Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 52-1 (2019) 64–69

An ADMM algorithm for incorporating structural constraints in self-optimizing control

Jonatan Ralf Axel Klemets * **Morten Hovd** **

* *Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway (e-mail: jonatan.klemets@ntnu.no).*

** *Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway (e-mail: morten.hovd@ntnu.no)*

Abstract: An ADMM algorithm is proposed for selecting structurally constrained measurement combinations as controlled variables (CVs). The CV selection is based on the self-optimizing control principle, where the goal is to choose CVs such that the steady-state operation is optimized when they are kept at constant set-point. When CV selection incorporates structural constraints, it becomes a non-convex optimization problem and thus, finding the optimal solution is difficult. However, using an ADMM algorithm for a given measurement set together with specified structural constraints, a local solution can be obtained. The resulting CVs seem to give similar or better performance when compared to other existing methods. The proposed method was evaluated on case studies, consisting of a binary distillation column and an evaporator.

Keywords: Control variable selection, self-optimizing control, structural constraints

1. INTRODUCTION

An increasingly competitive global market, together with stricter environmental and safety regulations make it necessary for chemical process plants to operate close to its optimum. As a result, there has been a growing interest in online optimization methods, e.g., model predictive control (MPC), real-time optimization (RTO), and economic MPC (EMPC). However, implementing such techniques remains challenging, mainly due to the computational complexity and lack of accurate dynamic models.

Another approach is to use simple control structures that keep specific controlled variables (CVs) at a constant value, also known as self-optimizing control (Skogestad, 2000). The central idea of self-optimizing control is to select CVs such that in the presence of disturbances, the loss is minimized by holding them at constant set-points.

Besides using single measurements, selecting linear combinations of measurements as CVs will further improve the self-optimizing control performance. Two methods that achieve this are the exact local method (Halvorsen et al., 2003) and the null-space method (Alstad and Skogestad, 2007). The null-space method makes the assumption that no implementation errors (e.g., no measurement noise) are present, which can be seen as unrealistic in practice. The exact local method, on the other hand, doesn't require any such assumptions, as it handles both implementation errors and disturbances.

Using all measurements available will, in theory, result in the best self-optimizing control performance, but increases the risk of getting sensor failures and makes implementing the control structure more difficult. However, finding the optimal subset of measurements can be computational demanding as it is a combinatorial optimization problem and, every possible alternative needs to be evaluated. Additionally, when using decentralized control, it is often desirable to impose some structural constraints on the CVs. E.g., by only combining manipulated variables (MVs) with CVs associated with certain units or parts of the process. Unfortunately, when including structural constraints on the measurement selection problem, it becomes non-convex. Therefore, finding the optimal measurement subset with structural constraints is still an open problem.

For the first part of the problem, a branch and bound method were derived in (Kariwala and Cao, 2009) to find the best measurement subset. A mixed integer quadratic programming (MIQP) approach was used by Yelchuru and Skogestad (2012). The MIQP formulation was expanded on in (Yelchuru and Skogestad, 2011) to handle the second part of the problem (incorporating structural constraint) by using a convex approximation. A generalized singular value decomposition (GSVD) method was proposed in (Heldt, 2010) for dealing with the structural constraints.

In this paper, the focus lies on the second part of the problem. An alternating direction method of multipliers (ADMM) algorithm is proposed for incorporating structural constraints in the CVs, assuming the measurement set has been given. The proposed ADMM algorithm is evaluated by obtaining structural CVs on a binary distillation column model, and an evaporator model, showing an improvement in the self-optimizing control performance compared to the other existing methods.

10.1016/j.ifacol.2019.06.038

2. SELF-OPTIMIZING CONTROL

Self-optimizing control is achieved when an acceptable loss is obtained with constant set-points without the need to reoptimize when (changes in) disturbances occur (Skogestad, 2000). More precisely, the aim is to select CVs rather than determining optimal set-points.

For specified disturbances (d), the problem for achieving optimal steady-state operation can be formulated as,

$$\min_{x,u} \; J\Big(x,u,d\Big) \tag{1}$$
$$\text{s.t. } f(x,u,d)=0 \tag{2}$$
$$g(x,u,d)\leq 0 \tag{3}$$
$$y=f_y(x,u,d) \tag{4}$$

where $x\in\mathbb{R}^{n_x}$, $u\in\mathbb{R}^{n_u}$, and $d\in\mathbb{R}^{n_d}$ are the states, inputs, and disturbances respectively. The equality constraints are represented by $f(\cdot)$ and contain the steady-state model equations; the inequality constraints in $g(\cdot)$ define the constraints on the operation, and the available measurements are given by y. The solution to the optimization problem usually results in some of the constraints being active, i.e., $g_i(x,u,d)=0$. To achieve optimal operation at steady-state, the variables related to the active constraints should be controlled and kept as close as possible to their optimal set-points. Stabilizing the plant and controlling the active constraints, therefore, requires a corresponding number of degrees of freedom. This results in a reduced space optimization problem:

$$\min_{u} J^*\Big(u,d\Big). \tag{5}$$

Here, the model equations and active constraints, are implicitly included in J^*. What remains is to determine which of the unconstrained variables should be kept constant by using the remaining inputs to minimize loss.

The loss L is defined as the difference between the actual value of a given cost function and the truly optimal value (accounting for the correct value of the disturbance), i.e.,

$$L(u,d)=J(u,d)-J_{opt}(d), \tag{6}$$

where the truly optimal operation is achieved when $L=0$. In general, $L\geq 0$ and thus smaller value for the loss, L implies that the plant is operating closer to its optimum.

To quantify the loss resulting from keeping the selected controlled variables at constant values, methods for calculating the worst case and average local loss were derived by Halvorsen et al. (2003), and Kariwala et al. (2008) respectively. The authors of (Kariwala et al., 2008) proved that selecting the controlled variables that minimize the average loss is super-optimal and hence, it also minimizes the worst case loss. The average loss is given by

$$L_{avg}=\frac{1}{2}\Big\|J_{uu}^{1/2}(HG^y)^{-1}HY\Big\|_F^2. \tag{7}$$

Here, $Y:=[FW_d\ W_n]$, with W_d and W_n representing the expected magnitudes of the disturbances and implementation errors respectively. $F=\frac{\partial y^{opt}}{\partial d}$ is the sensitivity matrix for the optimal deviations in the measurements (∂y^{opt}) with respect to changes in the disturbances (∂d); $J_{uu}=\frac{\partial^2 J}{\partial u^2}$ denotes the second derivative of the cost function in (5); and $G^y=\frac{\partial y}{\partial u}$, represents the gain from the inputs to the available measurements. The matrix $H\in\mathbb{R}^{n_u\times n_y}$ contains the measurements for the controlled variable $c=Hy$ that will be kept at a constant set-point.

2.1 Optimal Measurement Combination

Rather than selecting single measurements for the unconstrained optimization problem in (5), a further reduction in loss can be obtained by selecting the control variables as optimal linear measurements (with H being a full matrix).

The optimal linear combination can then be obtained by selecting H such that the loss in (7) gets minimized, i.e., the optimization problem can be formulated as

$$\min_H \frac{1}{2}\Big\|J_{uu}^{1/2}(HG^y)^{-1}HY\Big\|_F^2. \tag{8}$$

At first glance, this seems like a non-linear optimization problem. However, an important property was discovered by Alstad et al. (2009), who found that (8) can be recast as a convex optimization problem.

Theorem 1. If H is a full matrix (with no structural constraints) then (8) can be formulated as a convex constrained optimization problem (Alstad et al., 2009):

$$\min_H \frac{1}{2}\Big\|HY\Big\|_F^2 \tag{9}$$
$$\text{s.t. } HG^y=J_{uu}^{1/2} \tag{10}$$

Proof. From the original problem in (8), it can be shown that the solution for H is non-unique and for any non-singular matrix $Q\in\mathbb{R}^{n_u\times n_u}$, there exists $\hat{H}\in\mathbb{R}^{n_u\times n_y}$:

$$\hat{H}=Q^{-1}H \tag{11}$$

that gives the same loss. The non-uniqueness of H is used to add the constraint in (10), forcing the first part in (8) to become $J_{uu}^{1/2}(HG^y)^{-1}=I$. Hence, the nonlinear optimization in (8) can be recast as the convex optimization problem in (9) and (10). For a more detailed proof see e.g., (Jäschke et al., 2017), and (Alstad et al., 2009).

2.2 Measurement combinations with structural constraints

The lowest steady-state loss can be achieved when the measurement combination H is a full matrix. However, in many practical cases, it may be preferable to impose certain structural constraints on the measurement combinations. These structural constraints may be needed to avoid pairing MVs to CVs that are located far apart causing long time delay and thus, reducing the dynamic controllability. Furthermore, combining measurements that are of similar type (e.g., combining several temperature measurements) may be preferred by the operators as it has a more intuitive physical meaning.

These structural constraints on H the should be included in (8), such that the optimization problem becomes:

$$\min_H \frac{1}{2}\Big\|J_{uu}^{1/2}(HG^y)^{-1}HY\Big\|_F^2 \tag{12}$$
$$\text{s.t. } H\in\mathcal{S} \tag{13}$$

where $\mathcal{S}$ denotes the structural constraints that are imposed on H. Unfortunately, it is not possible to reformulate (12) and (13) such that they become a convex problem as in (9) and (10). This is due to there not being enough degrees of freedom to make $HG^y=J_{uu}^{1/2}$ when H is forced to have a certain structure.

To address this problem an iterative method was proposed by Heldt (2010) whereas in (Yelchuru and Skogestad, 2011), and (Yelchuru, 2012) two convex relaxation methods for solving (12) and (13) were suggested. However, these methods can't guarantee a global optimum, but rather an upper bound for the loss when using structural constraints. Therefore, in this paper, an alternative approach is suggested. The aim is to further reduce the loss for measurement combinations with structural constraints, by solving the following optimization problem:

$$\min_{H,Q,\hat{H}} \frac{1}{2}\left\|HY\right\|_F^2 \tag{14}$$

$$\text{s.t. } HG^y = J_{uu}^{1/2} \tag{15}$$

$$H = Q\hat{H} \tag{16}$$

$$\hat{H} \in \mathcal{S} \tag{17}$$

where $\hat{H} \in \mathbb{R}^{n_u \times n_y}$ contains the optimal measurement combination with the structural constraints $\mathcal{S}$. This, gives H enough degrees of freedom to satisfy the constraint in (15), while $Q \in \mathbb{R}^{n_u \times n_u}$ can be computed such that $Q^{-1}H$ (and $\hat{H}$) has the desired structure. Unfortunately, this is a non-convex problem and can, therefore, be difficult to solve. However, numerical experience indicates that this problem translates well to using alternating direction method of multipliers (ADMM).

3. ALTERNATING DIRECTION METHOD OF MULTIPLIERS

The alternating direction method of multipliers (ADMM) has extensively been studied since the 1970s and has shown to be a simple, yet a robust algorithm that is well suited for distributed convex optimization in large-scale problems. However, it can also be extended to non-convex problems, see Boyd et al. (2011) for a survey of ADMM and its different applications. Recently, there has also been an increasing interest in using ADMM for sparsity promoting functions (Lin et al., 2012), (Dhingra et al. (2014)).

Let's again consider the non-convex optimization in (14) − (17). The *augmented Lagrangian* associated with (14) and the constraint (16) can be formulated as

$$\mathcal{L}(H,Q,\hat{H},\Lambda) = \frac{1}{2}\left\|HY\right\|_F^2 + trace(\Lambda^T(H - Q\hat{H})) + \frac{\rho}{2}\left\|H - Q\hat{H}\right\|_F^2 \tag{18}$$

where Λ is the dual variable (*Lagrange multiplier*) and ρ is a positive scalar. The *augmented Lagrangian* in (18) can be expressed more conveniently by combining the linear and quadratic terms (Boyd et al., 2011):

$$\mathcal{L}(H,Q,\hat{H},\Lambda) = \frac{1}{2}\left\|HY\right\|_F^2 + \frac{\rho}{2}\left\|H - Q\hat{H} + \frac{1}{\rho}\Lambda\right\|_F^2. \tag{19}$$

The ADMM algorithm solves the problem in (14) − (17), by iteratively solving,

$$H, \hat{H}^{k+1} := \arg\min_{H,\hat{H}} \mathcal{L}(H, Q^k, \hat{H}, \Lambda^k), \tag{20}$$

$$\text{s.t. } (15), \text{and } (17)$$

$$H, Q^{k+1} := \arg\min_{H,Q} \mathcal{L}(H, Q, \hat{H}^{k+1}, \Lambda^k), \tag{21}$$

$$\text{s.t. } (15)$$

$$\Lambda^{k+1} := \Lambda^k + \rho(H - Q^{k+1}\hat{H}^{k+1}) \tag{22}$$

until $\|H - Q\hat{H}\|_F \leq \epsilon$. Thus, the optimal $\hat{H}$, H, and Q are solved in an alternating fashion, hence the name alternating direction.

There are two major benefits for using ADMM on (14) - (17). First, it temporarily relaxes the equality constraints in (16), thus, allows for more flexibility when searching for the optimal solution. Secondly, by separating $\hat{H}$ from Q, both the step in (20) and the step in (21) become convex quadratic optimization problems with equality constraints. Therefore, there exist analytical solutions to both these steps for improved computational efficiency.

3.1 Analytical solution to the (21) subproblem

After dropping superscripts for notational simplicity, the H, Q minimization step in (21) becomes

$$\min_{H,Q} \mathcal{L}(H,Q,\hat{H},\Lambda) \tag{23}$$

$$\text{s.t. } HG^y = J_{uu}^{1/2} \tag{24}$$

Theorem 2. Under the assumption that YY^T is full rank and $\hat{H}$ is of full row rank, then an analytical solution to (23) can be obtained for H and Q:

$$H^T = (\Phi^{-1} - \Phi^{-1}G^y(G^{y^T}\Phi^{-1}G^y)^{-1}G^{y^T}\Phi^{-1})\Omega + (\Phi^{-1}G^y(G^{y^T}\Phi^{-1}G^y)^{-1})J_{uu}^{1/2} \tag{25}$$

$$Q^T = (\hat{H}\hat{H}^T)^{-1}\hat{H}(\frac{1}{\rho}\Lambda^T + H^T) \tag{26}$$

where

$$\Phi := YY^T + \rho I - \rho\hat{H}^T(\hat{H}\hat{H}^T)^{-1}\hat{H}$$
$$\Omega := \hat{H}^T(\hat{H}\hat{H}^T)^{-1}\hat{H}\Lambda^T - \Lambda^T$$

Proof. Solving (23) with respect to Q is an unconstrained quadratic optimization problem and the optimal solution can be obtained from:

$$\frac{\partial\mathcal{L}}{\partial Q} = (\hat{H}\hat{H}^T)^T Q^T - \frac{1}{\rho}\hat{H}\Lambda^T - \hat{H}H^T = 0$$

for which the solution for Q^T is equivalent to (26). To find a solution for (23) and (24) w.r.t. H, the problem must satisfy the following KKT-conditions (Nocedal and Wright, 1999):

$$\begin{bmatrix} YY^T + \rho I & -G^y \\ G^{y^T} & 0 \end{bmatrix}\begin{bmatrix} H^T \\ \lambda^T \end{bmatrix} = \begin{bmatrix} \rho\hat{H}^T Q^T - \Lambda^T \\ J_{uu}^{1/2} \end{bmatrix} \tag{27}$$

where λ is the *Lagrange multiplier* for the constraint in (24). Replacing Q with the solution from (26), the KKT conditions in (27) can be rewritten to:

$$\begin{bmatrix} \Phi & -G^y \\ G^{y^T} & 0 \end{bmatrix}\begin{bmatrix} H^T \\ \lambda^T \end{bmatrix} = \begin{bmatrix} \Omega \\ J_{uu}^{1/2} \end{bmatrix} \tag{28}$$

The optimal H^T can then be found by inverting the KKT-matrix using the Schur complement for the inverse of block partitioned matrices (see e.g., Lu and Shiou (2002)).

3.2 Analytical solution to the (20) subproblem

The solution to (20) is identical to solving

$$\min_{H,\hat{H}} \mathcal{L}(H,Q,\hat{H},\Lambda) \tag{29}$$

$$\text{s.t. } HG^y = J_{uu}^{1/2} \tag{30}$$

$$\Gamma\ vec(\hat{H}^T) = 0 \tag{31}$$

Table 1. CVs and their respective loss for the distillation column when using a triangular and a block diagonal structure. In $Method$ 1 and $Method$ 2 ($Method$ 1, 2 is used if the results were almost identical) CVs were obtained using the methods in (Yelchuru, 2012). In $ADMM$, the CVs are computed using the proposed ADMM algorithm.

		Controlled variables (CVs)	Loss	(Full H)
Triangular Structure	S_1 : $Method$ 1, 2 :	$c = \left[T_{30} + 0.9887T_{31};\ T_{11} + 0.7365T_{30} + 0.7812T_{31}\right]$	0.4640	(0.4638)
	S_1 : $ADMM$:	$c = \left[T_{30} + 1.0239T_{31};\ T_{11} - 0.1300T_{30} - 0.1060T_{31}\right]$	0.4638	
	S_2 : $Method$ 1, 2 :	$c = \left[0.63T_{30} + 0.6229T_{31};\ -0.3463T_{10} - 0.3484T_{11} - 0.2390T_{30} - 0.2680T_{31}\right]$	0.3529	(0.3525)
	S_2 : $ADMM$:	$c = \left[0.63T_{30} + 0.6456T_{31};\ -0.3463T_{10} - 0.3720T_{11} + 0.0003T_{30} - 0.0237T_{31}\right]$	0.3526	
	S_3 : $Method$ 1 :	$c = \left[f(T_{21} \cdots T_{41});\ f(T_1 \cdots T_{41})\right]$	0.1410	
	S_3 : $Method$ 2 :	$c = \left[f(T_{21} \cdots T_{41});\ f(T_1 \cdots T_{41})\right]$	0.0940	(0.0813)
	S_3 : $ADMM$:	$c = \left[f(T_{21} \cdots T_{41});\ f(T_1 \cdots T_{41})\right]$	0.0881	
Block diagonal Structure	S_1 : $Method$ 1, 2 :	$c = \left[0.63T_{30} + 0.6229T_{31};\ 0.9675T_{12}\right]$	0.4427	(0.4425)
	S_1 : $ADMM$:	$c = \left[0.63T_{30} + 0.6282T_{31};\ 0.9675T_{12}\right]$	0.4427	
	S_2 : $Method$ 1, 2 :	$c = \left[0.63T_{30} + 0.6229T_{31};\ -0.5151T_{11} - 0.5110T_{12}\right]$	0.3441	(0.3437)
	S_2 : $ADMM$:	$c = \left[0.63T_{30} + 0.6272T_{31};\ -0.5151T_{11} - 0.5140T_{12}\right]$	0.3441	
	S_3 : $Method$ 1 :	$c = \left[f(T_{21} \cdots T_{41});\ f(T_1 \cdots T_{20})\right]$	0.1270	
	S_3 : $Method$ 2 :	$c = \left[f(T_{21} \cdots T_{41});\ f(T_1 \cdots T_{20})\right]$	0.1050	(0.0813)
	S_3 : $ADMM$:	$c = \left[f(T_{21} \cdots T_{41});\ f(T_1 \cdots T_{20})\right]$	0.1039	

with $vec(\hat{H}^T)$ being the vectorization of $\hat{H}^T$. Γ is a matrix consisting of only ones and zeros that is the orthogonal complement to the structural constraints imposed on $vec(\hat{H}^T)$. E.g., if the structural constraints for $\hat{H}$ are:

$$\hat{H} \in \mathcal{S} := \begin{bmatrix} \hat{h}_{1,1} & 0 & \hat{h}_{1,3} \\ 0 & \hat{h}_{2,2} & 0 \end{bmatrix},$$

then Γ should be chosen such that $\Gamma vec(\hat{H}^T)$ contains the elements that should be set to zero, i.e.,

$$\Gamma\ vec(\hat{H}^T) = 0 \equiv \left[\hat{h}_{1,2}\ \hat{h}_{2,1}\ \hat{h}_{2,3}\right]^T = [0\ 0\ 0]^T.$$

Theorem 3. Under the assumption of Q, and YY^T being full rank, then an analytical solution for (29) − (31) can be obtained for $\hat{H}$ and H:

$$vec(\hat{H}^T) = \zeta\ vec(\Xi) \tag{32}$$

$$H^T = \Upsilon(\rho \hat{H}^T Q^T - \Lambda^T) + \Psi^{-1} G^y (G^{y^T} \Psi^{-1} G^y)^{-1} J_{uu}^{1/2} \tag{33}$$

where Ξ, ζ, Ψ, and Υ are defined

$$\Xi := \Upsilon \Lambda^T Q - \Psi^{-1} G^y (G^{y^T} \Psi^{-1} G^y)^{-1} J_{uu}^{1/2} Q - \frac{1}{\rho} \Lambda^T Q \tag{34}$$

$$\zeta := \alpha^{-1} - \alpha^{-1} \Gamma^T (\Gamma \alpha^{-1} \Gamma^T)^{-1} \Gamma \alpha^{-1} \tag{35}$$

$$\Upsilon := \Psi^{-1} - \Psi^{-1} G^y (G^{y^T} \Psi^{-1} G^y)^{-1} G^{y^T} \Psi^{-1} \tag{36}$$

$$\Psi := YY^T + \rho I \tag{37}$$

and α is

$$\alpha = Q^T Q \otimes (\rho \Upsilon - I), \tag{38}$$

with $\otimes$ being the Kronecker product.

Proof. The proof for (32), and (33) follows a similar procedure as in the proof for Theorem 2, where the KKT conditions are first formulated to solve (29) − (31) w.r.t. H, and $vec(\hat{H}^T)$. The resulting KKT matrices w.r.t H is equivalent to the one in (27). Solving (27) for H and replacing the results in the KKT conditions for $vec(\hat{H}^T)$ gives,

$$\begin{bmatrix} \alpha & \Gamma^T \\ -\Gamma & 0 \end{bmatrix} \begin{bmatrix} vec(\hat{H}^T) \\ \lambda \end{bmatrix} = \begin{bmatrix} vec(\Xi) \\ 0 \end{bmatrix} \tag{39}$$

from which the results in (32) can be obtained.

3.3 The proposed ADMM algorithm

The resulting ADMM algorithm can be seen in Algorithm 1, where first an initial value for Q (e.g., $Q = I$) needs to be set, together with a positive scalar for the parameter ρ.

Algorithm 1 ADMM for structural constraints

Initialize: For $k = 1$, select, $Q^k = I$, and choose a positive scalar ρ.

1: Calculate $\hat{H}^{k+1}$ using (32).
2: Calculate H^{k+1}, and Q^{k+1} using (25) and (26).
3: Update Λ^{k+1} using (22).
4: If $\|H^{k+1} - Q^{k+1}\hat{H}^{k+1}\|_F \leq \epsilon$ or $k \geq MaxIter$ stop, else set k to $k+1$ and repeat step 1 to 4.

It is important to note that for non-convex problems the ADMM algorithm may not converge to the globally optimal solution, in fact, it may not converge at all. Thus, it should only be considered as a local optimization method. Whereas global convergence of ADMM can be guaranteed for convex problems, this is not the case when dealing with non-convex problems. However, the ADMM algorithm seems to be able to converge in most cases as long as the value of ρ is chosen to be sufficiently large. Furthermore, even when it converges, the final result can depend on how the initial values Q^k and ρ were chosen. However, since all the steps can be solved analytically, and a solution can be obtained relatively fast, it should be easy

Table 2. CVs for the Evaporator and their respective loss. In $Method\ 1,2$ the CVs have been obtained using the methods in (Yelchuru, 2012), in $GSVD$ the results from (Heldt, 2010) is used as CVs, and in $ADMM$ the CVs are computed using the proposed ADMM algorithm.

	Controlled variables (CVs)		Loss	(Full H)
S_1 :	$GSVD$:	$c = \left[-6.27F_2 + F_{100};\ F_{200} - 23.3F_1\right]$	11.9070	(10.679)
	$Method\ 1,2$:	$c = \left[-6.259F_2 + F_{100};\ F_{200} - 21.723F_1\right]$	11.8640	
	$ADMM$:	$c = \left[-6.238F_2 + F_{100};\ F_{200} - 21.530F_1\right]$	11.8473	
S_2 :	$Method\ 1,2$:	$c = \left[T_2;\ F_5 - 0.038F_{200}\right]$	58.6848	(56.788)
	$ADMM$:	$c = \left[T_2;\ F_5 - 0.080F_{200}\right]$	58.2074	
S_3 :	$Method\ 1,2$:	$c = \left[-6.259F_2 + F_{100};\ F_5 - 0.039F_{200}\right]$	11.9307	(9.9538)
	$ADMM$:	$c = \left[-6.2375F_2 + F_{100};\ F_5 - 0.0391F_{200}\right]$	11.9160	
S_4 :	$Method\ 1,2$:	$c = \left[P_2 + 117.795F_5 - 4.945F_{200};\ F_2 - 0.032F_3\right]$	31.7480	(9.2494)
	$ADMM$:	$c = \left[P_2 - 6.7287F_5 - 0.0998F_{200};\ F_2 - 0.0321F_3\right]$	20.3317	
S_5 :	$Method\ 1,2$:	$c = \left[f(T_2, F_2, F_{100}, F_3.F_1);\ f(P_2, T_3.T_{201}, F_5, F_{200})\right]$	9.2450	(7.5499)
	$ADMM$:	$c = \left[f(T_2, F_2, F_{100}, F_3.F_1);\ f(P_2, T_3.T_{201}, F_5, F_{200})\right]$	8.8847	

to try different initial values until a suitable solution has been found. Therefore, the hope is that it may be able to provide better results than other methods when obtaining structured measurement combinations.

4. CASE STUDIES

4.1 Distillation column case study

The proposed method for obtaining measurement combinations with specified structures was evaluated on a binary distillation column model (Skogestad, 1997). The column consists of 41 stages (with a temperature measurement on each stage) and uses an LV-configuration, where the reflux flow rate (L) and the vapor boilup rate (V) are the two remaining degrees of freedom. Thus, the available inputs (u) and measurements (y) are:

$$u = [L\ V]^T \tag{40}$$

$$y = [T_1\ \cdots\ T_{41}]^T. \tag{41}$$

The main disturbances (d) are changes in feed flow rate (F), feed composition (z_F), and feed liquid fraction (q_F):

$$d = [F\ z_F\ q_F]^T. \tag{42}$$

The objective is to get a top product with 99% light component (1% heavy) and a bottom product with 1% light component, i.e., the cost function is,

$$J = \left(\frac{x_H^{top} - x_H^{top,s}}{x_H^{top,s}}\right)^2 + \left(\frac{x_L^{btm} - x_L^{btm,s}}{x_L^{btm,s}}\right)^2 \tag{43}$$

where the specifications for the top and bottom products are denoted with the superscript, s.

In (Yelchuru, 2012) a MIQP formulation was used to acquire CVs with specified structures, that consisted of either a triangular or a block diagonal structure. Besides trying to find structural CVs, Yelchuru (2012) also tried to find the best measurement subsets for these structures. Using the same measurement sets with the same structural constraints, the proposed ADMM algorithm is used with the aim of finding CVs with a smaller loss compared to the results in (Yelchuru, 2012).

The resulting CVs and their respective loss are shown in Table 1. Three different measurement sets S_1, S_2, and S_3 are considered, consisting of 3, 4, and 41 measurements respectively when using both triangular and block diagonal structural constraints. Based on the results, the proposed ADMM algorithm is able to provide CVs with a loss that is at least on par with the ones proposed in (Yelchuru, 2012), but with a better result for the triangular case and when more measurements are used.

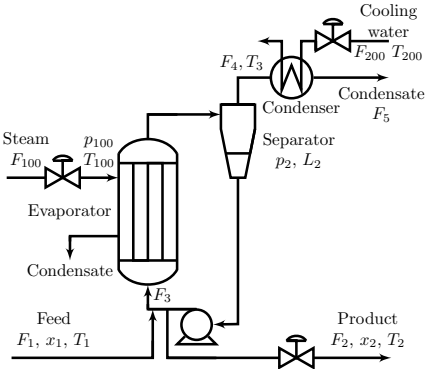

Fig. 1. The evaporator process flowsheet.

4.2 Evaporator case study

The next case study consists of the evaporation process represented in Fig. 1. The evaporator was originally treated by Newell and Lee (1989) and has been modified in (Kariwala et al., 2008). The process uses 2 inputs u, 10 potential measurements y, with the 3 disturbances d being changes in compositions and temperature of the inflows:

$$u = [F_{200}\ F_1]^T \tag{44}$$

$$y = [p_2\ T_2\ T_3\ F_2\ F_{100}\ T_{201}\ F_3\ F_5\ F_{200}\ F_1]^T \tag{45}$$

$$d = [x_1\ T_1\ T_{200}]^T \tag{46}$$

The economic objective of the evaporator is to maximize the operating profit [\$/h] and has been formulated in (Kariwala et al., 2008):

$$J = 600F_{100} + 0.6F_{200} + 1.009(F_2 + F_3) + 0.2F_1 - 4800F_2 \tag{47}$$

The CV selection containing structural constraints has been studied in (Heldt, 2010), and (Yelchuru, 2012). These results will be used as a benchmark to compare the performance with the proposed ADMM algorithm. Five different structures (S_1 - S_5) are investigated with the first structure being proposed in (Heldt, 2010). The remaining structures (S_2 - S_5) that were considered have been obtained by Yelchuru (2012), in which CVs had been designed to separate the measurements associated with the condenser $(p_2, T_3, T_{201}, F_5, F_{200})$ and the measurements associated to the evaporator $(T_2, F_2, F_{100}, F_3, F_1)$.

Using the proposed ADMM algorithm, CVs are found for the same measurements and structural constraints. The resulting CVs and their steady-state loss are shown in Table 2. The CVs computed using the ADMM algorithm are able to give a smaller loss, in particular for structures S_4, and S_5 compared to the other CVs.

4.3 Initialization of the ADMM algorithm

As briefly mentioned in section 3.3, the convergence and the final results for the ADMM algorithm depends on how it is initialized (i.e., how Q^k, and ρ are chosen). Thus, different values of ρ together with randomly generated initial values for Q^k has been investigated. In Table 3, the resulting losses (L) are displayed when using different ρ and initial values Q^k. As expected, if ρ is chosen to be too small ($\rho = 10$), the ADMM algorithm isn't able to converge. However, interestingly, the lowest loss is obtained when ρ is chosen to be between 10^2 and 10^6, whereas the loss seems to increase as ρ chosen to large (10^7). Furthermore, when ρ is chosen to be sufficiently small (while still large enough for the ADMM algorithm to converge), the initial values of Q^k seems to have little impact on the final result. This property has also been identified for the other structures in both the distillation and the evaporator case studies.

Table 3. Loss (L) for the Evaporator with structure S_5 using different initial values for the ADMM algorithm.

Q	$\rho = 10$	$\rho = 10^2$	$\rho = 10^6$	$\rho = 10^7$
$\begin{bmatrix}1 & 0; & 0 & 1\end{bmatrix}$	-	$L = 8.8847$	$L = 8.8847$	$L = 42.237$
$\begin{bmatrix}0.7 & 0.9; & 2 & 0.1\end{bmatrix}$	-	$L = 8.8847$	$L = 8.8847$	$L = 8.8873$
$\begin{bmatrix}-3 & 0.5; & -1 & 6\end{bmatrix}$	-	$L = 8.8847$	$L = 8.8847$	$L = 46.231$
$\begin{bmatrix}0.3 & -2; & 9 & -4\end{bmatrix}$	-	$L = 8.8847$	$L = 8.8847$	$L = 8.9586$

5. CONCLUSION

In this work, an ADMM algorithm for incorporating structural constraints in self-optimizing control variables has been investigated. It may thus, serve as an alternative to the approaches in (Heldt, 2010) and (Yelchuru, 2012) when trying to compute CVs with specified structures. A major benefit with the referenced methods is that they can also be used for finding the optimal measurement subset, which is something the ADMM algorithm is not able to do. However, it has been demonstrated on two different case studies that for a given measurement set and CV structure, the proposed method is able to obtain CVs with a smaller steady-state loss than the other existing methods.

REFERENCES

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

Alstad, V., Skogestad, S., and Hori, E.S. (2009). Optimal measurement combinations as controlled variables. *Journal of Process Control*, 19(1), 138–148.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1), 1–124.

Dhingra, N.K., Jovanović, M.R., and Luo, Z.Q. (2014). An ADMM algorithm for optimal sensor and actuator selection. *In Proc. of 53rd IEEE Conference on Decision and Control*, 4039–4044.

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

Heldt, S. (2010). Dealing with structural constraints in self-optimizing control engineering. *Journal of Process Control*, 20(9), 1049–1058.

Jäschke, J., Cao, Y., and Kariwala, V. (2017). Self-optimizing control – A survey. *Annual Reviews in Control*, 43, 199–223.

Kariwala, V. and Cao, Y. (2009). Bidirectional branch and bound for controlled variable selection. part II: Exact local method for self-optimizing control. *Computers and Chemical Engineering*, 33(8), 1402–1414.

Kariwala, V., Cao, Y., and Janardhanan, S. (2008). Local self-optimizing control with average loss minimization. *Industrial and Engineering Chemistry Research*, 47(4), 1150–1158.

Lin, F., Fardad, M., and Jovanovic, M.R. (2012). Sparse feedback synthesis via the alternating direction method of multipliers. *American Control Conference*, 4765–4770.

Lu, T.T. and Shiou, S.H. (2002). Inverses of 2×2 block matrices. *Computers and Mathematics with Applications*, 43(1-2), 119–129.

Newell, R.B. and Lee, P. (1989). Applied process control: A case study. *Prentice-Hall of Australia, New York; Sydney.*

Nocedal, J. and Wright, S. (1999). Numerical optimization. *Springer.*

Skogestad, S. (1997). Dynamics and control of distillation columns: A tutorial introduction. *Chemical Engineering Research and Design*, 75(6), 539–562.

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure. *Journal of Process Control*, 10(5), 487–507.

Yelchuru, R. (2012). Quantitative methods for controlled variables selection. *PhD thesis, NTNU.*

Yelchuru, R. and Skogestad, S. (2011). Optimal controlled variable selection with structural constraints using MIQP formulations. *Proceedings of the 18th World Congress The International Federation of Automatic Control*, 4977–4982.

Yelchuru, R. and Skogestad, S. (2012). Convex formulations for optimal selection of controlled variables and measurements using mixed integer quadratic programming. *Journal of Process Control*, 12(6), 995–1007.