I&EC research
Industrial & Engineering Chemistry Research

Article

pubs.acs.org/IECR

Self-Optimizing Control Structures with Minimum Number of Process-Dependent Controlled Variables

Alejandro G. Marchetti* and David Zumoffen[†]

French-Argentine International Center for Information and Systems Sciences (CIFASIS), CONICET-Universidad Nacional de Rosario (UNR), 27 de Febrero 210bis, S2000EZP Rosario, Argentina

ABSTRACT: In order to operate continuous processes near the economically optimal steady-state operating point, self-optimizing control schemes reformulate the optimization problem as a process control problem. The objective is to find controlled variables and constant set points such that the controller leads to optimal adjustments of the inputs in the presence of stable disturbances. In particular, the null space approach consists in selecting the self-optimizing controlled variables as linear combinations of the inactive output variables, based on the first-order variation of the necessary conditions of optimality. In the self-optimizing control structures proposed in the literature, the number of controlled variables required is typically equal to the number of degrees of freedom (inputs) that are available after all the equality and active inequality constrained variables are controlled. In this paper, we propose new self-optimizing control structures based on the null space approach, where depending on the number of disturbances, the number of active constraints, and the number of inputs, it is possible to decrease the number of process-dependent controlled variables by fixing linear combinations of the inputs. The effectiveness of the proposed self-optimizing control structures with minimum number of process-dependent controlled variables is demonstrated in simulation by means of a continuous stirred tank reactor and an evaporator.

1. INTRODUCTION

Steady-state optimization is of economic importance in many industrial process plants that run in continuous operation. Several techniques have been proposed in order to achieve optimal steady-state operation.[1] A typical approach is to use a hierarchical structure involving several layers that include plant scheduling, real-time optimization (RTO), and process control.[2,3] At the RTO layer, a detailed steady-state model is used to compute the optimal operating point by solving a nonlinear program (NLP). RTO typically proceeds using a *two-step approach*,[2,3] which consists of an iteration between parameter estimation and optimization. At the current operating point, the uncertain model parameters and disturbances are identified and used to update the model to generate new inputs via optimization.[2]

As an alternative to solving the parameter estimation and economic optimization problems online, several *reformulation methods* have been proposed, where the optimization problem is recast as the problem of tracking selected variables whose optimal values are approximately invariant to uncertainty. These include *constraint control*,[4] *optimizing control*,[5−7] *self-optimizing control*,[8,9] and tracking the necessary conditions of optimality (NCO), which is known as *NCO-tracking*.[10,11]

In self-optimizing control (SOC) schemes, the process model is exploited off-line in order to find a control structure such that an acceptable economic loss is achieved by keeping the controlled variables at constant set point values.[8] For given values of the disturbance variables, the economic loss associated with a given SOC structure is defined as the difference between the value of the cost function obtained upon implementing the SOC strategy and the true optimal value of the cost function. Using a first-order variation of the NCO for an unconstrained optimization problem and a linearized model of the output variables, Halvorsen et al.[12] derived a local expression for the loss in terms of the disturbance values and measurement error. Based on this result, *local methods* were proposed for finding an optimal linear combination of measurements that minimizes the (local) worst-case loss,[12] and the (local) average loss.[13] Many contributions to the SOC literature can be found that build on these two approaches. Kariwala[14] proposes an algorithm for finding the combination matrix that minimizes the worst-case loss. Heldt[15] considers a class of structural constraints on the combination matrix, wherein each controlled variable consists of the linear combination of an individual subset of measurements. The worst-case loss[12] and the average loss[13] approaches will, in general, result in smaller losses by increasing the number of measurements used in the linear combination matrix. However, since the improvements in the loss may quickly become negligable, it has been proposed to fix the number of measurements and to find the best subset of measurements that minimizes the loss.[16] This represents a combinatorial problem.

Also based on the first-order variation of the NCO for unconstrained optimization problems, Alstad and Skogestad[17] proposed the *null space method* for selecting self-optimizing controlled variables as linear combinations of the measurement variables. The approach consists in selecting linear combinations that lie in the null space of the sensitivity matrix of the optimal outputs with respect to the disturbances. In the absence of measurement error, Alstad et al.[18] have shown that the null space method zeroes the local loss expression derived by Halvorsen et al.[12] On the other hand, in the presence of

Received: December 19, 2013
Revised: April 30, 2014
Accepted: May 1, 2014

   dx.doi.org/10.1021/ie5010509 | *Ind. Eng. Chem. Res.* XXXX, XXX, XXX−XXX

measurement error and extra measurements, it is possible to select a subset of measurements to be used in the null space method such that the local loss due to measurement error is minimized.[18] Links between SOC using the null space method and the two-step approach of RTO have been discussed.[19] Most approaches for formulating a SOC problem are based on the assumption that the set of active constraints at the optimum does not change when the disturbances change. Changes in the active set have been considered by reformulating the SOC problem for each active set region.[20]

In the SOC structures proposed in the literature, the number of self-optimizing controlled variables is typically equal to the number of degrees of freedom that are available after controlling all the equality and active inequality constrained output variables.[8,9,12,13,17,18] In this paper, we show that in cases where the number of disturbances is lower than the number of inactive input variables, it is possible to decrease the number of process-dependent controlled variables in self-optimizing control strategies based on the null space method, by fixing linear combinations of the inputs. Based on this idea, this paper presents SOC structures based on the null space method with minimum number of process-dependent controlled variables, wherein as many linear combinations of the inputs as possible are fixed, and the remaining linear input combinations are used as manipulated variables for controlling the self-optimizing controlled variables, which include the active constrained output variables and linear combinations of inactive input and output variables. Reducing the number of process-dependent controlled variables permits to reduce the dimension of the dynamic control problem that must be solved in SOC strategies based on the null space method.

The paper is organized as follows. Section 2 formulates the constrained optimization problem and presents the first-order variation of the necessary conditions of optimality. Based on this variational analysis, the null space method proposed by Alstad and Skogestad[17] is reformulated in section 3 for the case of a constrained optimization problem. The self-optimizing control structures with minimum number of controlled variables are presented in section 4 for different scenarios that depend on the problem dimensions. The applicability of the proposed SOC structures is illustrated in simulation in section 5 for a continuous stirred tank reactor and a forced-circulation evaporator. Finally, section 6 concludes the paper.

2. PROBLEM FORMULATION

2.1. Optimization Problem. The behavior of the plant is represented by a nonlinear state-space model of the form:

$$\begin{aligned}\dot{\mathbf{x}}_p &= \mathbf{f}_p(\mathbf{x}_p, \mathbf{u}, \mathbf{d}_p)\\ \mathbf{y}_p &= \mathcal{F}_p(\mathbf{x}_p)\end{aligned} \tag{1}$$

where $\mathbf{x}_p \in \mathbf{R}^{n_{xp}}$ is the vector of state variables, $\mathbf{u} \in \mathbf{R}^{n_u}$ is the vector of decision (or input) variables, $\mathbf{y}_p \in \mathbf{R}^{n_y}$ is the vector of measured (or output) variables, and $\mathbf{d}_p \in \mathbf{R}^{n_{dp}}$ is the vector of process disturbances. The notation $(\cdot)_p$ is used for the variables associated with the plant.

In any practical application, the input–output mapping corresponding to the operation of the plant at steady state, which is represented here as $\mathbf{y}_p = \mathbf{H}_p(\mathbf{u},\mathbf{d}_p)$, is not known precisely, and only an approximate steady-state model is available:

$$\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{d}) = \mathbf{0} \tag{2a}$$

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \mathbf{d}) \tag{2b}$$

where $\mathbf{x} \in \mathbf{R}^{n_x}$ are the model state variables, and $\mathbf{d} \in \mathbf{R}^{n_d}$ is a set of model parameters. We denote by $\mathbf{y} = \mathbf{H}(\mathbf{u},\mathbf{d})$ the input–output mapping representing the steady-state behavior predicted by the model. In order to obtain $\mathbf{y}$ one has to first solve the model eq 2a in order to compute the states $\mathbf{x}$ and then obtain $\mathbf{y}$ by evaluating eq 2b.

The steady-state economic optimum of the plant is given by the solution of the following optimization problem:

$$\begin{aligned}\mathbf{u}_p^\star(\mathbf{d}_p) = \arg\min_{\mathbf{u}} \quad & \Phi_p(\mathbf{u}, \mathbf{d}_p) := \phi(\mathbf{u}, \mathbf{H}_p(\mathbf{u}, \mathbf{d}_p))\\ \text{s.t.} \quad & \mathbf{y}_p^{\text{eq}} = \mathbf{H}_p^{\text{eq}}(\mathbf{u}, \mathbf{d}_p) = \mathbf{y}^{\text{S}},\\ & \mathbf{y}^{\text{L}} \leq \mathbf{y}_p^{\text{in}} = \mathbf{H}_p^{\text{in}}(\mathbf{u}, \mathbf{d}_p) \leq \mathbf{y}^{\text{U}},\\ & \mathbf{u}^{\text{L}} \leq \mathbf{u} \leq \mathbf{u}^{\text{U}}\end{aligned} \tag{3}$$

where $\phi: \mathbf{R}^{n_u} \times \mathbf{R}^{n_y} \to \mathbf{R}$ is the scalar cost function to be minimized; $\mathbf{y}_p^{\text{eq}} \in \mathbf{R}^{n_y^{eq}}$ is the set of equality constrained outputs for which $\mathbf{y}^{\text{S}}$ are the set point values; $\mathbf{y}_p^{\text{in}} \in \mathbf{R}^{n_y^{in}}$ is the set of inequality constrained outputs for which $\mathbf{y}^{\text{L}}$ and $\mathbf{y}^{\text{U}}$ are the lower and upper bounds, respectively; and $\mathbf{u}^{\text{L}}$ and $\mathbf{u}^{\text{U}}$ are the lower and upper bounds on the decision variables. The measured outputs $\mathbf{y}_p$ are partitioned into the controlled outputs $\mathbf{y}_p^{\text{eq}}$ and the remaining outputs $\mathbf{y}_p^{\text{in}}$, for which lower and upper bounds can always be found. Hence, $n_y = n_y^{\text{eq}} + n_y^{\text{in}}$.

It is assumed that $n_y^{\text{eq}} < n_u$, that is, the number of equality constrained outputs is lower than the number of inputs. Without this assumption there are no degrees of freedom available for optimization. Also, we assume that there exists a feasible solution to problem 3 for any $\mathbf{d}_p \in \mathcal{D}$, where the disturbance set $\mathcal{D}$ comprises the disturbances that may be encountered during the operation of the plant.

The plant mapping $\mathbf{H}_p(\mathbf{u},\mathbf{d}_p)$ is not known accurately, and only the approximate model $\mathbf{y} = \mathbf{H}(\mathbf{u},\mathbf{d})$ is available. Using the model, the solution of the original problem 3 can be approached by solving the following NLP problem:

$$\begin{aligned}\mathbf{u}^\star(\mathbf{d}) = \arg\min_{\mathbf{u}} \quad & \Phi(\mathbf{u}, \mathbf{d}) := \phi(\mathbf{u}, \mathbf{H}(\mathbf{u}, \mathbf{d}))\\ \text{s.t.} \quad & \mathbf{y}^{\text{eq}} = \mathbf{H}^{\text{eq}}(\mathbf{u}, \mathbf{d}) = \mathbf{y}^{\text{S}},\\ & \mathbf{y}^{\text{L}} \leq \mathbf{y}^{\text{in}} = \mathbf{H}^{\text{in}}(\mathbf{u}, \mathbf{d}) \leq \mathbf{y}^{\text{U}},\\ & \mathbf{u}^{\text{L}} \leq \mathbf{u} \leq \mathbf{u}^{\text{U}}\end{aligned} \tag{4}$$

Let us consider the following model accuracy assumption:

Assumption 1 (No Structural Plant-Model Mismatch). The model equations are structurally correct, that is, $\Phi = \Phi_p$, and $\mathbf{H} = \mathbf{H}_p$. In other words, it is assumed that plant-model mismatch is only originated by differences in the values of the model parameters $\mathbf{d}$.

Since self-optimizing control is based on this assumption, only the optimization problem 4 will be considered in the development of the approach proposed in this paper.

2.2. Sensitivity Analysis. In order to conduct a sensitivity analysis of problem 4, the following invariant active set assumption is considered:

Assumption 2 (Invariant Active Set). Locally, the set of active constraints of problem 4 does not change with the process disturbances $\mathbf{d}$.

Let us assume that at the optimal point $\mathbf{u}^\star(\mathbf{d})$ there are m^L inequality constrained outputs at their lower bounds, m^U inequality constrained outputs at their upper bounds, n^L inputs at their lower bounds, and n^U inputs at their upper bounds. We define the matrix $P^{yL} \in \mathbf{R}^{m^L \times n_y^{in}}$ such that each row of P^{yL} has a one at the index number corresponding to an inequality constrained output that is active at its lower bound and zeros elsewhere. Similarly, the matrix $P^{yU} \in \mathbf{R}^{m^U \times n_y^{in}}$ is defined for the inequality constrained outputs that are active at their upper bounds, and the matrices $P^L \in \mathbf{R}^{n^L \times n_u}$ and $P^U \in \mathbf{R}^{n^U \times n_u}$ are defined for the inputs that are active at their lower and upper bounds, respectively. This way, the active inequality constrained outputs can be denoted collectively as the vector $\mathbf{y}^a$, with the corresponding set points $\mathbf{y}^{aS}$, and the mapping $\mathbf{H}^a(\mathbf{u},\mathbf{d})$:

$$\mathbf{y}^a = \begin{bmatrix} P^{yL}\mathbf{y}^{in} \\ P^{yU}\mathbf{y}^{in} \end{bmatrix} \in \mathbf{R}^{n_y^a}, \quad \mathbf{y}^{aS} = \begin{bmatrix} P^{yL}\mathbf{y}^{L} \\ P^{yU}\mathbf{y}^{U} \end{bmatrix} \in \mathbf{R}^{n_y^a} \tag{5}$$

$$\mathbf{H}^a(\mathbf{u}, \mathbf{d}) = \begin{bmatrix} P^{yL}\mathbf{H}^{in}(\mathbf{u}, \mathbf{d}) \\ P^{yU}\mathbf{H}^{in}(\mathbf{u}, \mathbf{d}) \end{bmatrix} \in \mathbf{R}^{n_y^a} \tag{6}$$

with $n_y^a = m^L + m^U$. In turn, the inactive inequality constrained outputs and the corresponding mapping are denoted as the vectors $\mathbf{y}^{na} = \mathbf{y}^{in}\backslash\mathbf{y}^a$ and $\mathbf{H}^{na} = \mathbf{H}^{in}\backslash\mathbf{H}^a$, respectively. The active inputs can be denoted collectively as the vector $\mathbf{u}^a$, with the corresponding set points $\mathbf{u}^{aS}$:

$$\mathbf{u}^a = \begin{bmatrix} P^{L}\mathbf{u} \\ P^{U}\mathbf{u} \end{bmatrix} \in \mathbf{R}^{n_u^a}, \quad \mathbf{u}^{aS} = \begin{bmatrix} P^{L}\mathbf{u}^{L} \\ P^{U}\mathbf{u}^{U} \end{bmatrix} \in \mathbf{R}^{n_u^a} \tag{7}$$

with $n_u^a = n^L + n^U$. In turn, the inactive inputs can be denoted as the vector $\mathbf{u}^{na} = \mathbf{u}\backslash\mathbf{u}^a$. Also, the equality and active inequality constrained variables can be denoted collectively as the vector $\mathbf{z}$, with the corresponding set points $\mathbf{z}^S$, and the mapping $\mathbf{Z}(\mathbf{u},\mathbf{d})$:

$$\mathbf{z} = \begin{bmatrix} \mathbf{y}^{eq} \\ \mathbf{y}^{a} \\ \mathbf{u}^{a} \end{bmatrix} \in \mathbf{R}^{n_z}, \quad \mathbf{z}^S = \begin{bmatrix} \mathbf{y}^{S} \\ \mathbf{y}^{aS} \\ \mathbf{u}^{aS} \end{bmatrix} \in \mathbf{R}^{n_z} \tag{8}$$

$$\mathbf{Z}(\mathbf{u}, \mathbf{d}) = \begin{bmatrix} \mathbf{H}^{eq}(\mathbf{u}, \mathbf{d}) \\ \mathbf{H}^{a}(\mathbf{u}, \mathbf{d}) \\ \mathbf{u}^{a} \end{bmatrix} \in \mathbf{R}^{n_z} \tag{9}$$

with $n_z = n_y^{eq} + n_y^a + n_u^a$.

By taking the active inequality constraints as equality constraints and removing the inactive constraints, it is possible to reformulate problem 4 as follows:

$$\mathbf{u}^\star(\mathbf{d}) = \arg\min_{\mathbf{u}} \ \Phi(\mathbf{u}, \mathbf{d}) \quad \text{s.t.} \ \mathbf{z} = \mathbf{Z}(\mathbf{u}, \mathbf{d}) = \mathbf{z}^S \tag{10}$$

The first-order NCO of problem 10 read (the notation $\mathbf{a_b} = \partial\mathbf{a}/\partial\mathbf{b}$ is used henceforth):

$$\mathcal{L}_{\mathbf{u}} = \Phi_{\mathbf{u}} + \boldsymbol{\mu}^T\mathbf{Z}_{\mathbf{u}} = \mathbf{0} \tag{11}$$

$$\mathcal{L}_{\boldsymbol{\mu}} = (\mathbf{z} - \mathbf{z}^S)^T = \mathbf{0} \tag{12}$$

with $\mathcal{L}(\mathbf{u}, \mathbf{d}, \boldsymbol{\mu}) = \Phi(\mathbf{u}, \mathbf{d}) + \boldsymbol{\mu}^T(\mathbf{Z}(\mathbf{u}, \mathbf{d}) - \mathbf{z}^S)$ being the Lagrangian function, and $\boldsymbol{\mu} \in \mathbf{R}^{n_z}$ the Lagrange multipliers.

Let $\mathbf{d}_{nom}$ be the nominal parameter values, and consider the parametric disturbance $\delta\mathbf{d} = \mathbf{d} - \mathbf{d}_{nom}$. The deviation of the optimal inputs induced by $\delta\mathbf{d}$ is $\delta\mathbf{u}^\star = \mathbf{u}^\star(\mathbf{d}) - \mathbf{u}^\star(\mathbf{d}_{nom})$, and the corresponding deviations of the Lagrange multipliers is $\delta\boldsymbol{\mu}^\star$. The first-order variations of the NCO eqs 11, 12 read:

$$\mathcal{L}_{\mathbf{uu}}\delta\mathbf{u} + \mathcal{L}_{\mathbf{ud}}\delta\mathbf{d} + \mathbf{Z}_{\mathbf{u}}^T\delta\boldsymbol{\mu} = \mathbf{0} \tag{13}$$

$$\mathbf{Z}_{\mathbf{u}}\delta\mathbf{u} + \mathbf{Z}_{\mathbf{d}}\delta\mathbf{d} = \mathbf{0} \tag{14}$$

where $\delta\mathbf{u}$ and $\delta\boldsymbol{\mu}$ are first-order approximations of $\delta\mathbf{u}^\star$, and $\delta\boldsymbol{\mu}^\star$, respectively, and $\mathcal{L}_{\mathbf{uu}}$, $\mathcal{L}_{\mathbf{ud}}$, $\mathbf{Z}_{\mathbf{u}}$, and $\mathbf{Z}_{\mathbf{d}}$ are all evaluated at $(\mathbf{u}^\star(\mathbf{d}_{nom}), \mathbf{d}_{nom})$. Equations 13 and 14 can be written as

$$M\begin{bmatrix} \delta\mathbf{u} \\ \delta\boldsymbol{\mu} \end{bmatrix} = -\begin{bmatrix} \mathcal{L}_{\mathbf{ud}} \\ \mathbf{Z}_{\mathbf{d}} \end{bmatrix}\delta\mathbf{d}, \quad \text{with } M = \begin{bmatrix} \mathcal{L}_{\mathbf{uu}} & \mathbf{Z}_{\mathbf{u}}^T \\ \mathbf{Z}_{\mathbf{u}} & \mathbf{0} \end{bmatrix} \tag{15}$$

Assuming that the linear independence constraint qualification (LICQ) holds at $\mathbf{u}^\star(\mathbf{d}_{nom})$, then $\mathbf{Z}_{\mathbf{u}}$ has full row rank, and if $n_z < n_u$ one can find $\mathbf{N} \in \mathbf{R}^{n_u \times (n_u - n_z)}$ such that the columns of $\mathbf{N}$ are an orthonormal basis of the null space of the rows of $\mathbf{Z}_{\mathbf{u}}$, that is, $\mathbf{Z}_{\mathbf{u}}\mathbf{N} = \mathbf{0}$. If in addition the reduced Hessian matrix $\mathbf{N}^T\mathcal{L}_{\mathbf{uu}}\mathbf{N}$ is positive definite, then the KKT matrix M is nonsingular,[21] and there is a unique vector pair $(\delta\mathbf{u},\delta\boldsymbol{\mu})$ satisfying

$$\begin{bmatrix} \delta\mathbf{u} \\ \delta\boldsymbol{\mu} \end{bmatrix} = -M^{-1}\begin{bmatrix} \mathcal{L}_{\mathbf{ud}} \\ \mathbf{Z}_{\mathbf{d}} \end{bmatrix}\delta\mathbf{d}, \quad \text{with } M^{-1} = \begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} \tag{16}$$

Hence, the (approximate) optimal inputs are a linear function of the disturbances:

$$\delta\mathbf{u} = K\delta\mathbf{d}, \quad \text{with } K = -[M_1\mathcal{L}_{\mathbf{ud}} + M_2\mathbf{Z}_{\mathbf{d}}] \tag{17}$$

3. SELF-OPTIMIZING CONTROL USING THE NULL SPACE METHOD

In self-optimizing control the active input and output variables are controlled to their optimal boundary values, and for the remaining degrees of freedom, $n_c = (n_u - n_z)$ controlled variables $\mathbf{c} \in \mathbf{R}^{n_c}$ are selected as functions of the inactive input and output variables, such that near optimal operation is achieved by keeping $\mathbf{c}$ at constant set points $\mathbf{c}^S$.[8,12] The null space method by Alstad and Skogestad[17] is based on the first-order variation of the NCO. In this section, the null space method for constrained optimization problems is presented. The first-order variation of the output variables is

$$\delta\mathbf{y} = \mathbf{H}_u\delta\mathbf{u} + \mathbf{H}_d\delta\mathbf{d} \tag{18}$$

Using eq 17 in eq 18, we have

$$\delta\mathbf{y} = [\mathbf{H}_uK + \mathbf{H}_d]\delta\mathbf{d} \tag{19}$$

From eqs 19 and 17, we have

$$\begin{bmatrix} \delta\mathbf{y} \\ \delta\mathbf{u} \end{bmatrix} = S\delta\mathbf{d}, \quad \text{with } S = \begin{bmatrix} \mathbf{H}_uK + \mathbf{H}_d \\ K \end{bmatrix} \tag{20}$$

Next, let us analyze some characteristics of matrix S. For the case of the equality and active inequality constrained variables we have

$$\delta\mathbf{z} = \mathbf{Z}_u\delta\mathbf{u} + \mathbf{Z}_d\delta\mathbf{d} \tag{21}$$

Using eq 14 in eq 21, we have $\boldsymbol{\delta}\mathbf{z} = \mathbf{0}$. Since this is valid for all $\boldsymbol{\delta}\mathbf{d}$, it must be that the rows in K corresponding to the active input variables are equal to $\mathbf{0}$, and the rows in $[\mathbf{H_u}K + \mathbf{H_d}]$ corresponding to the equality constrained outputs and the active inequality constrained outputs are equal to $\mathbf{0}$.

Removing from eq 20 the equality constrained outputs and all the active inequality constrained inputs and outputs, one can write

$$\begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{\text{na}} \\ \boldsymbol{\delta}\mathbf{u}^{\text{na}} \end{bmatrix} = S^{\text{na}}\boldsymbol{\delta}\mathbf{d} \tag{22}$$

where $S^{\text{na}} \in \mathbb{R}^{n^{\text{na}} \times n_d}$, with $n^{\text{na}} = (n_y^{\text{na}} + n_u^{\text{na}})$, is obtained by eliminating from S all the rows corresponding to the equality constrained outputs and active variables.

Let us assume that S^{na} is full column rank, and let the columns in $\mathcal{N}^{\text{na}} \in \mathbb{R}^{n^{\text{na}} \times n_c}$ be a set of n_c orthonormal vectors that lie in the left null space of S^{na}. Hence, $(\mathcal{N}^{\text{na}})^{\text{T}}S^{\text{na}} = \mathbf{0}$, and from eq 22 we have

$$(\mathcal{N}^{\text{na}})^T \begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{\text{na}} \\ \boldsymbol{\delta}\mathbf{u}^{\text{na}} \end{bmatrix} = (\mathcal{N}^{\text{na}})^{\text{T}}S^{\text{na}}\boldsymbol{\delta}\mathbf{d} = \mathbf{0} \tag{23}$$

Notice that, the dimension of the null space of S^{na} should be greater than or equal to n_c, which requires that $n^{\text{na}} - n_d \geq n_c$. This last inequality reduces to $n_y^{\text{eq}} + n_y^{\text{in}} \geq n_d$, or equivalently, $n_y \geq n_d$. That is, the number of measured output variables should be greater than or equal to the number of disturbances (regardless of whether these outputs are controlled at fixed set points or not).

Based on eq 23, the null space method—initially proposed by Alstad and Skogestad[17]—consists in selecting the controlled variables and set point values as follows:

Controlled variables	Set points
$\mathbf{z}$	$\mathbf{z}^{\text{S}}$
$\mathbf{c} = (\mathcal{N}^{\text{na}})^T \begin{bmatrix} \mathbf{y}^{\text{na}} \\ \mathbf{u}^{\text{na}} \end{bmatrix}$	$\mathbf{c}^{\text{S}} = (\mathcal{N}^{\text{na}})^{\text{T}} \begin{bmatrix} \mathbf{H}^{\text{na}}(\mathbf{u}^{\star}(\mathbf{d}_{\text{nom}}), \mathbf{d}_{\text{nom}}) \\ \mathbf{u}^{\text{na}\star}(\mathbf{d}_{\text{nom}}) \end{bmatrix}$

(24)

where $\mathbf{u}^{\text{na}\star}(\mathbf{d}_{\text{nom}})$ includes the optimal values in $\mathbf{u}^{\star}(\mathbf{d}_{\text{nom}})$ corresponding to the inactive inputs. Notice that eq 24 represents a square control problem with n_u controlled variables and n_u input variables. Since the active input variables can be fixed to their boundary values, the number of controlled variables is actually $(n_y^{\text{eq}} + n_y^{\text{a}} + n_c) = n_u^{\text{na}}$. The SOC structure corresponding to the control problem 24 is depicted in Figure 1. The active input variables $\mathbf{u}^{\text{a}}$ are applied directly to the plant, and the inactive inputs $\mathbf{u}^{\mathbf{na}}$ are decoupled into the inputs $\mathbf{u}' \in \mathbb{R}^{(n_y^{\text{eq}}+n_y^{\text{a}})}$ and $\mathbf{u}'' \in \mathbb{R}^{n_c}$. The inputs $\mathbf{u}'$ are used to control the equality and active inequality constrained outputs by means of the *constraint controller* $\mathcal{K}^{\text{c}}$. Meanwhile, the inputs $\mathbf{u}''$ are used to control the SOC variables $\mathbf{c}$ by means of the *sensitivity controller* $\mathcal{K}^{\text{s}}$. The dashed-line box in Figure 1 represents the plant as viewed by the sensitivity controller. This is the same controlled plant as viewed by the SOC controller proposed by Alstad and Skogestad.[17]

Remark 1 The formulation of the null space method given in this section presents some differences with respect to the original formulation presented in the literature.[17,18] In the original formulation, the inactive output and input variables are considered collectively as output measurements, while here they are considered separately in equation 22, the same as in Francois et al.[1] Another difference is that in previous work[1,17,18] it is assumed that a number of degrees of freedom (inputs) are consumed in controlling the active constraints (for instance, the inputs $\mathbf{u}'$ and $\mathbf{u}^{\text{a}}$ in Figure 1), and the optimization problem is formulated as an unconstrained optimization problem in terms of the remaining inputs (i.e., in terms of the inputs $\mathbf{u}''$ in Figure 1). In contrast, in this paper, we consider the constrained optimization problem in terms of all the inputs $\mathbf{u}$, and the null space method is derived based on the first-order variations of the NCO for the constrained optimization problem. In principle, the approach is the same. However, the alternative formulation used here is justified because it is key in deriving the new SOC structures that will be presented in section 4.

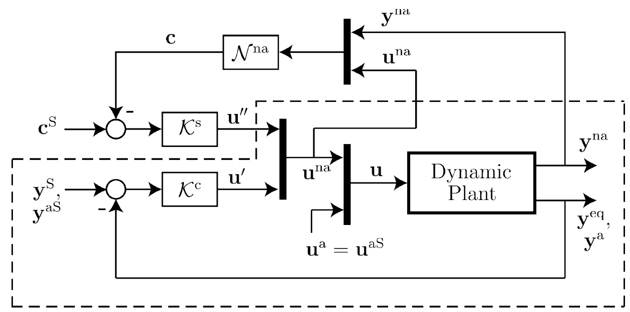

Figure 1. Self-optimizing control structure, including the constraint controller $\mathcal{K}^{\text{c}}$ and the sensitivity controller $\mathcal{K}^{\text{s}}$.

Remark 2 (Condition $n_y \geq n_d$) Notice that the condition $n_y \geq n_d$ found in this paper is similar to the condition found in Alstad and Skogestad.[17] In Alstad and Skogestad,[17] it was found that the number of independent measurement variables should be greater than or equal to the number of extra degrees of freedom plus the number of disturbances. The apparent difference in this condition is due to differences in the notation used in Alstad and Skogestad[17] with respect to the notation used in this paper. The measurement variables considered in Alstad and Skogestad[17] include the inactive output and input variables. Hence, using the notation of this paper, the condition given in Alstad and Skogestad[17] can be written as $n_u^{\text{na}} + n_y^{\text{na}} \geq n_c + n_d$. Recalling that $n_c = n_u^{\text{na}} - n_y^{\text{eq}} - n_y^{\text{a}}$, the condition reduces to $n_y \geq n_d$.

Equivalent Formulation. An equivalent way of arriving to the control problem 24 is by selecting a particular matrix $\mathcal{N} \in \mathbb{R}^{(n_u+n_y) \times n_u}$, such that the columns of $\mathcal{N}$ are a set of n_u normal vectors that lie in the left null space of S (defined in eq 20), that is, $\mathcal{N}^{\text{T}}S = \mathbf{0}$. These column vectors are linearly independent (by construction) but they are not all orthogonal. They are selected as follows: The first n_z columns in $\mathcal{N}$ correspond to an orthonormal set of unit vectors. Each unit vector has a one at the index number corresponding to a variable in $\mathbf{z}$, and zeros elsewhere. The remaining n_c columns in $\mathcal{N}$ are constructed from the columns in $\mathcal{N}^{\text{na}}$ by adding zeros in the positions corresponding to the variables in $\mathbf{z}$. Using this choice of $\mathcal{N}$ it is possible to arrive to the control problem 24 by selecting the controlled variables and set points as follows:

Controlled variables Set points

$$\mathbf{r} = \mathcal{N}^{\mathrm{T}}\begin{bmatrix}\mathbf{y}\\ \mathbf{u}\end{bmatrix} \qquad \mathbf{r}^{\mathrm{S}} = \mathcal{N}^{\mathrm{T}}\begin{bmatrix}\mathbf{H}(\mathbf{u}^{\star}(\mathbf{d}_{\mathrm{nom}}),\ \mathbf{d}_{\mathrm{nom}})\\ \mathbf{u}^{\star}(\mathbf{d}_{\mathrm{nom}})\end{bmatrix} \tag{25}$$

This equivalent formulation of the control problem shows that the selection of the active constraints as controlled variables can also be viewed as being based on a null space approach. Notice that, the dimension of the null space of S should be greater than or equal to n_u. Assuming that S is full column rank, this requires that $n_y \geq n_d$. This is the same condition found previously. The following example shows how to construct $\mathcal{N}^{\mathrm{na}}$ and $\mathcal{N}$ from S.

Example 1 Consider the case in which there are four inputs, two outputs, and two disturbances, with the second output and the second input being active at the nominal optimum. Let us assume that the matrix S is given by

$$\mathrm{S} = \begin{bmatrix}1 & 2\\ 0 & 0\\ 2 & -1\\ 0 & 0\\ 3 & 1\\ 2 & 4\end{bmatrix}, \quad \mathrm{S}^{\mathrm{na}} = \begin{bmatrix}1 & 2\\ 2 & -1\\ 3 & 1\\ 2 & 4\end{bmatrix}$$

where the rows of S corresponding to the active output and input variables are equal to **0**. These rows are eliminated in order to obtain the matrix S^{na}. Next, the null space matrix $\mathcal{N}^{\mathrm{na}}$ can be obtained, for example, by singular value decomposition of $(\mathrm{S}^{\mathrm{na}})^{\mathrm{T}}$.

$$\mathcal{N}^{\mathrm{na}} \approx \begin{bmatrix}-0.5050 & -0.7504\\ -0.6092 & 0.2888\\ 0.6092 & -0.2888\\ -0.05208 & 0.5196\end{bmatrix},$$

$$\mathcal{N} \approx \begin{bmatrix}0 & 0 & -0.5050 & -0.7504\\ 1 & 0 & 0 & 0\\ 0 & 0 & -0.6092 & 0.2888\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0.6092 & -0.2888\\ 0 & 0 & -0.05208 & 0.5196\end{bmatrix}$$

Finally, the null space matrix $\mathcal{N}$ is obtained from $\mathcal{N}^{\mathrm{na}}$ and the knowledge of the active constraints.

First-Order Approximation of Optimal Operation. The null space method is based on the first-order variation of the NCO, and as discussed by Marchetti and Zumoffen,[19] it provides a first-order approximation of optimal operation. The equation $\mathcal{N}^{T}S = \mathbf{0}$ can be rewritten as

$$\mathcal{N}^{T}\begin{bmatrix}\mathbf{H}_{\mathbf{u}}\\ \mathbf{I}\end{bmatrix}\mathrm{K} = \mathcal{N}^{T}\begin{bmatrix}-\mathbf{H}_{\mathrm{d}}\\ \mathbf{0}\end{bmatrix} \tag{26}$$

Now, if we assume that eq 18 is exact, and that the variables **r** are controlled to their set point values $\mathbf{r}^{\mathrm{S}}$ in eq 25 with integral control action, then at steady state we will have

$$\delta\mathbf{r} = \mathcal{N}^{T}\begin{bmatrix}\delta\mathbf{y}\\ \delta\mathbf{u}\end{bmatrix} = \mathcal{N}^{T}\begin{bmatrix}\mathbf{H}_{\mathrm{u}}\delta\mathbf{u} + \mathbf{H}_{\mathrm{d}}\delta\mathbf{d}\\ \delta\mathbf{u}\end{bmatrix} = \mathbf{0} \tag{27}$$

from where

$$\mathcal{N}^{\mathrm{T}}\begin{bmatrix}\mathbf{H}_{\mathrm{u}}\\ \mathbf{I}\end{bmatrix}\delta\mathbf{u} = \mathcal{N}^{\mathrm{T}}\begin{bmatrix}-\mathbf{H}_{\mathrm{d}}\\ \mathbf{0}\end{bmatrix}\delta\mathbf{d} \tag{28}$$

Using eq 26 in eq 28, we have

$$\mathcal{N}^{T}\begin{bmatrix}\mathbf{H}_{\mathrm{u}}\\ \mathbf{I}\end{bmatrix}\delta\mathbf{u} = \mathcal{N}^{T}\begin{bmatrix}\mathbf{H}_{\mathrm{u}}\\ \mathbf{I}\end{bmatrix}\mathrm{K}\delta\mathbf{d} \tag{29}$$

The linear system (eq 29) has the unique solution $\delta\mathbf{u} = \mathbf{K}\delta\mathbf{d}$ if and only if the square matrix $\mathcal{N}^{\mathrm{T}}[\mathbf{H}_{\mathbf{u}}^{\mathrm{T}}\ \mathbf{I}]^{\mathrm{T}}$ is nonsingular. Notice that the rank of $\mathcal{N}$ is n_u, and that the rank of $[\mathbf{H}_{\mathbf{u}}^{\mathrm{T}}\ \mathbf{I}]$ is also n_u, which are necessary conditions for $\mathcal{N}^{\mathrm{T}}[\mathbf{H}_{\mathbf{u}}^{\mathrm{T}}\ \mathbf{I}]^{\mathrm{T}}$ to be nonsingular. The operating point $\delta\mathbf{u} = \mathbf{K}\delta\mathbf{d}$ is a first-order approximation of optimal operation. In practice, since eq 18 is not exact, it is not possible to write the second equality in 27, and a (slightly) different operating point will be reached at steady-state.

4. SELF-OPTIMIZING CONTROL STRUCTURES WITH MINIMUM NUMBER OF CONTROLLED VARIABLES

In the SOC structures proposed in the literature, the number of additional self-optimizing controlled variables is typically equal to the number of degrees of freedom that are available after controlling all the equality and active inequality constrained output variables (i.e., equal to n_c, with $n_c = n_u^{\mathrm{na}} - n_y^{\mathrm{eq}} - n_y^{\mathrm{a}}$). In this section, we show that if the number of disturbances is lower than the number of inactive inputs, then it is possible to decrease the number of self-optimizing controlled variables by fixing linear combinations of the input variables. In this paper, the following distinction is made between controlled variables and fixed variables:

Controlled Variables. We refer to process-dependent controlled variables, that is, variables that depend on the outputs **y**, which are functions of the states and disturbances as per eq 2b. A controller with integral action is in general required in order to meet the set points with zero offset.

Fixed Variables. We refer to input variables and combinations of input variables that can be fixed at their target values at the input of the process; that is, they do not depend on the process dynamics.

Three different scenarios will be analyzed depending on the dimensions of the problem considered.

4.1. Scenario 1: $n_d \leq (n_y^{\mathrm{eq}} + n_y^{\mathrm{a}}) < n_u^{\mathrm{na}}$. If the number of equality constrained outputs plus the number of active inequality constrained outputs is greater than or equal to the number of disturbances, then the condition $n_y \geq n_d$ is already satisfied by controlling $\mathbf{y}^{\mathrm{eq}}$ and $\mathbf{y}^{\mathrm{a}}$. Therefore, it is not necessary to include any inactive output variables in eq 22, which reduces to

$$\delta\mathbf{u}^{\mathrm{na}} = \mathrm{K}^{\mathrm{na}}\delta\mathbf{d} \tag{30}$$

where $\mathrm{K}^{\mathrm{na}} \in \mathbf{R}^{n_u^{na} \times n_d}$ is obtained by eliminating from K all the rows corresponding to the active input variables. The columns in $\mathcal{N} \in \mathbf{R}^{n_u^{na} \times n_c}$ orthonormal vectors that lie in the left null space of $\mathbf{K}^{\mathrm{na}}$, so that $(\mathcal{N}^{\mathrm{na}})^{\mathrm{T}}\mathbf{K}^{\mathrm{na}} = \mathbf{0}$. In this case, the variables **c** are linear combinations of the input variables only. Therefore, they can be fixed to their optimal values $\mathbf{c}^{\mathrm{S}}$, and the only variables that need to be controlled are $\mathbf{y}^{\mathrm{eq}}$ and $\mathbf{y}^{\mathrm{a}}$. The controlled variables and their set points, as well as the fixed variables and their targets, are the following:

Controlled variables	Set points
$\mathbf{y}^{eq}$	$\mathbf{y}^{S}$
$\mathbf{y}^{a}$	$\mathbf{y}^{aS}$
Fixed variables	Targets
$\mathbf{u}^{a}$	$\mathbf{u}^{aS}$
$\mathbf{c} = (\mathcal{N}^{na})^T \mathbf{u}^{na}$	$\mathbf{c}^S = (\mathcal{N}^{na})^T \mathbf{u}^{na\star}(\mathbf{d}_{nom})$ (31)

Next, we shall specify the manipulated variables used to control $\mathbf{y}^{eq}$ and $\mathbf{y}^{a}$. The fixed variables can be obtained collectively from the inputs $\mathbf{u}$ as

$$\begin{bmatrix} \mathbf{u}^a \\ \mathbf{c} \end{bmatrix} = (\mathcal{N}^u)^T \mathbf{u} \tag{32}$$

where the first n_u^a columns in $\mathcal{N}^u \in \mathbf{R}^{n_u \times (n_u^a + n_c)}$ are the unit vectors that have a one at the index number corresponding an active input variable, while the remaining n_c columns in $\mathcal{N}^u$ are obtained from the columns in $\mathcal{N}^{na}$ by adding zeros in the positions corresponding to the active input variables. Let the columns in matrix $\mathcal{R} \in \mathbf{R}^{n_u \times (n_y^{eq} + n_y^a)}$ be a set of orthonormal vectors that lie in the left null space of $\mathcal{N}^u$, so that $\mathcal{R}^T\mathcal{N}^u = \mathbf{0}$. Notice that the dimension of the null space of $\mathcal{N}^u$ is $(n_u - n_u^a - n_c) = (n_z - n_u^a) = (n_y^{eq} + n_y^a)$. Using $\mathcal{R}$, the manipulated variables used to control $\mathbf{y}^{eq}$ and $\mathbf{y}^a$ can be selected as

$$\mathbf{v} = \mathcal{R}^T \mathbf{u} \tag{33}$$

which are linear combinations of the inactive input variables (notice that the columns in $\mathcal{R}$ corresponding to the active input variables are zero columns). Equations 32 and 33 can be written jointly as

$$\begin{bmatrix} \mathbf{u}^a \\ \mathbf{c} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} (\mathcal{N}^u)^T \\ \mathcal{R}^T \end{bmatrix} \mathbf{u} \tag{34}$$

Noticing that the square matrix $[\mathcal{N}^u \mathcal{R}]$ is invertible (since it has linearly independent columns), it follows that the input $\mathbf{u}$ can be uniquely reconstructed from $\mathbf{v}$ and from the targets $\mathbf{u}^{aS}$ and $\mathbf{c}^S$ as follows:

$$\mathbf{u} = \begin{bmatrix} (\mathcal{N}^u)^T \\ \mathcal{R}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}^{aS} \\ \mathbf{c}^S \\ \mathbf{v} \end{bmatrix} = [Q_1 \;\; Q_2] \begin{bmatrix} \mathbf{u}^{aS} \\ \mathbf{c}^S \\ \mathbf{v} \end{bmatrix} = Q_1 \begin{bmatrix} \mathbf{u}^{aS} \\ \mathbf{c}^S \end{bmatrix} + Q_2 \mathbf{v} \tag{35}$$

The resulting self-optimizing control structure is depicted in Figure 2. Notice that only the constraint controller is required. The dashed-line box in Figure 2 represents the plant as viewed by the constraint controller.

If the number of disturbances is lower than or equal to the number of controlled output variables (equality constrained outputs and active inequality constrained outputs), then it is not necessary to include additional controlled variables in order to obtain a self-optimizing control structure. Instead, it is possible to use linear combinations of the input variables (i.e., directions in the input space) in order to control the equality and active inequality constrained outputs, while the input variables for the remaining degrees of freedom are fixed at their optimal values along computed directions in the input space, which are locally invariant to disturbances. In this scenario, self-optimizing control is achieved with only $(n_y^{eq} + n_y^a)$ controlled variables. No additional inactive output variables (or combinations of output variables) need to be measured online and controlled.

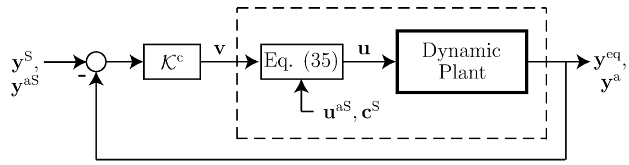

Figure 2. Self-optimizing control structure with minimum number of controlled variables when scenario 1 holds.

4.2. Scenario 2: $(n_y^{eq} + n_y^a) < n_d < n_u^{na}$.

If the number of disturbances is greater than the number of equality and active inequality constrained outputs, but smaller than the number of inactive inputs, then the minimum number of inactive output variables $\mathbf{y}^{na}$ required to satisfy $n_y \geq n_d$ is $(n_y^{na})_{min} = n_d - n_y^{eq} - n_y^a$. On the other hand, the dimension of the left null space of K^{na} is $(n_u^{na} - n_d)$, which means that it is possible to fix $(n_u^{na} - n_d)$ linear combinations of the inputs. Let the columns in $\mathcal{N}^{una} \in \mathbf{R}^{n_u^{na} \times (n_u^{na} - n_d)}$ be taken as a set of $(n_u^{na} - n_d)$ orthonormal vectors that lie in the left null space of K^{na}, so that $(\mathcal{N}^{una})^T K^{na} = \mathbf{0}$. Using $\mathcal{N}^{una}$, the linear combinations of inputs $\mathbf{c}^u = (\mathcal{N}^{una})^T \mathbf{u}^{na}$ can be fixed at the locally invariant optimal values $\mathbf{c}^{uS} = (\mathcal{N}^{una})^T \mathbf{u}^{na\star}(\mathbf{d}_{nom})$.

In this case, $\mathcal{N}^{na} \in \mathbf{R}^{n^{na} \times n_c}$ can be selected in the left null space of $S^{na} \in R^{n^{na} \times n_d}$ as follows:

$$\mathcal{N}^{na} = \left[\mathcal{N}^{yna} \middle| \begin{matrix} \mathbf{0} \\ \mathcal{N}^{una} \end{matrix} \right] \tag{36}$$

where $\mathcal{N}^{yna}$ includes the first $n_d - n_y^{eq} - n_y^a$ columns in $\mathcal{N}^{na}$, which can be chosen to be orthonormal. The whole set of column vectors in $\mathcal{N}^{na}$ should be linearly independent. Using $\mathcal{N}^{una}$ and $\mathcal{N}^{yna}$, the controlled variables and their set points, as well as the fixed variables and their targets, can be selected as follows:

Controlled variables	Set points
$\mathbf{y}^{eq}$	$\mathbf{y}^{S}$
$\mathbf{y}^{a}$	$\mathbf{y}^{aS}$
$\mathbf{c} = (\mathcal{N}^{yna})^T \begin{bmatrix} \mathbf{y}^{na} \\ \mathbf{u}^{na} \end{bmatrix}$	$\mathbf{c}^S = (\mathcal{N}^{yna})^T \begin{bmatrix} \mathbf{H}^{na}(\mathbf{u}^\star(\mathbf{d}_{nom}), \mathbf{d}_{nom}) \\ \mathbf{u}^{na\star}(\mathbf{d}_{nom}) \end{bmatrix}$
Fixed variables	Targets
$\mathbf{u}^{a}$	$\mathbf{u}^{aS}$
$\mathbf{c}^u = (\mathcal{N}^{una})^T \mathbf{u}^{na}$	$\mathbf{c}^{uS} = (\mathcal{N}^{una})^T \mathbf{u}^{na\star}(\mathbf{d}_{nom})$ (37)

Notice that the controlled variables $\mathbf{c}$ and the fixed variables $\mathbf{c}^u$ can be written jointly as

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{c}^u \end{bmatrix} = (\mathcal{N}^{na})^T \begin{bmatrix} \mathbf{y}^{na} \\ \mathbf{u}^{na} \end{bmatrix} \tag{38}$$

On the other hand, the fixed variables can be obtained collectively from the inputs $\mathbf{u}$ as

$$\begin{bmatrix} \mathbf{u}^{\mathrm{a}} \\ \mathbf{c}^{\mathrm{u}} \end{bmatrix} = (\mathcal{N}^{\mathrm{u}})^{\mathrm{T}}\mathbf{u} \tag{39}$$

where the first n_u^{a} columns in $\mathcal{N}^u \in \mathbf{R}^{n_u \times (n_u - n_d)}$ are the unit vectors that have a one at the index number corresponding an active input variable, while the remaining $(n_u^{\mathrm{na}} - n_d)$ columns in $\mathcal{N}^u$ are obtained from the columns in $\mathcal{N}^{\mathrm{una}}$ by adding zeros in the positions corresponding to the active input variables. Let the columns in matrix $\mathcal{R} \in \mathbf{R}^{n_u \times n_d}$ be a set of orthonormal vectors that lie in the left null space of $\mathcal{N}^{\mathrm{u}}$, so that $\mathcal{R}^{\mathrm{T}}\mathcal{N}^{\mathrm{u}} = \mathbf{0}$. Using $\mathcal{R}$, the manipulated variables used to control $\mathbf{y}^{\mathrm{eq}}$, $\mathbf{y}^{\mathrm{a}}$, and $\mathbf{c}$ can be selected as

$$\mathbf{v} = \mathcal{R}^{\mathrm{T}}\mathbf{u} \tag{40}$$

which are linear combinations of the inactive input variables. Equations 39 and 40 can be written jointly as

$$\begin{bmatrix} \mathbf{u}^{\mathrm{a}} \\ \mathbf{c}^{\mathrm{u}} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} (\mathcal{N}^{\mathrm{u}})^{\mathrm{T}} \\ \mathcal{R}^{\mathrm{T}} \end{bmatrix}\mathbf{u} \tag{41}$$

The square matrix $[\mathcal{N}^{\mathrm{u}}\ \mathcal{R}]$ being nonsingular, it follows that the input $\mathbf{u}$ can be uniquely reconstructed from $\mathbf{v}$ and from the targets $\mathbf{u}^{\mathrm{aS}}$ and $\mathbf{c}^{\mathrm{uS}}$ as follows:

$$\mathbf{u} = \begin{bmatrix} (\mathcal{N}^{\mathrm{u}})^{T} \\ \mathcal{R}^{T} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{u}^{\mathrm{aS}} \\ \mathbf{c}^{\mathrm{uS}} \\ \mathbf{v} \end{bmatrix} = [\mathrm{Q}_1 \quad \mathrm{Q}_2] \begin{bmatrix} \mathbf{u}^{\mathrm{aS}} \\ \mathbf{c}^{\mathrm{uS}} \\ \mathbf{v} \end{bmatrix} = \mathrm{Q}_1 \begin{bmatrix} \mathbf{u}^{\mathrm{aS}} \\ \mathbf{c}^{\mathrm{uS}} \end{bmatrix} + \mathrm{Q}_2\mathbf{v} \tag{42}$$

The manipulated variables $\mathbf{v}$ can be decoupled into the variables $\mathbf{v}' \in \mathbf{R}^{(n_y^{\mathrm{eq}} + n_y^{\mathrm{a}})}$, which are used to control the equality and active inequality constrained outputs, and the variables $\mathbf{v}'' \in \mathbf{R}^{(n_d - n_y^{\mathrm{eq}} - n_y^{\mathrm{a}})}$, which are used to control the SOC variables $\mathbf{c}$. The proposed self-optimizing control structure with minimum number of controlled variables in this scenario is depicted in Figure 3. Self-optimizing control is achieved with a minimum of n_d controlled variables.

4.3. Scenario 3: $(n_y^{\mathrm{eq}} + n_y^{\mathrm{a}}) < n_u^{\mathrm{na}} \le n_d$. If the number of disturbances is greater than or equal to the number of inactive input variables, then the left null space of K^{na} is empty, which means that it is not possible to fix any linear combinations of the inputs. In order to satisfy the condition $n_y \ge n_d$, the minimum number of inactive output variables required is $n_d - n_y^{\mathrm{eq}} - n_y^{\mathrm{a}}$. In this case, the controlled variables and set points required to obtain a self-optimizing control structure are those given in eq 24 (the active input variables are fixed). The number of additional controlled variables required is n_c. The self-optimizing control structure for this scenario is that depicted in Figure 1. Self-optimizing control is achieved with $n_u^{\mathrm{na}} = (n_y^{\mathrm{eq}} - n_y^{\mathrm{a}} + n_c)$ controlled variables, and it is not possible to reduce the number of controlled variables by fixing linear combinations of the input variables.

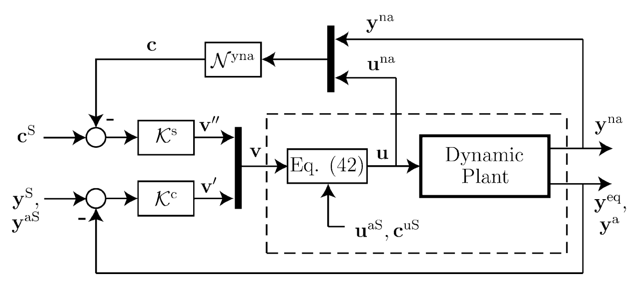

Figure 3. Self-optimizing control structure with minimum number of controlled variables when scenario 2 holds.

4.4. Effect of Measurement Error on the Optimality Loss. The NCO of problem 4 can be decomposed into two parts: the sensitivity part (eq 11) and the constraint part (eq 12). The sensitivity part is equivalent to zeroing the reduced gradient of the cost function. A parametric variational analysis[22] shows that—in the presence of parametric (or disturbance) variations of magnitude $\delta\mathbf{d}$—failure to enforce the active constraints would lead to a loss of optimality $O(\delta\mathbf{d})$. In contrast, upon enforcing the active constraints, the loss of optimality would be $O(\delta\mathbf{d}^2)$, which is due to failure in zeroing the reduced gradient. In other words, when disturbances take place, there is often much more to win (locally) in terms of optimality by controlling the active constrained variables to their optimal boundary values than by zeroing the reduced gradient. This justifies why controlling the active constrained quantities[4] has become widely adopted in basically all optimizing control approaches.[5−7,10,17,23] Where these approaches differ is on how the sensitivity part of the NCO is dealt with. Different sensitivity control strategies have been proposed, such as gradient control,[24] neighboring-extremal control (NEC),[23] extremum-seeking control,[25] and self-optimizing control approaches, such as the null space method,[17] which is revisited in this paper.

When considering the effect of measurement error, it is important to keep in mind that, in terms of optimality loss, measurement error is in general more detrimental to the constraint part of the NCO than to the sensitivity part of the NCO. This follows as a result of the previously mentioned parametric variational analysis.[22] In addition, measurement error in the active constrained variables may result in violation of the constraints, which calls for implementing constraint backoffs,[26] with the consequent loss in optimality.

The effect of measurement error and disturbances on the performance of the sensitivity controller should be analyzed in order to determine whether it is justified or not to include the sensitivity controller. In Gros et al.,[23] the cost improvement of using the NEC scheme is compared to using the nominal inputs on the perturbed plant (the active constraints are assumed to be perfectly controlled). If this cost improvement is too small, or if it is comparable to the cost variation due to measurement noise, then it is not justified to apply NEC.[23] A similar type of analysis should be carried out for all sensitivity control strategies, including SOC approaches such as the null space method.

In this section, a local sensitivity analysis is conducted in order to analyze the effect of disturbances and measurement error on the optimality loss incurred by SOC. This analysis is analogous to the sensitivity analysis conducted in previous works.[12,18] A difference with previous work is that here the analysis is conducted for the constrained optimization problem that is formulated in the space of all the input (manipulated) variables. Most importantly, in the following analysis, a clear distinction is made between the effect of measurement error in the input and output variables. This discrimination will enable us to arrive to important conclusions concerning the SOC scenarios discussed in this paper. Initially, the analysis will be carried out for the general SOC problem given in eq 25. The special cases concerning the SOC structures with minimum number of controlled variables for scenarios 1 and 2 will be discussed later. Instead of using the null space combination matrix $\mathcal{N}^{\mathrm{na}}$, the general combination matrix Q^{na} will be used, which will enable us to establish links with other SOC approaches different from the null space method. In order to incorporate the active constraints, the matrix Q is constructed from Q^{na} in the same way $\mathcal{N}$ is constructed from $\mathcal{N}^{\mathrm{na}}$.

We shall assume that the KKT second-order sufficient conditions for a strict local minimum of problem 4 are satisfied (see e.g., Theorem 4.4.2 in Bazaraa et al.[27]). Since we are performing a local analysis, we shall consider that the first-order variations of the NCO (eqs 13 and 14) and of the outputs (eq 18) are exact. The null space method is concerned with the sensitivity part of the NCO. Hence, we will focus only on the optimality loss due to failure in zeroing the reduced gradient; that is, under the assumption that the active constraints are perfectly controlled.

Optimality Loss due to Sensitivities. A second-order Taylor expansion of the Lagrangian function at the optimum point $\mathbf{u}^\star(\mathbf{d})$ gives

$$\mathcal{L}(\mathbf{u}, \mathbf{d}, \boldsymbol{\mu}^\star(\mathbf{d})) = \mathcal{L}(\mathbf{u}^\star(\mathbf{d}), \mathbf{d}, \boldsymbol{\mu}^\star(\mathbf{d})) + \mathcal{L}_{\mathbf{u}}(\mathbf{u} - \mathbf{u}^\star(\mathbf{d})) + \frac{1}{2}(\mathbf{u} - \mathbf{u}^\star(\mathbf{d}))^T \mathcal{L}_{\mathrm{uu}}(\mathbf{u} - \mathbf{u}^\star(\mathbf{d})) \tag{43}$$

Since $\mathbf{u}^\star(\mathbf{d})$ is an optimum point, we have $\mathcal{L}_{\mathbf{u}} = \mathbf{0}$ and $\mathbf{Z}(\mathbf{u}^\star(\mathbf{d}),\mathbf{d}) = \mathbf{z}^{\mathrm{S}}$. Assuming that the active constraints are perfectly controlled at $\mathbf{u}$, we have $\mathbf{Z}(\mathbf{u},\mathbf{d}) = \mathbf{z}^{\mathrm{S}}$. Therefore,

$$\mathcal{L}(\mathbf{u}, \mathbf{d}, \boldsymbol{\mu}^\star(\mathbf{d})) = \Phi(\mathbf{u}, \mathbf{d})$$

$$\mathcal{L}(\mathbf{u}^\star(\mathbf{d}), \mathbf{d}, \boldsymbol{\mu}^\star(\mathbf{d})) = \Phi(\mathbf{u}^\star(\mathbf{d}), \mathbf{d})$$

From 43 it follows that the optimality loss is given by

$$\mathrm{Loss} = \Phi(\mathbf{u}, \mathbf{d}) - \Phi(\mathbf{u}^\star(\mathbf{d}), \mathbf{d}) = \frac{1}{2}(\mathbf{u} - \mathbf{u}^\star(\mathbf{d}))^T \mathcal{L}_{\mathrm{uu}}(\mathbf{u} - \mathbf{u}^\star(\mathbf{d})) \tag{44}$$

In this analysis, the constraints are assumed to be linear, and the effect of the curvature of the constraints on the cost is taken into account by using the Hessian of the Lagrangian function. Since the active constraints are assumed to be linear, we have that $\mathbf{u} - \mathbf{u}^\star(\mathbf{d})$ belongs to the tangent space $\mathcal{T} = \{\mathbf{w} \in \mathbf{R}^{n_u}: \mathbf{Z}_{\mathbf{u}}w = \mathbf{0}\}$. The Hessian of the Lagrangian function is not in general positive definite, but it is positive definite when restricted to the tangent space.[27]

Input Reached by the SOC Controller. Let $\mathbf{n}^{\mathbf{u}} \in \mathbf{R}^{n_u}$ and $\mathbf{n}^{\mathrm{y}} \in \mathbf{R}^{n_y}$ denote the input and output measurement error vectors, respectively. The input reached by the SOC controller in the presence of disturbances and measurement error (assuming the controller has integral action) can be derived by adding measurement error to eq 27:

$$Q^{\mathrm{T}}\begin{bmatrix} \mathbf{H}_{\mathbf{u}}(\boldsymbol{\delta}\mathbf{u} + \tilde{\mathbf{n}}^{\mathrm{u}}) + \mathbf{H}_{\mathrm{d}}\boldsymbol{\delta}\mathbf{d} + \mathbf{n}^{\mathrm{y}} \\ \boldsymbol{\delta}\mathbf{u} + \tilde{\mathbf{n}}^{\mathrm{u}} \end{bmatrix} = \mathbf{0} \tag{45}$$

which can be written as

$$Q^{\mathrm{T}}\begin{bmatrix} \mathbf{H}_{\mathrm{u}} \\ \mathbf{I} \end{bmatrix}(\boldsymbol{\delta}\mathbf{u} + \tilde{\mathbf{n}}^{\mathrm{u}}) + Q^{\mathrm{T}}\begin{bmatrix} \mathbf{n}^{\mathrm{y}} \\ \mathbf{0} \end{bmatrix} + Q^{\mathrm{T}}\begin{bmatrix} \mathbf{H}_{\mathrm{d}} \\ \mathbf{0} \end{bmatrix}\boldsymbol{\delta}\mathbf{d} = \mathbf{0} \tag{46}$$

Since we are only interested in the effect of measurement error on the sensitivity part of the NCO, the projection of the input measurement error on the tangent space is included in 46:

$$\tilde{\mathbf{n}}^{\mathrm{u}} = \mathrm{BB}^{\mathrm{T}}\mathbf{n}^{\mathrm{u}}$$

where the columns of the matrix $\mathrm{B} \in \mathbf{R}^{n_u \times (n_u - n_z)}$ are an orthonormal basis that spans the tangent space $\mathcal{T}$ (thus, $\mathrm{Z_u B} = \mathbf{0}$). In addition, the output measurement error corresponding to the equality and active inequality constrain output variables is assumed to be equal to zero. That is, the elements in $\mathbf{n}^{\mathbf{y}}$ corresponding to these controlled output variables are equal to zero. Because of this, and due to the structure of Q, it follows that the first n_z equations in 46 reduce to

$$\mathbf{Z}_{\mathrm{u}}(\boldsymbol{\delta}\mathbf{u} + \tilde{\mathbf{n}}^{\mathrm{u}}) + \mathbf{Z}_{\mathbf{d}}\boldsymbol{\delta}\mathbf{d} = \mathbf{0}$$

which means that the active constraints are active at $(\mathbf{u} + \tilde{\mathbf{n}}^{\mathbf{u}})$. Given that $\tilde{\mathbf{n}}^{\mathbf{u}} \in \mathcal{T}$, it follows that the active constraints are active at $\mathbf{u}$. Therefore, the input $\mathbf{u}$ computed from eq 46 can be used in eq 44. Considering that

$$\boldsymbol{\delta}\mathbf{u}^\star = \mathbf{u}^\star(\mathbf{d}) - \mathbf{u}^\star(\mathbf{d}_{\mathrm{nom}}) = \mathrm{K}\boldsymbol{\delta}\mathbf{d}$$

and recalling that SOC requires that the matrix $Q^{\mathrm{T}}[\mathbf{H}_{\mathrm{u}}^{\mathrm{T}}\ \mathbf{I}]^{\mathrm{T}}$ be invertible, we have

$$\mathbf{u} - \mathbf{u}^\star(\mathbf{d}) = \boldsymbol{\delta}\mathbf{u} - \boldsymbol{\delta}\mathbf{u}^\star = -\mathrm{M}^{\mathrm{y}}\begin{bmatrix} \mathbf{n}^{\mathrm{y}} \\ \mathbf{0} \end{bmatrix} - \tilde{\mathbf{n}}^{\mathrm{u}} - \mathrm{M}^{\mathrm{d}}\boldsymbol{\delta}\mathbf{d} \tag{47}$$

with

$$\mathrm{M}^{\mathrm{y}} = \left(Q^{\mathrm{T}}\begin{bmatrix} \mathbf{H}_{\mathrm{u}} \\ \mathbf{I} \end{bmatrix}\right)^{-1} Q^{\mathrm{T}},$$

$$\mathrm{M}^{\mathrm{d}} = \left(\left(Q^{\mathrm{T}}\begin{bmatrix} \mathbf{H}_{\mathrm{u}} \\ \mathbf{I} \end{bmatrix}\right)^{-1} Q^{\mathrm{T}}\begin{bmatrix} \mathbf{H}_{\mathrm{d}} \\ \mathbf{0} \end{bmatrix} + \mathrm{K}\right)$$

Loss in Terms of the Disturbances and Measurement Error. In the presence of parametric disturbance $\boldsymbol{\delta}\mathbf{d}$ and zero-mean measurement noise $\mathbf{n}^{\mathbf{u}}$, $\mathbf{n}^{\mathbf{y}}$, the (local) optimality loss associated with the SOC controller that uses the combination matrix Q^{na} is given by

$$E(\mathrm{Loss}) = \frac{1}{2}E\left(\begin{bmatrix} \mathbf{n}^{\mathrm{y}} \\ \mathbf{0} \end{bmatrix}^{\mathrm{T}} (\mathrm{M}^{\mathrm{y}})^{\mathrm{T}} \mathcal{L}_{\mathrm{uu}} \mathrm{M}^{\mathrm{y}} \begin{bmatrix} \mathbf{n}^{\mathrm{y}} \\ \mathbf{0} \end{bmatrix}\right) + \frac{1}{2}E((\tilde{\mathbf{n}}^{\mathrm{u}})^T \mathcal{L}_{\mathrm{uu}} \tilde{\mathbf{n}}^{\mathrm{u}}) + \frac{1}{2}\boldsymbol{\delta}\mathbf{d}^{\mathrm{T}}(\mathrm{M}^{\mathrm{d}})^{\mathrm{T}} \mathcal{L}_{\mathrm{uu}} \mathrm{M}^{\mathrm{d}} \boldsymbol{\delta}\mathbf{d} \tag{48}$$

where $E(\cdot)$ represents the mathematical expectation operator. The first and second terms in eq 48 represent the expected loss associated with the output and input measurement noise, respectively. The third term represents the loss associated with the disturbances. Notice that the output noise loss and the disturbance loss depend on the choice of the controlled variables through the combination matrix Q^{na}, while the loss due to input noise is independent of Q^{na}.

If the null space method is used, then $Q = \mathcal{N}$, and from eq 26, it follows that $\mathbf{M}^{\mathbf{d}} = \mathbf{0}$. This is exactly the same result obtained in Alstad et al.,[18] where it is shown that by using the null space method the (local) loss due to disturbances is equal to zero. Yet, the null space method does not in general minimize the loss eqs 44 or 48 in the presence of measurement error. In order to take measurement error into account, alternative SOC approaches have been proposed, which find the combination matrix Q^{na} that minimizes the worst-case loss,[12] and the average loss.[13]

Loss for the SOC Structures with Minimum Number of Controlled Variables. Scenario 1. If $n_d \leq (n_y^{eq} + n_y^{a}) < n_u^{na}$, then it is possible to select a null space matrix $\mathcal{N}^{na}$ that fixes n_c linear combinations of the input variables. Since no inactive output variables are needed, the expected value of the loss in eq 48 reduces to

$$E(\text{Loss}) = \frac{1}{2}E((\tilde{\mathbf{n}}^{u})^{T}\mathcal{L}_{uu}\tilde{\mathbf{n}}^{u})$$

Notice that, in this scenario it is not possible to improve this loss by any combination matrix Q^{na} that incorporates inactive output variables and(or) does not belong to the null space of S^{na}, as this would make the output noise loss and(or) the disturbance loss greater than zero.

Scenario 2. If $(n_y^{eq} + n_y^{a}) < n_d < n_u^{na}$, then it is possible to select the inactive output variables $\mathbf{y}^{na}$ and the null space matrices $\mathcal{N}^{yna}$ and $\mathcal{N}^{una}$ in eq 36 so as to minimize the optimality loss due to output measurement error, that is, the first term in eq 48. This problem falls outside the scope of this paper.

4.5. Fixing Linear Combinations of the Input Variables Using Alternative SOC Approaches. The SOC strategies based on the worst-case loss[12] and the average loss[13] will not in general result in fixing linear combinations of the input variables, even if the optimality loss in doing so might be negligible. However, using these methods, it is possible to fix linear combinations of the input variables by imposing certain structure to the combination matrix Q^{na}. Since in this case Q^{na} is not restricted to be a null space matrix, we may (in principle) fix from one to n_c linear combinations of the inputs, regardless of the number of disturbances:

Fixing n_c Linear Combinations of the Inputs. If $(n_y^{eq} + n_y^{a}) \geq 1$, it is possible to fix n_c linear combinations of the inputs, $\mathbf{c} = (Q^{na})^{T}\mathbf{u}^{na}$, by selecting $Q^{na} \in \mathbf{R}^{n_u^{na} \times n_c}$ such that the second and third terms in eq 48 are minimized. In this case, the SOC control structure in Figure 2 can be used. Notice that, if $(n_y^{eq} + n_y^{a}) = 0$ it makes no sense to fix $n_c = n_u^{na}$ linear combinations of the inputs, as this would be equivalent to applying the nominal optimal inputs directly to the plant.

Fixing Less than n_c Linear Combinations of the Inputs. It is possible to fix n_u^f, $1 \leq n_u^f \leq (n_c - 1)$ linear combinations of the inputs by selecting the controlled variables $\mathbf{c}$ and the fixed variables $\mathbf{c}^{u}$ as

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{c}^{u} \end{bmatrix} = (Q^{na})^{T}\begin{bmatrix} \mathbf{y}^{na} \\ \mathbf{u}^{na} \end{bmatrix} \tag{49}$$

where Q^{na} has the following structure:

$$Q^{na} = \left[\begin{array}{c|c} Q^{yna} & \mathbf{0} \\ & Q^{una} \end{array}\right] \tag{50}$$

The inactive output variables $\mathbf{y}^{na}$, as well as the matrices $Q^{yna} \in \mathbf{R}^{n^{na} \times (n_c - n_u^f)}$ and $Q^{una} \in \mathbf{R}^{n_u^{na} \times n_u^f}$ can be selected so as to minimize the loss expression in 48. In this case, the SOC control structure depicted in Figure 3 can be implemented.

The analysis of such alternative SOC structures with reduced number of process-dependent controlled variables may be a subject of future research.

4.6. Discussion. A number of remarks are in order:

- **Minimum number of controlled variables.** The self-optimizing control structures proposed use the minimum number of controlled variables without compromising local optimality, in the sense that the optimum input $\boldsymbol{\delta}\mathbf{u} = \mathbf{K}\boldsymbol{\delta}\mathbf{d}$ is reached (locally) at steady state. Nevertheless, it might be possible to find self-optimizing control structures with acceptable optimality loss by fixing additional input combinations that are not invariant to disturbances, but that do not result in significant optimality loss. The possibility of further reducing the number of controlled variables, with the acceptance of the additional optimality loss, may be a subject of future research.
- Notice that, it is always possible to use the self-optimizing control structure in Figure 1 in the case of scenarios 1 and 2. This corresponds to the original SOC structure proposed by Alstad and Skogestad.[17]
- The null space approach allows for important flexibility in selecting the null space matrix $\mathcal{N}$. In this paper, we have taken advantage of this flexibility in order to select in the columns of $\mathcal{N}$ as many linearly independent combinations of the inputs as possible, with the aim of reducing the number of required controlled variables by fixing these input combinations to their invariant optimal values.
- In the special case where there are no equality or active inequality constrained outputs (i.e., $n_y^{eq} = n_y^{a} = 0$), the following statements hold: (i) scenario 1 cannot take place; (ii) if $n_d < n_u^{na}$, then scenario 2 takes place, but without including the constraint controller $\mathcal{K}^{c}$, and with $\mathbf{v} = \mathbf{v}''$; (iii) if $n_u^{na} \leq n_d$, then scenario 3 takes place, without including the constraint controller $\mathcal{K}^{c}$, and with $\mathbf{u}^{na} = \mathbf{u}''$.

5. ILLUSTRATIVE EXAMPLES

5.1. SELF-OPTIMIZING CONTROL OF A CONTINUOUS STIRRED TANK REACTOR

The reactor in the Williams–Otto plant is considered.[28] It consists of an ideal CSTR in which the following reactions occur:

$$A + B \rightarrow C, \quad k_1 = 1.6599 \times 10^{6}\ e^{-6666.7/(T_R+273)}$$

$$B + C \rightarrow P + E, \quad k_2 = 7.2117 \times 10^{8}\ e^{-8333.3/(T_R+273)}$$

$$C + P \rightarrow G, \quad k_3 = 2.6745 \times 10^{12}\ e^{-11111/(T_R+273)}$$

where the reactants A and B are fed with the mass flow rates F_A and F_B, respectively. The desired products are P and E. C is an intermediate product and G is an undesired product. The reaction rates are

$$r_1 = k_1 X_A X_B, \quad r_2 = k_2 X_B X_C, \quad r_3 = k_3 X_C X_P$$

where X_i is the mass fraction of species i, and k_j is the kinetic coefficient of reaction j, which is dependent on the reactor temperature. Operation is isothermal at the temperature T_R.

 dx.doi.org/10.1021/ie5010509 | Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

Assuming that the level control is perfect (i.e., $F = F_A + F_B$), the dynamic behavior of the reactor is described by the following set of differential equations:

$$\dot{X}_A = \frac{F_A}{W} - \frac{(F_A + F_B)}{W}X_A - r_1 \quad (51)$$

$$\dot{X}_B = \frac{F_B}{W} - \frac{(F_A + F_B)}{W}X_B - r_1 - r_2 \quad (52)$$

$$\dot{X}_C = -\frac{(F_A + F_B)}{W}X_C + 2r_1 - 2r_2 - r_3 \quad (53)$$

$$\dot{X}_P = -\frac{(F_A + F_B)}{W}X_P + r_2 - \frac{1}{2}r_3 \quad (54)$$

$$\dot{X}_G = -\frac{(F_A + F_B)}{W}X_G + \frac{3}{2}r_3 \quad (55)$$

$$\dot{X}_E = -\frac{(F_A + F_B)}{W}X_E + 2r_2 \quad (56)$$

$$WC_p\dot{T}_R = F_BC_pT_{inB} + F_AC_pT_{inA} - \Delta H_1Wr_1 - \Delta H_2Wr_2 - \Delta H_3Wr_3 - \frac{W}{1000}\frac{A_o}{V_o}U(T_R - T_J) - (F_A + F_B)C_pT_R \quad (57)$$

$$W_JC_{pJ}\dot{T}_J = F_JC_{pJ}(T_{Jin} - T_J) + \frac{W}{1000}\frac{A_o}{V_o}U(T_R - T_J) \quad (58)$$

where 51-56 are the species mass balances; 57 is the heat balance in the reactor, and 58 is the heat balance in the jacket. A PI controller regulates the reactor temperature by adjusting the jacket inlet temperature T_{Jin}.

Variables and parameters: T_{inA},T_{inB} = inlet temperatures of A and B, T_J = jacket temperature, T_{Jin} = inlet temperature of the fluid entering the jacket, F_J = flow rate of the fluid entering the jacket, W = reactor mass holdup, W_J = jacket mass holdup, C_p = heat capacity of the reactants, C_{pJ} = heat capacity of the fluid in the jacket, ΔH_i = enthalpy of reaction i, U = heat transfer coefficient, A_o/V_o = specific heat exchange area.

The decision variables are the set point of the reactor temperature, T_R^{sp}, and F_B, that is, $\mathbf{u} = [T_R^{sp}\ F_B]^T$. The disturbance variable considered is F_A (i.e., $d = F_A$). The set of eqs 51−56 is $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},d)$, with $\mathbf{x} = [X_A\ X_B\ X_C\ X_P\ X_G\ X_E]^T$. The nominal model uses F_A = 1.4 kg/s. Since the temperature controller is designed with zero offset, we have $T_R^{sp} = T_R$ at steady state. Hence, the steady-state model is given by $\mathbf{0} = \mathbf{f}(\mathbf{x},\mathbf{u},d)$.

The objective is to maximize profit at steady state operation, which is expressed as the price difference between the products and the reactants:

Table 1. Model Variables and Parameters

variable	value	unit	variable	value	unit
T_{inA}	60	°C	T_{inB}	60	°C
F_A	1.4	kg/s	F_J	3	kg/s
W	2105	kg	W_J	200	kg
C_p	4.184	kJ/(kg °C)	C_{pJ}	4.184	kJ/(kg °C)
ΔH_1	−263.8	kJ/kg	ΔH_2	−158.3	kJ/kg
ΔH_3	−226.3	kJ/kg	A_o	12.2	m^2
V_o	2.1052	m^3	U	0.72	kJ/(m^2 °C s)

$$\Phi = 1200X_PF + 80X_EF - 76F_A - 114F_B \quad (59)$$

The steady-state optimization problem reads:

$$\begin{aligned} \max_{F_B, T_R^{sp}} \quad & \Phi \\ \text{s.t.} \quad & \mathbf{f}(\mathbf{x}, \mathbf{u}, F_A) = \mathbf{0} \\ & X_B \le 0.25 \\ & 2 \le F_B \le 4,\ 70 \le T_R \le 100, \end{aligned} \quad (60)$$

where Φ is defined in eq 59 and X_B is constrained to be lower than or equal to 0.25. This optimization problem formulation is the same as that used in Marchetti et al.[29]

The constraint $X_B \le 0.25$ is the only active constraint at the optimum for the nominal value of F_A, and also for any $F_A \in [1.069, 2.322]$. Therefore, assuming that the concentration X_B can be measured online, a constraint controller should be implemented for controlling X_B. Notice that, in this example we have $n_u^{na} = 2$, $n_y^{a} = 1$, and $n_d = 1$, which corresponds to scenario 1. Hence, we can use the self-optimizing control structure depicted in Figure 2.

Let us denote by $U^\star(F_A)$ the optimal solution map of problem 60 as a function of the disturbance value F_A. The nominal optimum is $\bar{\mathbf{u}}^\star = U^\star(1.4) = [85.6, 2.549]^T$. Figure 4

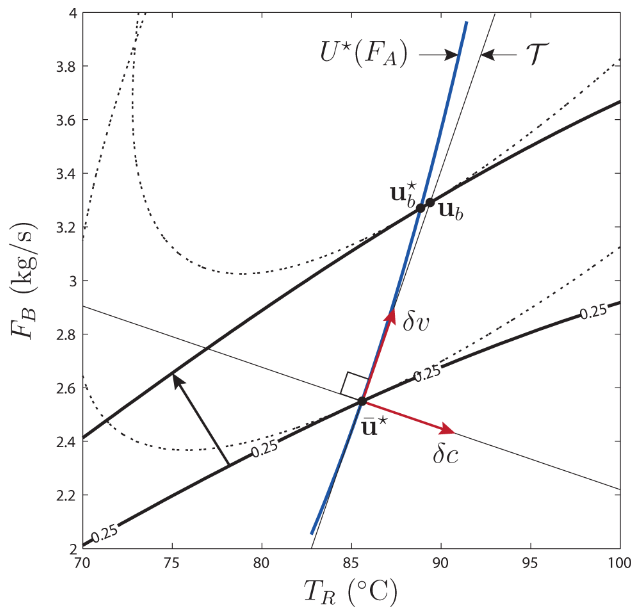

Figure 4. Steady-state map for problem 60. Thick solid curves: constraint boundary for X_B. Dotted curves: contours of the profit function.

shows the location of $\bar{\mathbf{u}}^\star$ on the constraint boundary $X_B = 0.25$ for the nominal value of F_A. The disturbance considered is an increase in the value of F_A from 1.4 kg/s to 1.85 kg/s. As a result, the constraint boundary shifts as indicated in Figure 4, and the new optimum is $\mathbf{u}_b^\star = U^\star(1.85)$. Notice that the optimal solutions are always located on the curve $U^\star(F_A)$ for any $F_A \in [1.069,2.322]$.

In order to apply the null space approach, the inputs are scaled as $(F_B)_{\text{scaled}} = (F_B - 2)/2$ and $(T_R^{sp})_{\text{scaled}} = (T_R^{sp} - 70)/30$. The vector of scaled inputs is denoted by $\mathbf{u}_s$, and the scaled nominal optimum is denoted by $\bar{\mathbf{u}}_s^\star$. The gain matrix in eq 17 is computed at the nominal point as K = $[8.377,\ 1.631]^T$. The corresponding gain matrix for the scaled inputs is K_s = [0.2792,

$0.8154]^T$. Since there are no active input variables, we have $\mathbf{u}^{na} = \mathbf{u}$, and the null space of K_s is $\mathcal{N}^{na} = \mathcal{N}^{u} = [-0.9461, 0.3239]^T$. Therefore, the combination of input variables $c = (\mathcal{N}^{na})^T \mathbf{u}_s$ can be fixed at the target value $c^S = (\mathcal{N}^{na})^T \bar{\mathbf{u}}_s^{\star} = -0.40224$. In this example, the control problem 31 reads:

Controlled variable	Set point	Manipulated variable
$y^a = X_B$	$y^{aS} = 0.25$	$v = \mathcal{R}^T \mathbf{u}_s$
Fixed variable	Target	
$c = (\mathcal{N}^{na})^T \mathbf{u}_s$	$c^S = -0.40224$	

The manipulated variable used to control X_B is $v = \mathcal{R}^T \mathbf{u}_s$, with $\mathcal{R} = [0.3239, 0.9461]^T$. This way, eq 35 becomes $\mathbf{u}_s = Q_1 c^S + Q_2 v$, with $Q_1 = [-0.9461, 0.3239]^T$, and $Q_2 = [0.3239, 0.9461]^T$.

When F_A increases from 1.4 kg/s to 1.85 kg/s, the self-optimizing controller will take the operation to point $\mathbf{u}_b$ in Figure 4, which is at the intersection of the line $\mathcal{T}$ and the constraint boundary $X_B = 0.25$. Here, $\mathcal{T}$ is the affine subspace for which $\delta c = c - c^S = 0$. Notice that $\mathcal{T}$ is tangent to $U^{\star}(F_A)$ at $\bar{\mathbf{u}}^{\star}$. The optimality loss of point $\mathbf{u}_b$ with respect to the optimum $\mathbf{u}_b^{\star}$ is negligible. The simulation results obtained using a PI controller with parameters $K_{pi} = 5$ and $\tau_{pi} = 1$ (min) are shown in Figure 5. The disturbance in F_A takes place at time $t = 50$

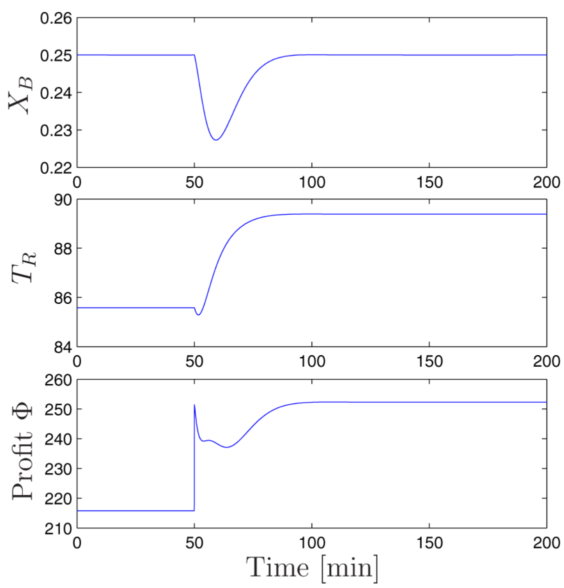

Figure 5. Time response of key variables in the self-optimizing control of the CSTR example.

(min). The self-optimizing constraint controller brings X_B back to the active value of 0.25, and reaches at steady state exactly the operating point $\mathbf{u}_b$ in Figure 4.

Notice that, if in this optimization scenario a linear combination of the inputs is not fixed, then the null space method would require measuring at least one inactive output variable (e.g., a state variable different from X_B). In this case, the control structure depicted in Figure 1 should be implemented, requiring two (process-dependent) controlled variables instead of one.

5.2. Self-Optimizing Control of an Evaporator. We consider the forced-circulation evaporator described by Newell and Lee,[30] as modified by Kariwala et al.[13] This evaporator example has been used to illustrate the performance of self-optimizing control schemes in numerous studies.[13,15,31] It has also been used to illustrate the implementation of model predictive control.[32] The liquid feed is mixed with recirculating liquor, which is pumped through the evaporator. The evaporator is a heat exchanger, which is heated by steam. The mixture of feed and recirculating liquor boils inside the evaporator, and a vapor–liquid mixture flows to the separator, where the liquid and vapor are separated. Most of the separated liquid, which is more concentrated than when it entered the evaporator, becomes recirculating liquor, and a small proportion of it is drawn off as product. The evaporator is depicted in Figure 6 and the main variables are listed in Table

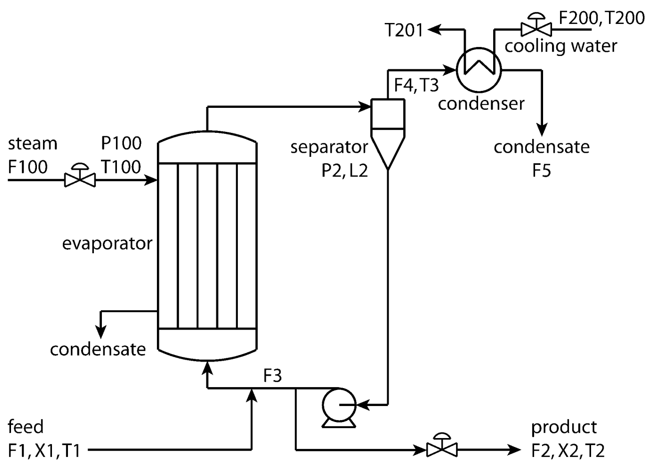

Figure 6. Forced-circulation evaporator.

2. The process model equations are given in Kariwala et al.[13] The process model has three state variables, L_2, X_2, and P_2. In order to stabilize the plant operation, a controller is needed for the separator level L_2. A PI controller is used with a set point for L_2 of 1 m, which manipulates the product flow rate F_2.

The objective is to maximize profit at steady-state operation, which is expressed as[13]

Table 2. Evaporator Variables and Values at the Nominal Optimum

variable	description	value
F_1	feed flow rate	10.155 kg/min
F_2	product flow rate	1.430 kg/min
F_3	circulating flow rate	23.545 kg/min
F_4	vapor flow rate	8.725 kg/min
F_5	condensate flow rate	8.725 kg/min
X_1	feed comp.	5.00%
X_2	product comp.	35.50%
T_1	feed temp.	38 °C
T_2	product temp.	83.26 °C
T_3	vapor temp.	76.42 °C
L_2	separator level	1 m
P_2	operating pressure	42.25 kPa
F_{100}	steam flow rate	10.06 kg/min
T_{100}	steam temp.	151.52 °C
P_{100}	steam pressure	400.00 kPa
Q_{100}	heat duty	368.07 kW
F_{200}	cooling water flow rate	232.63 kg/min
T_{200}	cooling water inlet temp.	17 °C
T_{201}	cooling water outlet temp.	37.63 °C
Q_{200}	condenser duty	335.90 kW

$$\Phi = 4800F_2 - 0.2F_1 - 600F_{100} - 0.6F_{200} - 1.009(F_2 + F_3) \quad (61)$$

where the first term is the product value, while the last four terms are operational costs related to the steam and cooling water utilities, pumping effort, and raw material cost. The decision variables, disturbances, and measured output variables are

$$\mathbf{u} = [P_{100} \quad F_{200} \quad F_3 \quad F_1]^{\mathbf{T}}$$
$$\mathbf{d} = [X_1 \quad T_1 \quad T_{200}]^{\mathbf{T}}$$
$$\mathbf{y} = [X_2 \quad F_4 \quad F_5 \quad T_2 \quad T_3 \quad P_2 \quad F_{100} \quad T_{201}]^{\mathbf{T}}$$

The nominal disturbance values are $X_1 = 5\%$, $T_1 = 38$ °C, and $T_{200} = 17$ °C, that is, $\mathbf{d}_{\text{nom}} = [5 \quad 38 \quad 17]^{\text{T}}$, which is different from the nominal disturbance values used in Kariwala et al.[13]

The steady-state optimization problem reads:

$$\begin{aligned} \max_{\mathbf{u}} \quad & \Phi \\ \text{s.t.} \quad & X_2 \geq 35.5\% \\ & 35 \text{ kPa} \leq P_2 \leq 80 \text{ kPa} \\ & P_{100} \leq 400 \text{ kPa} \\ & 0 \text{ kg/min} \leq F_{200} \leq 400 \text{ kg/min} \\ & 0 \text{ kg/min} \leq F_1 \leq 20 \text{ kg/min} \\ & 0 \text{ kg/min} \leq F_3 \leq 100 \text{ kg/min} \end{aligned} \quad (62)$$

where Φ is defined in eq 61. Let us denote by $U^\star(\mathbf{d})$ the optimal solution map of problem 62 as a function of the disturbance values $\mathbf{d}$. The nominal optimum is $\bar{\mathbf{u}}^\star = U^\star(\mathbf{d}_{\text{nom}}) = [400, 232.62, 23.545, 10.155]^{\text{T}}$. The values of all the process variables at the nominal optimum are given in Table 2. At the optimal point, there is one active input, $P_{100} = 400$ kPa, and one active inequality constrained output, $X_2 = 35.5\%$. In this work, it is assumed that the disturbance values can vary within the following ranges:

$$\begin{aligned} 4.95\% &\leq X_1 \leq 5.1\% \\ 36°\text{C} &\leq T_1 \leq 40°\text{C} \\ 15°\text{C} &\leq T_{200} \leq 20°\text{C} \end{aligned} \quad (63)$$

The following remarks are in order:

- The degrees of freedom and constraints are similar to those adopted in Kariwala et al.[13] The evaporator pressure P_2 is very sensitive to the value of the disturbance X_1. In industrial practice it would be advisible to control P_2, which would consume an additional degree of freedom. However, since the purpose here is to illustrate the different SOC scenarios, we choose to leave P_2 free, the same as in Kariwala et al.[13]
- The disturbance ranges in eq 63 were selected such that the set of active constraints in problem 62 does not change with the disturbance values. Notice that the ranges are quite tight, as it happens that outside these ranges P_2 quickly becomes active, either at its lower or upper boundary value. These changes in the active set were handled by using a cascade control strategy.[13] However, in this work we want to limit our study to the case where the set of active constraints does not change.

In this problem we have the following description of the variables:

- Active inputs: $u^{\text{a}} = P_{100}$ ($n_u^{\text{a}} = 1$)
- Active outputs: $y^{\text{a}} = X_2$ ($n_y^{\text{a}} = 1$)
- Inactive inputs: $\mathbf{u}^{\text{na}} = [F_{200} \; F_3 \; F_1]^{\text{T}}$ ($n_u^{\text{na}} = 3$)
- Inactive outputs: $\mathbf{y}^{\text{na}} = [F_4 \; F_5 \; T_2 \; T_3 \; P_2 \; F_{100} \; T_{201}]^{\text{T}}$ ($n_y^{\text{na}} = 7$)
- Equality outputs: none ($n_y^{\text{eq}} = 0$)

Hence, depending on whether 1, 2, or 3 disturbances are considered, the problem falls in the framework of scenarios 1, 2, and 3, respectively.

Scenario 1. If the only disturbance is $d = X_1$ ($n_d = 1$), then it is possible to implement the SOC structure described in section 4.1, which is depicted in Figure 2. The matrices involved in the control structure design are the following:

$$\text{K}^{\text{na}} = \begin{bmatrix} 75.9411 \\ 10.2685 \\ 2.4522 \end{bmatrix}, \quad \mathcal{N}^{\text{na}} = \begin{bmatrix} -0.1339 & -0.0320 \\ 0.9910 & -0.0022 \\ -0.0022 & 0.9995 \end{bmatrix}$$

$$\mathcal{N}^{\text{u}} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathcal{N}^{\text{na}} \end{bmatrix}, \quad \mathcal{R} = \begin{bmatrix} 0.0000 \\ 0.9905 \\ 0.1339 \\ 0.0320 \end{bmatrix}$$

In this example, the control problem 31 reads:

Controlled variable	Set point	Manipulated Variable
$y^{\text{a}} = X_2$	$y^{\text{aS}} = 35.5$	$v = \mathcal{R}^T\mathbf{u}$
Fixed variables	Targets	
$u^{\text{a}} = P_{100}$	$u^{\text{aS}} = 400$	
$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = (\mathcal{N}^{\text{na}})^{\text{T}}\mathbf{u}^{\text{na}}$	$\begin{bmatrix} c_1^S \\ c_2^S \end{bmatrix} = \begin{bmatrix} -7.8447 \\ 2.6590 \end{bmatrix}$	

The only controlled variable is the product composition X_2, which is controlled at its lower boundary value of 35.5 using as the manipulated variable the input combination $v = \mathcal{R}^{\text{T}}\mathbf{u}$. In this case, a PI controller is selected and the tuning procedure is the classical internal model control (IMC) approach. The proportional gain and the reset time constant result $K_P = -2$, and $\tau_I = 50$, respectively.

The dynamic response obtained when the composition X_1 changes from 5.0% to 5.1% at $t = 1000$ min is shown in Figure 7. The active constraint $X_2 = 35.5$ is effectively controlled using the manipulated variable v. The optimality loss at steady state ($t = 6000$) is 6.0542. Notice that, in this optimization scenario the original SOC structure would require three controlled variables instead of only one, if no linear combinations of the input variables are fixed.

Scenario 2. If we have $\mathbf{d} = [X_1 \; T_{200}]^{\text{T}}$ ($n_d = 2$), then it is possible to implement the SOC structure described in section 4.2, which is depicted in Figure 3. The minimum number of inactive output variables required is $n_y^{\text{na}} = n_d - n_y^{\text{eq}} - n_y^{\text{a}} = 2 - 0 - 1 = 1$. In this case, we choose $y^{na} = T_{201}$. Next, the matrices involved in the control structure design are computed at the nominal optimum operating point:

dx.doi.org/10.1021/ie5010509 | Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

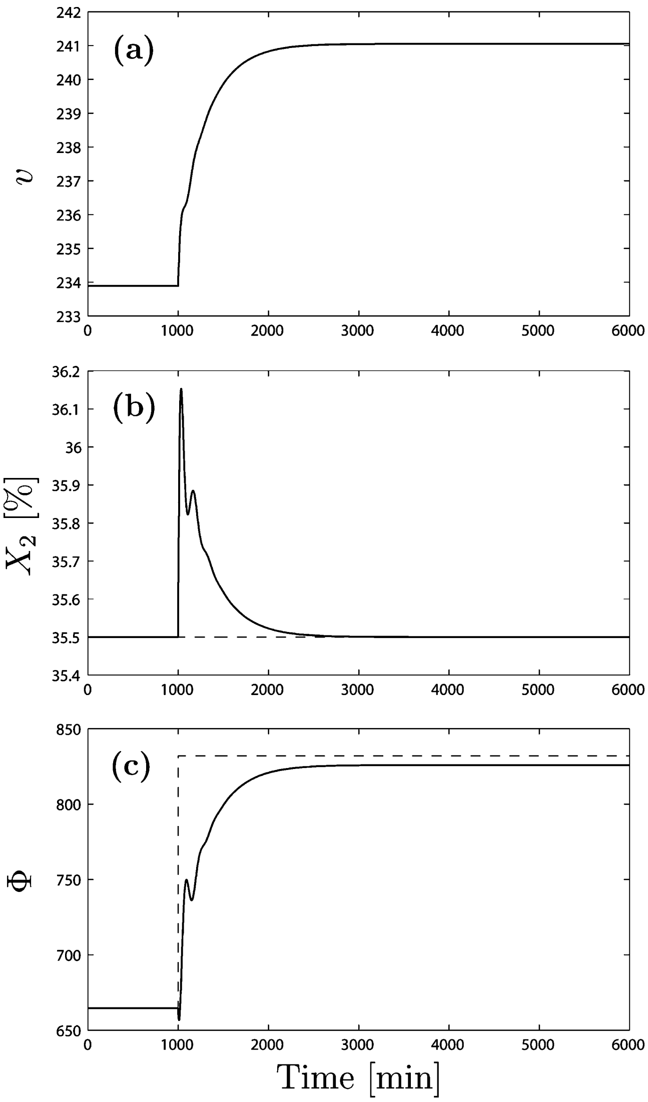

Figure 7. Self-optimizing control of the evaporator example in scenario 1. Positive disturbance in X_1. Plot a: manipulated variable−v. Plot b: controlled variable−X_2. Plot c: Profit−Φ.

$$K^{na} = \begin{bmatrix} 75.9411 & -2.3685 \\ 10.2685 & -0.1449 \\ 2.4522 & -0.1395 \end{bmatrix}, \quad \mathcal{N}^{una} = \begin{bmatrix} -0.0759 \\ 0.3372 \\ 0.9384 \end{bmatrix}$$

$$\mathcal{N}^{u} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathcal{N}^{una} \end{bmatrix}, \quad \mathcal{R} = \begin{bmatrix} 0.0000 & 0.0000 \\ 0.3372 & 0.9384 \\ 0.8943 & -0.2941 \\ -0.2941 & 0.1816 \end{bmatrix}$$

$$S^{yna} = [-2.4292 \quad 0.9266], \quad S^{na} = \begin{bmatrix} S^{yna} \\ K^{na} \end{bmatrix}$$

$$\mathcal{N}^{yna} = \begin{bmatrix} -0.1990 \\ -0.1380 \\ 0.9702 \\ 0.0133 \end{bmatrix}, \quad \mathcal{N}^{na} = \begin{bmatrix} -0.1990 & 0.0000 \\ -0.1380 & -0.0759 \\ 0.9702 & 0.3372 \\ 0.0133 & 0.9384 \end{bmatrix}$$

where S^{yna} corresponds to the output variable T_{201}.

In this example, the control problem 37 reads:

Controlled variables	Set point	Manipulated variables
$y^a = X_2$	$y^{aS} = 35.5$	$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mathcal{R}^T \mathbf{u}$
$c = (\mathcal{N}^{yna})^T \begin{bmatrix} T_{201} \\ \mathbf{u}^{na} \end{bmatrix}$	$c^S = -16.6066$	
Fixed variables	**Targets**	
$u^a = P_{100}$	$u^{aS} = 400$	
$c^u = (\mathcal{N}^{una})^T \mathbf{u}^{na}$	$c^{uS} = -0.1871$	

The controlled variables are $y^a = X_2$ and c, while the manipulated variables are v_1 and v_2. In order to define the input-output pairing between $[\delta y^a \ \delta c]^T$ and $[\delta v_1 \ \delta v_2]^T$ we use the relative gain array (RGA) approach. For this purpose, the matrix P_2 given in A-8 and the RGA matrix $\Lambda = P_2 \otimes (P_2^T)^{-1}$ are computed:

$$P_2 = \begin{bmatrix} 5.4481 & -2.7687 \\ 0.8387 & -0.4168 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} -43.8841 & 44.8841 \\ 44.8841 & -43.8841 \end{bmatrix} \tag{64}$$

The RGA approach suggests the following input-output pairing: $v_1 - c$ and $v_2 - X_2$. The high absolute value of the entries in the matrix Λ is a clear signal of high interaction between both control loops. Both SISO feedback control loops are implemented via PI controllers and tuned using the classical IMC approach. This procedure gives the following setting: $K_p^1 = 2$ and $\tau_I^1 = 10$ for the first loop, and $K_p^2 = -1$, $\tau_I^2 = 20$ for the second one.

The dynamic response obtained when at time $t = 1000$ min the composition X_1 changes from 5.0% to 4.95% and the cooling water inlet temperature T_{200} changes from 17 to 15 °C is shown in Figures 8 and 9. The active constraint $X_2 = 35.5$ and the SOC variable c are effectively controlled using the manipulated variables v_1 and v_2. The peeks observed in Figure 9 are in part due to unmodeled dynamics (i.e., due to the

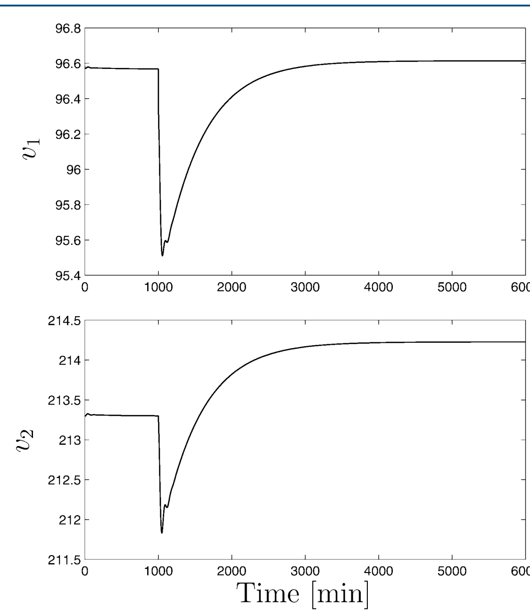

Figure 8. Self-optimizing control of the evaporator example in scenario 2. Negative disturbances in X_1 and T_{200}.

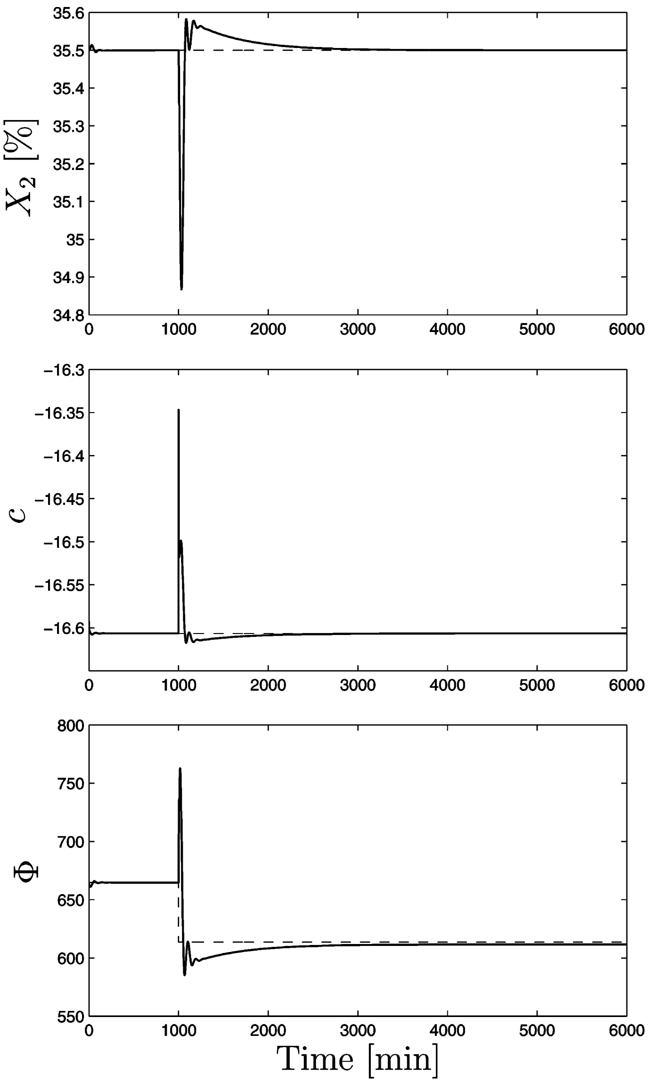

Figure 9. Self-optimizing control of the evaporator example in scenario 2. Negative disturbances in X_1 and T_{200}.

presence of algebraic equations in the simulation model). The optimality loss at steady state ($t = 6000$) is 1.9524. Notice that, in this optimization scenario the original SOC structure would require three controlled variables instead of two, if one does not fix a linear combination of the input variables.

Scenario 3. If the three disturbances $\mathbf{d} = [X_1\ T_1\ T_{200}]^{\mathrm{T}}$ are considered ($n_d = 3$), then the alternative SOC structures with reduced number of controlled variables do not apply. In this case, the original null space method described in section 3 is applied. The minimum number of inactive output variables required is $n_y^{\mathrm{na}} = 2$. In this case, we select $y_1^{\mathrm{na}} = F_4$, and $y_2^{\mathrm{na}} = T_2$.

The sensitivity matrix and the null space matrix are given by

$$\mathrm{S}^{\mathrm{na}} = \begin{bmatrix} \mathrm{S}^{y\mathrm{na}} \\ \mathrm{K}^{\mathrm{na}} \end{bmatrix} = \begin{bmatrix} 1.8208 & 0.0204 & -0.1199 \\ 10.0067 & 0.1194 & 0.3196 \\ 75.9411 & 0.7024 & -2.3685 \\ 10.2685 & 0.0567 & -0.1449 \\ 2.4522 & 0.0237 & -0.1395 \end{bmatrix}$$

$$\mathcal{N}^{\mathrm{na}} = \begin{bmatrix} 0.0140 & -0.0115 \\ 0.0033 & 0.0018 \\ -0.8027 & -0.2503 \\ 0.5819 & -0.1309 \\ -0.1299 & 0.9592 \end{bmatrix} \quad (65)$$

The corresponding control problem reads:

Controlled variables	Setpoint	Manipulated variables
$y^{\mathrm{a}} = X_2$	$y^{\mathrm{aS}} = 35.5$	F_1
$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = (\mathcal{N}^{\mathrm{na}})^{\mathrm{T}} \begin{bmatrix} \mathbf{y}^{\mathrm{na}} \\ \mathbf{u}^{\mathrm{na}} \end{bmatrix}$	$\begin{bmatrix} c_1^{\mathrm{S}} \\ c_2^{\mathrm{S}} \end{bmatrix} = \begin{bmatrix} -4.7045 \\ 6.3579 \end{bmatrix}$	F_3, F_{200}
Fixed variables	Targets	
$u^{\mathrm{a}} = P_{100}$	$u^{\mathrm{aS}} = 400$	

The matrix P_3 given in eq A-11 in the Appendix, and the RGA matrix Λ are computed as

$$\mathbf{P} = \begin{bmatrix} 0.0564 & 2.0554 & -12.2095 \\ 0.0436 & -0.2774 & -0.2157 \\ -0.0575 & 0.1986 & 0.9302 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} -0.6410 & -3.0567 & 4.6977 \\ -9.9877 & 9.5141 & 1.4737 \\ 11.6287 & -5.4574 & -5.1714 \end{bmatrix}$$

The RGA approach suggests the following input-output pairing: F_1−X_2, F_3−c_1, and F_{200}−c_2. The high absolute value of the entries in the matrix Λ is a clear signal of high interaction among the control loops. The three SISO feedback control loops previously defined are implemented via PI controllers and tuned using the classical IMC approach. This procedure gives the following setting: $K_P^1 = -0.09$ and $\tau_I^1 = 50$ for the first loop and $K_P^2 = 150$, $\tau_I^2 = 50$ for the second loop, and $K_P^3 = -1500$, $\tau_I^3 = 50$ for the last one. It is worth noting that the last two control loops (c_1 and c_2) were implemented by normalizing the corresponding data matrices (to zero mean and divided by the maximum expected values), so the PI tuning parameters correspond to this normalized situation.

At time $t = 1000$ min the composition X_1 changes from 5.0% to 4.95% and the cooling water inlet temperature T_{200} changes from 17 to 15 °C. Meanwhile, a ramp disturbance is imposed on the feed temperature T_1, starting at $t_1 = 1000$ min with $T_1(t_1) = 38$ °C, and ending at $t_2 = 2000$ min with $T_1(t_2) = 36$ °C. For brevity, only the dynamic responses of the product composition X_2 (controlled active constrained variable) and the profit are shown in Figure 10. The optimality loss at steady state ($t = 5000$) is 2.9225.

Comparison with Constraint Control. Let us introduce a fourth control policy called ACO. This control policy is based on controlling the active constrained outputs only. In this case, the control loop identified as F_1−X_2 is implemented, while the remaining degrees of freedom F_3, and F_{200}, are fixed at their nominal optimal values. In Table 3, the performance in terms of optimality loss of the different SOC schemes implemented is compared with ACO for the same disturbance cases considered for each scenario. It can be viewed that in this evaporator case study there is relatively little to win by using the null space approach with respect to simply controlling the active constraint.

 dx.doi.org/10.1021/ie5010509 | Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

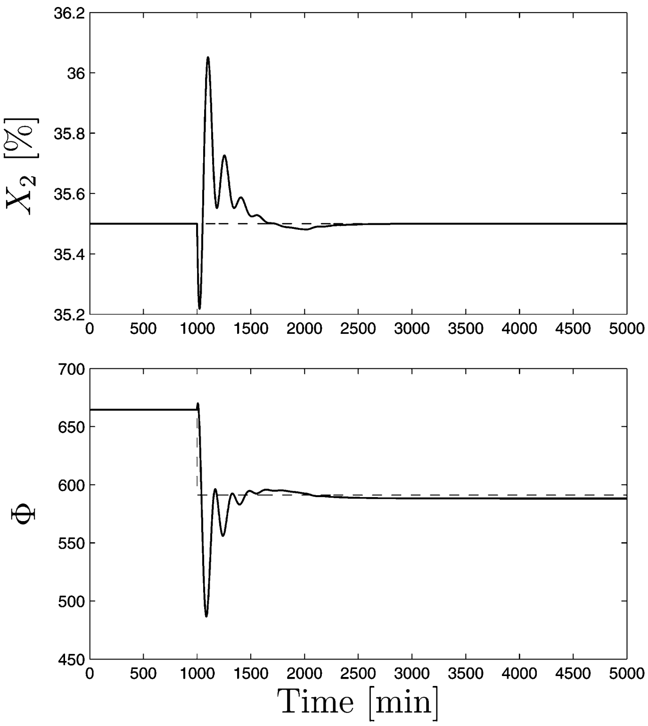

Figure 10. Self-optimizing control of the evaporator example in scenario 3. Negative disturbances in X_1, T_1, and T_{200}.

Table 3. Comparison with Active Constraint Control

disturbances			scenarios			
X_1	T_1	T_{200}	1	2	3	ACO
5.1	38	17	6.0542			8.7340
4.95	38	15		1.9524		2.2123
4.95	36	15			2.9225	3.5104

6. CONCLUSIONS

This paper has shown that, in cases where the number of disturbances is lower than the number of inactive input variables, it is possible to decrease the number of process-dependent controlled variables required in SOC structures based on the null space approach. This can be done by fixing linear combinations of the inputs to their invariant optimal values, while using the remaining input combinations for controlling both, the active constrained output variables, and additional self-optimizing controlled variables that are linear combinations of the inactive input and output variables. The minimum number of process-dependent controlled variables is obtained by fixing as many linear combinations of the inputs as possible. Based on this idea, this paper presented SOC structures with minimum number of process-dependent controlled variables for two different scenarios. Scenario 1 considered the case where the number of disturbances is lower than or equal to the number of (linearly independent) active constrained output variables, whereas scenario 2 considered the case where the number of disturbances is greater than the number of active constrained output variables and lower than the number of inactive input variables. Fixing linear combinations of the inputs by means of the proposed control structures permits to reduce the dimension of the dynamic control problem that must be solved. One of the advantages of SOC is that in many cases one can achieve near optimal operation using simple PID controllers. Hence, minimizing the number of process-dependent controlled variables permits to minimize the required number of control loops.

In the case of scenario 1, a sensitivity analysis shows that the proposed SOC structure, based on the null space method, not only minimizes the required number of control loops, but also minimizes the expected value of the optimality loss due to disturbances and measurement error (with respect to other choices of the controlled variables that include additional output variables and(or) do not satisfy the null space condition).

The applicability of the proposed SOC structures was tested in simulation in a CSTR reactor and an evaporator.

APPENDIX. RELATIVE-GAIN ARRAY AND THE SELECTION OF CONTROL LOOPS

If we take $\boldsymbol{\delta}\mathbf{d} = \mathbf{0}$, eq 18 can be decomposed as follows:

$$\begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{eq} \\ \boldsymbol{\delta}\mathbf{y}^{a} \\ \boldsymbol{\delta}\mathbf{y}^{na} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{eq}_{u^{na}} \\ \mathbf{H}^{a}_{u^{na}} \\ \mathbf{H}^{na}_{u^{na}} \end{bmatrix} \boldsymbol{\delta}\mathbf{u}^{na} \tag{A-1}$$

where $\mathbf{H}^{eq}_{u^{na}}$, $\mathbf{H}^{a}_{u^{na}}$, and $\mathbf{H}^{na}_{u^{na}}$ are obtained by eliminating from $\mathbf{H_u}$ the columns corresponding to $\boldsymbol{\delta}\mathbf{u}^a$ (i.e., the active input variables). Next, let us derive the expressions of the static open-loop gain matrix P associated with each of the controller scenarios described in section 4.

Scenario 1. In this case we have

$$\begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{eq} \\ \boldsymbol{\delta}\mathbf{y}^{a} \end{bmatrix} = P_1 \boldsymbol{\delta}\mathbf{v} \tag{A-2}$$

From eq 35 we have

$$\boldsymbol{\delta}\mathbf{u}^{na} = Q_2^{na} \boldsymbol{\delta}\mathbf{v} \tag{A-3}$$

where Q_2^{na} is obtained by eliminating all the rows in Q_2 corresponding to the active input variables. From eqs A-1—A-3 we get

$$P_1 = \begin{bmatrix} \mathbf{H}^{eq}_{u^{na}} \\ \mathbf{H}^{a}_{u^{na}} \end{bmatrix} Q_2^{na} \tag{A-4}$$

Scenario 2. In this case we have

$$\begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{eq} \\ \boldsymbol{\delta}\mathbf{y}^{a} \\ \boldsymbol{\delta}\mathbf{c} \end{bmatrix} = P_2 \boldsymbol{\delta}\mathbf{v} \tag{A-5}$$

From eq 37 we can write

$$\boldsymbol{\delta}\mathbf{c} = (\mathcal{N}^{yna})^T \begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{na} \\ \boldsymbol{\delta}\mathbf{u}^{na} \end{bmatrix} \tag{A-6}$$

Meanwhile, from eq 42 we have

$$\boldsymbol{\delta}\mathbf{u}^{na} = Q_2^{na} \boldsymbol{\delta}\mathbf{v} \tag{A-7}$$

where Q_2^{na} is obtained by eliminating all the rows in Q_2 corresponding to the active input variables. From A-1, A-5—A-7, we get

dx.doi.org/10.1021/ie5010509 | Ind. Eng. Chem. Res. XXXX, XXX, XXX−XXX

$$P_2 = \begin{bmatrix} \begin{bmatrix} \mathbf{H}^{eq}_{u^{na}} \\ \mathbf{H}^{a}_{u^{na}} \end{bmatrix} Q_2^{na} \\ (\mathcal{N}^{yna})^T \begin{bmatrix} \mathbf{H}^{na}_{u^{na}} \\ \mathbf{I} \end{bmatrix} Q_2^{na} \end{bmatrix} \tag{A-8}$$

Scenario 3. In this case we have

$$\begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{eq} \\ \boldsymbol{\delta}\mathbf{y}^{a} \\ \boldsymbol{\delta}\mathbf{c} \end{bmatrix} = P_3 \boldsymbol{\delta}\mathbf{u}^{na} \tag{A-9}$$

From eq 24 we can write

$$\boldsymbol{\delta}\mathbf{c} = (\mathcal{N}^{na})^T \begin{bmatrix} \boldsymbol{\delta}\mathbf{y}^{na} \\ \boldsymbol{\delta}\mathbf{u}^{na} \end{bmatrix} \tag{A-10}$$

Hence, from A-1, A-9, A-10, we get

$$P_3 = \begin{bmatrix} \mathbf{H}^{eq}_{u^{na}} \\ \mathbf{H}^{a}_{u^{na}} \\ (\mathcal{N}^{yna})^T \begin{bmatrix} \mathbf{H}^{na}_{u^{na}} \\ \mathbf{I} \end{bmatrix} \end{bmatrix} \tag{A-11}$$

The matrices P_i, $i = 1,2,3$, can be used to define the input-output pairing between the controlled and manipulated variables by using the well-known relative gain array (RGA) approach. Indeed, the RGA can be computed as $\Lambda = P_i \otimes (P_i^T)^{-1}$with $\otimes$ the element-by-element product.[33]

The well-conditioning of P_i guarantees confidence for the RGA approach. In this context, a design criteria used in plantwide control for selecting among many possible control structures is to find a control structure for which the matrix P_i is well conditioned.

In the cases of scenarios 2 and 3, there might be more nonactive output variables than the minimum number required. In this case, the choice of which (and how many) output variables are included in $\mathbf{y}^{na}$ will influence the condition number of P_i. On the other hand, in the three scenarios, the choice of the null space matrix $\mathcal{N}^{na}$ will also influence the condition number of P_i. Hence, a design criterion for selecting $\mathbf{y}^{na}$ and $\mathcal{N}^{na}$ could be to minimize the condition number of matrix P_i. These design decisions, which are important in order to define the final SOC structure and the pairing between the controlled and manipulated variables, fall outside the scope of this paper and may constitute a direction of future research.

■ AUTHOR INFORMATION

Corresponding Author

*Email: marchetti@cifasis-conicet.gov.ar.

Present Address

†In addition to being affiliated with CIFASIS-CONICET, David Zumoffen is also with Universidad Tecnológica Nacional, FRRo.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the financial support of CONICET (National Scientific and Technical Research Council), UNR-FCEIA (National University of Rosario), and ANPCYT PICT 2009-0044 (Agencia Nacional de Promoción Científica y Técnica) from Argentina.

■ REFERENCES

(1) François, G.; Srinivasan, B.; Bonvin, D. Comparison of six implicit real-time optimization schemes. *J. Eur. Syst. Autom.* **2012**, *46*, 291−305.

(2) Marlin, T. E.; Hrymak, A. N. Real-time operations optimization of continuous processes. *AIChE Symp. Ser.* **1997**, 156−164.

(3) Darby, M. L.; Nikolaou, M.; Jones, J.; Nicholson, D. RTO: An overview and assessment of current practice. *J. Process Contr.* **2011**, *21*, 874−884.

(4) Maarleveld, A.; Rijnsdorp, J. E. Constraint control on distillation columns. *Automatica* **1970**, *6*, 51−58.

(5) Garcia, C. E.; Morari, M. Optimal operation of integrated processing systems. Part II: Closed-loop on-line optimizing control. *AIChE J.* **1984**, *30*, 226−234.

(6) Arkun, Y.; Stephanopoulos, G. Studies in the synthesis of control structures for chemical processes: Part IV. Design of steady-state optimizing control structures for chemical process units. *AIChE J.* **1980**, *26*, 975−991.

(7) Tatjewski, P.; Brdyś, M. A.; Duda, J. Optimizing control of uncertain plants with constrained feedback controlled outputs. *Int. J. Control* **2001**, *74*, 1510−1526.

(8) Skogestad, S. Self-optimizing control: The missing link between steady-state optimization and control. *Comput. Chem. Eng.* **2000**, *24*, 569−575.

(9) Skogestad, S. Plantwide control: The search for the self-optimizing control structure. *J. Process Contr.* **2000**, *10*, 487−507.

(10) François, G.; Srinivasan, B.; Bonvin, D. Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty. *J. Process Contr.* **2005**, *15*, 701−712.

(11) Srinivasan, B.; Biegler, L. T.; Bonvin, D. Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function. *Comput. Chem. Eng.* **2008**, *32*, 572−579.

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

(13) Kariwala, V.; Cao, Y.; Janardhanan, S. Local self-optimizing control with average loss minimization. *Ind. Eng. Chem. Res.* **2008**, *47*, 1150−1158.

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

(15) Heldt, S. Dealing with structural constraints in self-optimizing control engineering. *J. Process Contr.* **2010**, *20*, 1049−1058.

(16) Yelchuru, R.; Skogestad, S. Convex formulations for optimal selection of controlled variables and measurements using Mixed Integer Quadratic Programming. *J. Process Contr.* **2012**, *22*, 995−1007.

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

(18) Alstad, V.; Skogestad, S.; Hori, E. S. Optimal measurement combinations as controlled variables. *J. Process Contr.* **2009**, *19*, 138−148.

(19) Marchetti, A. G.; Zumoffen, D. On the links between real-time optimization, neighboring-extremal control, and self-optimizing control. *European Control Conference, ECC'13*; Zurich, Switzerland, 2013.

(20) Manum, H.; Skogestad, S. Self-optimizing control with active set changes. *J. Process Contr.* **2012**, *22*, 873−883.

(21) Nocedal, J.; Wright, S. J. *Numerical Optimization*; Springer-Verlag: New York, 1999.

(22) Chachuat, B.; Marchetti, A.; Bonvin, D. Process optimization via constraints adaptation. *J. Process Contr.* **2008**, *18*, 244−257.

(23) Gros, S.; Srinivasan, B.; Bonvin, D. Optimizing control based on output feedback. *Comput. Chem. Eng.* **2009**, *33*, 191−198.
(24) Bamberger, W.; Isermann, R. Adaptive On-line steady state optimization of slow dynamic processes. *Automatica* **1978**, *14*, 223−230.
(25) Ariyur, K. B.; Krstić, M. *Real-Time Optimization by Extremum-Seeking Control*; John Wiley and Sons: Hoboken, NJ, 2003.
(26) Loeblein, C.; Perkins, J. D. Economic analysis of different structures of on-line process optimization systems. *Comput. Chem. Eng.* **1998**, *22*, 1257−1269.
(27) Bazaraa, M. S.; Sherali, H. D.; Shetty, C. M. *Nonlinear Programming: Theory and Algorithms*, 3rd ed.; John Wiley and Sons: Hoboken, NJ, 2006.
(28) Williams, T. J.; Otto, R. E. A generalized chemical processing model for the investigation of computer control. *AIEE Trans.* **1960**, *79*, 458.
(29) Marchetti, A. G.; Ferramosca, A.; González, A. H. Steady-state target optimization designs for integrating real-time optimization and model predictive control. *J. Process Contr.* **2014**, *24*, 129−145.
(30) Newell, R. B.; Lee, P. L. *Applied Process Control: A Case Study*; Prentice Hall: Englewood Cliffs, NJ, 1989.
(31) Hu, W.; Umar, L. M.; Xiao, G.; Kariwala, V. Local self-optimizing control of constrained processes. *J. Process Contr.* **2012**, *22*, 488−493.
(32) Maciejowski, J. M. *Predictive Control with Constraints*; Prentice Hall: Upper Saddle River, NJ, 2002.
(33) Skogestad, S.; Postlethwaite, I. *Multivariable Feedback Control. Analysis and Design*, 2nd ed.; John Wiley & Sons: New York, 2005.

dx.doi.org/10.1021/ie5010509 | *Ind. Eng. Chem. Res.* XXXX, XXX, XXX−XXX