/2} & \mathbf{J}_{uu}^{1/2}\mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \end{bmatrix}}^{\tilde{\mathbf{J}}} \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} \qquad (33)$$

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

$$\mathbf{z} = -\mathbf{M}_n\mathbf{H}\mathbf{n}^y = \mathbf{M}_n\mathbf{H}\Delta\mathbf{y} = \mathbf{M}_n\mathbf{H}\tilde{\mathbf{G}}^y \begin{bmatrix} \Delta\mathbf{u} \\ \Delta\mathbf{d} \end{bmatrix} \qquad (34)$$

where $\tilde{\mathbf{G}}^y = \begin{bmatrix} \mathbf{G}^y & \mathbf{G}_d^y \end{bmatrix}$ is the augmented plant. Comparing eqs. (33) and (34) yields

$$\mathbf{M}_n\mathbf{H}\tilde{\mathbf{G}}^y = \tilde{\mathbf{J}} \qquad (35)$$

where $\tilde{J}$ is defined in (33)

~~and 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

$$\mathbf{M}_n\mathbf{H} = \tilde{\mathbf{J}}[\tilde{\mathbf{G}}^y]^{-1} \qquad (36)$$

This expression gives $\mathbf{M}_n\mathbf{H}$ for a case with zero disturbance sensitivity ($\mathbf{M}_d = 0$), and thus gives the same result as (32). 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).

4.2 Extended nullspace method

The explicit solution for $\mathbf{H}$ in (36) forms the basis for the extending the nullspace method to cases where we have extra measurements ($n_y > n_u + n_d$) or too few measurements ($n_y < n_u + n_d$).

we want to derive a
~~To find the~~ relationship between E and M_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. (35):

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

From (15) and (9) the optimal sensitivity can be written

$$\mathbf{F} = -\tilde{\mathbf{G}}^y \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \tag{38}$$

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

$$\mathbf{M}_d = \mathbf{E} \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \mathbf{W}_d \tag{39}$$

least square

To see this note that minimizing $\|E\|_2$ is equivalent to ~~solving the least square problem XA=B~~ finding the ~~best~~ solution X ~~[illegible]~~ to $XA=B$ where $A=-M_n^y$ and $B=\tilde{J}$

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

Let $\|\mathbf{E}\|_2 = \sqrt{\sum_{i,j} e_{ij}^2}$ denote the Euclidean norm of a matrix, and let $\dagger$ denote the pseudo-inverse of a matrix. Then we have the following theorem:

Theorem ~~4.1~~ 2 Explicit expression for H in extended nullspace method. *Selecting*

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}(\mathbf{W}_{n^y}^{-1}\tilde{\mathbf{G}}^y)^\dagger\mathbf{W}_{n^y}^{-1} \tag{40}$$

minimizes $\|\mathbf{E}\|_2$, and in addition minimizes the noise sensitivity $\|\mathbf{M}_{n^y}\|_2$ among all solutions that minimize $\|\mathbf{E}\|_2$.

Proof: Rewrite the definition (37) ~~for E as~~

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

$-M_n^y$

From the theory of linear algebra [14], the solution for ~~[illegible]~~ that minimizes $\|\mathbf{E}\|_2$ and in addition ~~at the same time~~ minimizes $\|\mathbf{M}_{n^y}\|_2$ is ~~then~~ given by $-\mathbf{M}_{n^y} = \tilde{\mathbf{J}}(\mathbf{W}_{n^y}^{-1}\tilde{\mathbf{G}}^y)^\dagger$, which gives (40). □

~~to consider~~

~~among all s~~

To see this more clearly ~~note~~ note that (41) knows ~~[illegible]~~ $XA=B$ with ~~X=~~ [illegible] and $B=\tilde{J}$

among all solutions that minimize $\|E\|_2$.

Remark 1 If we have "enough" measurements ($n_y \geq n_u + n_d$) then the choice for $\mathbf{H}$ in eq. (40) gives $\mathbf{E} = 0$ and thus $\mathbf{M}_d = 0$. However, for the case with "too few" meassurements the above choice for $\mathbf{H}$ minimizes $\|\mathbf{E}\|_2$, whereas it ~~would seem more reasonable~~ we really want to minimize $\|\mathbf{M}_d\|_2$. ~~Unfortunately, we have no simple explicit solution for H in this case~~. Nevertheless, since $\|\mathbf{M}_d\|_2 \leq \|\mathbf{E}\|_2 \cdot \| \begin{bmatrix} \mathbf{J}_{uu}^{-1}\mathbf{J}_{ud} \\ -\mathbf{I} \end{bmatrix} \mathbf{W}_d\|_2$, we see that minimizing $\|\mathbf{E}\|_2$ will result in a small value of $\|\mathbf{M}_d\|_2$.

Remark 2 The matrix $\mathbf{H}$ is non-unique and the matrix $\mathbf{M}_n$ in (40) 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 (33) gives the measurement noise sensitivity, will not be affected as it is given by (35) and (40).

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 solution of $\mathbf{XA} = \mathbf{B}$ as $\mathbf{X} = \mathbf{BA}^\dagger$ where the following points 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 $\|\mathbf{X}\|_2$

(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 Euclidean norm of $\mathbf{E} = \mathbf{B} - \mathbf{XA}$ (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

We have some important special cases of the Theorem 4.1:

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 (40) becomes

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{-1} \tag{42}$$

as derived earlier in (36). This choice gives $\mathbf{M}_d = 0$ (zero disturbance loss) and the resulting effect of the measurement noise is

from (36)

$$\mathbf{M}_{n^y} = \tilde{\mathbf{J}}[\tilde{\mathbf{G}}^y]^{-1}\mathbf{W}_{n^y} \tag{43}$$

Note that we in this case have no degrees of freedom left for affecting the matrix $\mathbf{M}_{n^y}$. ~~the matrix $[\tilde{\mathbf{G}}^y]^{-1}$ has large elements, or equivalently (within a~~

constant factor)

4.3.2 Extra measurements: ~~Use~~ select "just enough" subset

If we have extra measurements ($n_y > n_u + n_d$), then one ~~alternative~~ 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 $\mathbf{H}$ ~~is selected~~ from ~~from~~ (42) ~~such that we have~~ to achieve zero disturbance loss ($\mathbf{M}_d = 0$). The degrees of freedom in selecting the measurement subset ~~should~~ 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 eq. (43). ~~Note that~~ The worst-case loss ~~is~~ caused by measurement noise n^y is

$$\mathbf{L}_{wc} = \max_{\|\mathbf{n}'^y\|_2 \leq 1} \mathbf{L} = \tfrac{1}{2}\bar{\sigma}(\mathbf{M}_{n^y})^2 = \tfrac{1}{2}\bar{\sigma}(\tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{-1})^2 \leq \tfrac{1}{2}\left(\bar{\sigma}(\tilde{\mathbf{J}})\underline{\sigma}(\tilde{\mathbf{G}}^y)\bar{\sigma}(\mathbf{W}_{n^y})\right)^2 \quad (44)$$

why from (33)

The selection of measurements does not affect the matrix $\tilde{\mathbf{J}}$, since it 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. (44), 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 ~~In this case we should~~ obtain $\mathbf{H}$ from (40) in Theorem 4.1. 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 (40), where † denotes the left inverse, minimizes $\|\mathbf{M}_{n^y}\|_2$ (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 (40).

We have that $M_{n^y} = J(\tilde{G}^y)^{+} W_{n^y}$ and to minimize the effect of measurement ~~noise~~ error

optimal H given in (40) in Theorem ? is not affected by the noise weight, and (40) ~~bec~~ becomes

(or even better minimize $\bar{\sigma}(M_{n^y}) = \bar{\sigma}(J(\tilde{G}^y)^{+} W_{n^y})$)

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 ~~noise weight does not affect $\mathbf{H}$ and (40) in Theorem 4.1 becomes~~

$$\mathbf{H} = \mathbf{M}_n^{-1}\tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{\dagger} \tag{45}$$

where † denotes the right inverse amd $\mathbf{M}_n$ as before is free to choose. This explicit expression ~~of~~ for $\mathbf{H}$ minimizes $\|\mathbf{E}\|_2$, whereas, as noted in Remark 1, we really want to minimize $\|\mathbf{M}_d\|_2$. ~~However, we have no explicit expression for $\mathbf{H}$ in this case, so we would need to obtain $\mathbf{H}$ numerically, for example, by solving the following optimization problem:~~ Minimizing $\|M_d\|_2$ is equivalent to solving the following optimization problem

$$\mathbf{H} = \arg\min_{\mathbf{H}} \|\mathbf{HFW}_d\|_2 \text{ subject to } \mathbf{HG}^y = \mathbf{J}_{uu}^{1/2} \tag{46}$$

However, for practical applications eq. (45) is simpler and most likely acceptable, at least provided we have scaled the system (i.e. $\mathbf{G}_d^y$) such that $\mathbf{W}_d = I$. There may also be cases where we have enough measurements, but we nevertheless want to use "too few" measurements to simplify implementation. In this case, we may first select the set of measurements that maximizes $\underline{\sigma}(\tilde{\mathbf{G}}^y)$, and then select $\mathbf{H}$ according to eq. (45).

(where the norm ~~is~~ may be the singular value ($\bar{\sigma}$) ~~or~~ or the Frobenius ($\|\cdot\|_F$)

$\bar{\sigma}$ is also ok?

5 Example

As a simple example, we consider a scalar problem with $n_u = 1$ and $n_d = 1$ [3]

~~This example with $n_u = 1$ and $n_d = 1$ is an extension of the example found in Halvorsen et al. [3]. Assume that we have a SISO system with one disturbance and the following objective function~~ The cost function to be minimized is

$$J = (u - d)^2 \tag{47}$$

where ~~with~~ the nominal disturbance is $d^* = 0$. We have $\mathbf{J}_{uu} = 2$ ~~and~~ and $\mathbf{J}_{ud} = -2$. From eq. (47) it is clear that $J_{opt}(d) = 0\ \forall\ d$ and the optimal input is $u^{opt}(d) = d$. ~~Assume that the following measurements are available:~~ For this problem we have four measurements

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

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

~~or equivalently~~ and

$$\mathbf{G}^{yT} = \begin{bmatrix}0.1 & 20 & 10 & 1\end{bmatrix} \quad \text{and} \quad \mathbf{G}_d^{yT} = \begin{bmatrix}-0.1 & 0 & -5 & 0\end{bmatrix} \tag{48}$$

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

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

?? Seems wrong; H should depend on W_{n^y} for this case!

cut?

For the case with $n_y > n_d$ the solution to this problem is (see Appendix ?; Sridhar's derivation)

$$H^T = P G^y (G^{yT} P G^y)^{-1} J_{uu}^{1/2}$$

with $P = I - F(F^T F)^{-1} F^T$

The optimal sensitivity matrix $\mathbf{F}$ is obtained from (15) or (38). 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 ($c = y_i$). For the four single measurement candidates the losses are [3]

$$L_1 = 100 \qquad L_2 = 1.0025 \qquad L_3 = 0.26 \qquad L_4 = 2$$

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 this choice actually has the largest loss. y_3 is the best single measurement candidate. This illustrates the importance of taking into account the implementation error (measurement noise).

5.2 *Measurement combinations: Use two of the four measurements*

Let us first consider combinations that give zero disturbance loss (M_d) = 0 which is possible since $n_u + n_d = n_y = 2$

~~Since $n_u + n_d = 2$, it is possible to get zero disturbance loss ($\mathbf{M}_d = 0$)~~ by

consider

combining two measurements, $c = \mathbf{H}y = h_1 y_i + h_2 y_j$. The "null space" combination ($\mathbf{H} = (h_1\ h_2)$) is most easily obtained using (32). For example, for measurements $(2, 3)$, $\mathbf{F} = [20\ 5]^T$ and

$$\mathbf{H} = [h_1\ h_2] = \mathcal{N}([20\ 5]) = [-0.2425\ 0.9701] \tag{50}$$

The controlled variable is then $c = -0.2425 y_2 + 0.9701 y_3$. The same result is obtained from (40).

since $M_d = 0$

The results for 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 $\mathbf{M} = \mathbf{M}_{n^y} = \tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{-1}\mathbf{W}_{n^y}$. 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. Combination $(2, 4)$ yields infinite noise sensitivity to noise with the nullspace method, since $\tilde{\mathbf{G}}^y$ is singular.

(1) from next page

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.

y_i	y_j	$\mathbf{H}$ from (40): 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

D

The reason is that

~~Remark 1~~ ~~Note that for the combinations using measurement y_1, we get $\mathcal{N}(F^T) =$~~ $[1 \;\; 0]$, so that only measurement y_1 is used. ~~This is the reason why the loss in Table 2 is the same for all combinations with y_1.~~

Next, consider the optimal combination of two measurements that minimizes the combined effect of disturbances and measurement error. We use measurement y_2 and y_3 which are the most promising from Table 2.

From (22) we get

~~Remark 2~~ ~~Using only measurements y_2 and y_3, the optimal combination that minimizes the loss L_{wc} with measurement noise included, may be obtained using (30). We get~~ [3] $\mathbf{H}^{opt}_{23} = [-0.2323 \; 0.9727]$ with a loss $L^{23,opt}_{wc} = 0.0406$ [3]. This gives $\mathbf{M}_d = -0.0606$ so, as expected, the disturbance loss is non-zero. ~~Nevertheless,~~ In this case the result is very similar to the extended nullspace method which gave $L^{23}_{wc} = 0.0425$ and $\mathbf{M}_d = 0$.

CHECK! with formula

5.3 Measurement combinations: Use all four measurements

Consider again first the case when we want zero disturbance loss ($\mathbf{M}_d = 0$).

~~Using all measurements,~~ ~~eq.~~ (40) ~~in~~ for the extended nullspace method gives (after normalizing the 2-norm of $\mathbf{H}$ to 1):

$$\mathbf{H}^{all} = \begin{bmatrix} 0.0206 & -0.2419 & 0.9700 & -0.0121 \end{bmatrix} \tag{51}$$

which gives $\mathbf{G} = 4.852$ and $\mathbf{M}_n = 0.2915$. The loss contribution from the disturbance and the noise is $\mathbf{M}_d = 0$ and $\mathbf{M}_{n^y} = [-0.0060 \; 0.0705 \; -0.2827 \; 0.0035]$. The corresponding loss is $L^{all}_{wc} = \bar{\sigma}^2[\mathbf{M}_d \;\; \mathbf{M}_{n^y}]/2 = 0.04248$, which is only marginally improved compared to using only two measurements ($L_{wc} = 0.0425$).

To compare, the "optimal" combination with respect to combined disturbance and measurement noise (with minimum loss L_{wc}), is obtained solving ~~(24) or using~~ (30) (22) (after normalizing the 2-norm of $\mathbf{H}$ to 1) is [3]

CHECK!

$$\mathbf{H}^{opt} = \begin{bmatrix} 0.0208 & -0.2317 & 0.9725 & -0.0116 \end{bmatrix} \tag{52}$$

which gives $\mathbf{G} = 5.082$ and $\mathbf{M}_n = 0.2783$. The loss contribution from the disturbance and the noise is $\mathbf{M}_d = -0.0606$ and $\mathbf{M}_{n^y} = [-0.0057\ 0.0645\ -0.2706\ 0.0032]$. The resulting loss is $L_{wc}^{opt} = 0.0405$. ~~However,~~ The reduction in loss ($L_{opt} = 0.0405$) is small compared to using c_{23} from the nullspace method using only two measurements ($L_{wc}^{23} = 0.0425$).

In summary, the single ~~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 eqs. (42) or (32) to make $\mathbf{M}_d = 0$, works well for the example.

Example 2 in here (O from email)

6 Discussion

6.1 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, we cannot guarantee that the proposed controlled variables are globally optimal. ~~However, using the above expressions should give an indication on how sensitive the candidates are to measurement error~~. For a final validation, we should always check the loss for the proposed structures using the non-linear models of the process.

6.2 Eliminating measurements

We have extended the null space method to the case where we want to use all available measurements. In general, using all measurements should be optimal. However, in many cases many of the measurements are closely correlated or have large measurement errors. In such cases the advantages of using additional measurements and the increased complexity of the control structure may not be justified, and we may want to use fewer measurements as discussed in Section 4.3.4, and also illustrated in the example.

6.3 Minimize disturbance loss (6) in here (from email)

6.~~3~~4 Relationship to indirect control

(A) from next page

Indirect control is when we want to find a set of controlled variables $c = Hy$ such that the primary variables y_1 are indirectly kept at ~~control a set of primary variables $\mathbf{y}_1$, at~~ constant setpoints, and these results are ~~This is~~ a special case of the results in this paper if we select

$$J = \frac{1}{2}\|\mathbf{y}_1 - \mathbf{y}_1^s\|_2 = \frac{1}{2}[\mathbf{y}_1 - \mathbf{y}_1^s]^T[\mathbf{y}_1 - \mathbf{y}_1^s] \tag{53}$$

Consider the case with $n_y = n_u + n_d$ ~~out~~ where we can achieve perfect indirect control with respect to disturbances. Substituting (55) and (56) into (36) ~~from~~ ~~the original nullspace method~~ gives

$$H = P_{c0}^{-1} \tilde{G}_1 (\tilde{G}^y)^{-1} \quad (57)$$

where we have introduced the new "free" parameter $P_{c0} = G_1 J_{uu}^{-1/2}$, $M_n = G_1 G^{-1}$. This is identical to the results of Hori et al. [4].

We ~~have that~~ write

$$\Delta \mathbf{y}_1 = \mathbf{G}_1 \Delta \mathbf{u} + \mathbf{G}_{d1} \Delta \mathbf{d} = \tilde{\mathbf{G}}_1 \begin{bmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{d} \end{bmatrix} \tag{54}$$

We assume that $n_{y1} = n_u$ so G_1 is a square matrix. We ~~and~~ find that

$$\mathbf{J}_{uu} = \mathbf{G}_1^T \mathbf{G}_1 \tag{55}$$

$$\mathbf{J}_{ud} = \mathbf{G}_1^T \mathbf{G}_{d1} \tag{56}$$

(A)

The case of indirect control is discussed in more detail by Hori et al. [4]. ~~The variables $\Delta \mathbf{c} = \mathbf{G}\Delta \mathbf{u} + \mathbf{G}_d \Delta \mathbf{d}$ are selected in order to indirectly control $\mathbf{y}_1$, and the results in Hori et al. [4] follow directly from the results in this paper by using~~

$$\cancel{\mathbf{P}_d = \mathbf{M}_d = (\mathbf{G}_{d1} - \mathbf{G}_1 \mathbf{G}^{-1} \mathbf{G}_d)} \tag{57}$$

$$\cancel{\mathbf{P}_c = \mathbf{M}_n = \mathbf{G}_1^{-1} \mathbf{G}} \tag{58}$$

Explicit expressions have been derived for ~~the~~ deriving the optimal measurement combination $c = Hy$. ~~For combined meas dis~~

7 Conclusion

The null space method [1] for selecting linear measurement combinations $\mathbf{c} = \mathbf{Hy}$ has been extended to the general case with extra measurements ~~or too few measurements~~ ($n_y > n_u + n_d$), see eq. (40). 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.~~ For the case with extra measurements, the use of eq. (40) to obtain $\mathbf{H}$ minimizes the loss with respect to implementation/measurement error (minimizes $\|\mathbf{M}_n^y\|_2$) subject to achieving zero disturbance loss ($\mathbf{M}_d = 0$). For the case with too few measurements, the use of eq. (40) minizes $\|\mathbf{E}\|_2$, wheras one really would like to minimize $\|\mathbf{M}_d\|_2$. Altough $\mathbf{E}$ and $\mathbf{M}_d$ are related by eq. (39), one may in this case instead obtain $\mathbf{H}$ numerically using (46).

Refer to Thiago's in conclusion!

An alternative approach for the case with extra measurements, which is usually preferred in practice,, is to use only a subset of the measurements. In the two-step nullspace method, one first obtains a "just-enough" subset with small sensitivity to implementation error by maximizing $\underline{\sigma}((\tilde{\mathbf{G}}^y)$ (or even better minimizing $\overline{\sigma}(\mathbf{M}_n^y) = \overline{\sigma}(\tilde{\mathbf{J}}(\tilde{\mathbf{G}}^y)^{-1})$), and then obtains $\mathbf{H}$ from (42) or (32) to get zero disturbance loss ($\mathbf{M}_d = 0$). ~~with~~

This expression applies to any number of measurements including $n_y < n_u + n_d$.

Alternatively, one may minimize the effect of ~~The effect~~ of combined disturbances and measurements ~~are minimized by selecting H according to~~ errors, ~~and~~ which is the "exact local method" of Halvorsen et al (2003). In this paper we ~~provide a~~ derive an explicit solution for H for this problem, see eq. (44).

② To simplify ~~implementation~~, one often uses only a subset of the available measurements when obtaining the combination c=Hy.
~~To simplify~~

Refs: Self-optimizing Original Skogestad (2000)

A simple rule, which aims at minimizing the effect of measurement error, is to select measurements to maximize ~~such that~~ $\underline{\sigma}(G^y)$ (or even better to minimize $\bar{\sigma}(J(G^y)^{-1}W_n^y)$). Here G^y is the steady-state gain matrix for 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, S. Skogestad, and V. Alstad. Perfect steady state indirect control. *Ind. Chem. Res.*, 44(4):291–309, 2005.

[5] 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.

[6] 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.

[7] 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.

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

[9] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control.* John Wiley & Sons, 1996.

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

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

[12] 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.

[13] 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.

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

[15] Y. Zhang, D. Monder, and J.F. Forbes. Real-time optimization under

V. Kariwala, "Optimal measurement combination for local self-optimizing control", Ind. Eng. Chem. Res., Vol. 46, 3629-3634 (2007).

Add Appendix - Proof of Theorem (two parts; more details on scalar case + Edward's generalization)

parametric uncertainty: A probability constrained approach. *J. Proc. Control*, 12(3):373–389, 2002.

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