2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

On the Links Between Real-Time Optimization, Neighboring-Extremal Control, and Self-Optimizing Control

Alejandro G. Marchetti and David Zumoffen

Abstract— **In the context of process optimization, several techniques have been proposed in order to operate the plant near the economically optimal steady-state operating point. This paper discusses the links and differences between three of these techniques: (i) real-time optimization using the classical two-step approach, (ii) neighboring-extremal control with output feedback, and (iii) self-optimizing control using the null space method. It is argued that, under certain assumptions, the technique (ii) can be viewed as a first-order approximation of (i). Also, while (ii) and (iii) enforce similar optimal operations, their dynamic implementation and design features are very different.**

I. INTRODUCTION

Most industrial chemical processes are characterized by continuously operating plants, for which optimal steady-state operation is of economic importance. Several techniques have been proposed in order to achieve optimal operation. In highly automated plants, optimal operation is typically addressed by a decision hierarchy involving several layers that include plant scheduling, real-time optimization (RTO), and process control [1], [2]. At the RTO layer, the optimal steady-state operating point is computed by solving a nonlinear programming (NLP) problem which uses a detailed steady-state model of the plant. The RTO executions take place when the plant is operating at (pseudo) steady-state conditions, which requires a steady-state identification subsystem [3], [4]. Several RTO approaches have been proposed in the literature . The classical two-step approach implies an iteration between parameter estimation and optimization [5], [1], [2]. The idea is to estimate repeatedly the disturbance and parameter values of the nonlinear steady-state model, and to use the updated model in the model optimization to generate new inputs. This way, the model is expected to represent the plant at its current operating point more accurately. This approach works well provided that there is little structural plant-model mismatch, and the parameters are identifiable from the available measurements [6].

Neighboring extremal control (NEC) with output feedback [7], and self-optimizing control (SOC) [8],[9] are two process optimization techniques that do not require solving an optimization problem online. This paper analyzes the similarities and differences between NEC with output feedback, SOC using the null space method (see [10]), and the classical two-step approach of RTO. The relation between the three techniques is mathematically formalized, and the underlying assumptions are highlighted.

A comparison of how the gradients are estimated in these three techniques, together with other three model-free techniques, was performed in [11], [12]. In these papers, the performance of an integral gradient controller is compared using the different gradient estimates obtained by each technique. This comparison is different from the comparison done in this paper, since here the two-step approach is implemented at the RTO layer, and the NEC technique given in [7] is considered, which does not use a dynamic controller.

This paper is organized as follows. Section II formulates the optimization problem and presents the first-order variations of the necessary conditions of optimality (NCO). The three process optimization techniques analyzed in this work are described in Section III. The links between these techniques are discussed in Section IV, and their performances are evaluated via the reactor of the Williams-Otto plant in Section V . Finally, Section VI concludes the paper.

II. PROBLEM FORMULATION

A. Optimization Problem

The steady-state optimization problem consists of an NLP problem where an economic cost function is minimized subject to equality and inequality constraints. Some of the inequality constraints might be active at the optimum solution. Neighboring-extremal control and self-optimizing control (in its original form) are both based on the following invariant active set assumption:

Assumption 1 (Invariant Active Set) *The set of active constraints is known and does not change with process disturbances. All the active constraints are kept active using feedback controllers.*

Controlling the active constrained variables has been widely adopted in implicit process optimization approaches such as constraint control [13], optimizing control [14],[15], and NCO-traking [16]. Using Assumption 1, and considering that a number of degrees of freedom (inputs) are consumed in controlling the active constraints, the steady-state optimization problem for the plant can be formulated in terms of the remaining inputs as follows:

$$\mathbf{u}_p^\star = \arg\min_{\mathbf{u}} \quad \phi_p(\mathbf{x}_p, \mathbf{u}, \mathbf{d}_p) \tag{1}$$

$$\text{s.t.} \quad \mathbf{F}_p(\mathbf{x}_p, \mathbf{u}, \mathbf{d}_p) = \mathbf{0}, \tag{2}$$

where ϕ_p is the scalar cost function to be minimized, $\mathbf{u} \in \mathbb{R}^{n_u}$ denotes the decision (or input) variables, $\mathbf{x}_p \in \mathbb{R}^{n_{x_p}}$ are

A. G. Marchetti and D. Zumoffen are with the Computer-Aided Process Engineering Group, French-Argentine International Center for Information and Systems Sciences (CIFASIS-CONICET-UNR-AMU), 27 de Febrero 210 bis, S2000EZP Rosario, Argentina.
marchetti/zumoffen@cifasis-conicet.gov.ar

the state variables, $\mathbf{d}_p \in \mathbb{R}^{n_{d_p}}$ is the set of plant parameters and process disturbances, and $\mathbf{F}_p \in \mathbb{R}^{n_{x_p}}$ is the vector of algebraic equations that describe the steady-state operation of the plant. The notation $(\cdot)_p$ is used for the variables associated with the plant.

Remark 1 *Notice that, in practice there is no mode representing exactly the plant, which should be seen as a map between the inputs and plant outputs. Hence, equation* (2) *should be viewed as an hypothetical representation that will be used here for the purpose of proving certain properties.*

Using an approximate model, the solution of the original problem (1, 2) can be approached by solving the following NLP problem:

$$\mathbf{u}^\star(\mathbf{d}) = \arg\min_{\mathbf{u}} \quad \phi(\mathbf{x},\mathbf{u},\mathbf{d}) \tag{3}$$

$$\text{s.t.} \quad \mathbf{F}(\mathbf{x},\mathbf{u},\mathbf{d}) = \mathbf{0}, \tag{4}$$

where ϕ is the scalar cost function predicted by the model, $\mathbf{x} \in \mathbb{R}^{n_x}$ are the model state variables, $\mathbf{d} \in \mathbb{R}^{n_d}$ is the set of model parameters and disturbances, and $\mathbf{F} \in \mathbb{R}^{n_x}$ is the set of model equations.

B. Necessary Conditions of Optimality

For a fixed value of the disturbances $\mathbf{d}$, the NCO for Problem (3, 4) read:[1]

$$\mathcal{L}_{\mathbf{u}} = \phi_{\mathbf{u}} + \boldsymbol{\mu}^\mathsf{T}\mathbf{F}_{\mathbf{u}} = \mathbf{0}, \tag{5}$$

$$\mathcal{L}_{\mathbf{x}} = \phi_{\mathbf{x}} + \boldsymbol{\mu}^\mathsf{T}\mathbf{F}_{\mathbf{x}} = \mathbf{0}, \tag{6}$$

$$\mathcal{L}_{\boldsymbol{\mu}} = \mathbf{F}^\mathsf{T} = \mathbf{0}. \tag{7}$$

with $\mathcal{L}(\mathbf{x},\mathbf{u},\mathbf{d},\boldsymbol{\mu}) = \phi(\mathbf{x},\mathbf{u},\mathbf{d}) + \boldsymbol{\mu}^\mathsf{T}\mathbf{F}(\mathbf{x},\mathbf{u},\mathbf{d})$ being the Lagrangian function, and $\boldsymbol{\mu} \in \mathbb{R}^{n_x}$ the Lagrange multipliers.

C. Sensitivity Analysis

Let $\mathbf{d}_{\text{nom}}$ be the nominal parameter values, and consider the parametric disturbance $\delta\mathbf{d} = \mathbf{d} - \mathbf{d}_{\text{nom}}$. The deviation of the optimal inputs induced by $\delta\mathbf{d}$ is $\delta\mathbf{u}^\star = \mathbf{u}^\star(\mathbf{d}) - \mathbf{u}^\star(\mathbf{d}_{\text{nom}})$, and the corresponding deviations of the states and Lagrange multipliers are $\delta\mathbf{x}^\star$ and $\delta\boldsymbol{\mu}^\star$, respectively. The first-order variations of the NCO (5-7) read:

$$\mathcal{L}_{\mathbf{ux}}\delta\mathbf{x} + \mathcal{L}_{\mathbf{uu}}\delta\mathbf{u} + \mathbf{F}_{\mathbf{u}}^\mathsf{T}\delta\boldsymbol{\mu} + \mathcal{L}_{\mathbf{ud}}\delta\mathbf{d} = \mathbf{0}, \tag{8}$$

$$\mathcal{L}_{\mathbf{xx}}\delta\mathbf{x} + \mathcal{L}_{\mathbf{xu}}\delta\mathbf{u} + \mathbf{F}_{\mathbf{x}}^\mathsf{T}\delta\boldsymbol{\mu} + \mathcal{L}_{\mathbf{xd}}\delta\mathbf{d} = \mathbf{0}, \tag{9}$$

$$\mathbf{F}_{\mathbf{x}}\delta\mathbf{x} + \mathbf{F}_{\mathbf{u}}\delta\mathbf{u} + \mathbf{F}_{\mathbf{d}}\delta\mathbf{d} = \mathbf{0}, \tag{10}$$

where $\delta\mathbf{x}$, $\delta\mathbf{u}$, and $\delta\boldsymbol{\mu}$ are first-order approximations of $\delta\mathbf{x}^\star$, $\delta\mathbf{u}^\star$, and $\delta\boldsymbol{\mu}^\star$, respectively. From (8-10) one can obtain the first-order variation of the optimal inputs [7]:

$$\delta\mathbf{u} = \mathsf{K}\delta\mathbf{d}, \tag{11}$$

with

$$\begin{aligned}\mathsf{K} = -\Big(&\mathcal{L}_{\mathbf{uu}} - \mathcal{L}_{\mathbf{ux}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{u}} - \mathbf{F}_{\mathbf{u}}^\mathsf{T}\mathbf{F}_{\mathbf{x}}^{-1^\mathsf{T}}\mathcal{L}_{\mathbf{xu}} \\ &+ \mathbf{F}_{\mathbf{u}}^\mathsf{T}\mathbf{F}_{\mathbf{x}}^{-1^\mathsf{T}}\mathcal{L}_{\mathbf{xx}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{u}}\Big)^{-1}\left(\mathcal{L}_{\mathbf{ud}} - \mathcal{L}_{\mathbf{ux}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{d}}\right. \\ &\left. - \mathbf{F}_{\mathbf{u}}^\mathsf{T}\mathbf{F}_{\mathbf{x}}^{-1^\mathsf{T}}\mathcal{L}_{\mathbf{ud}} + \mathbf{F}_{\mathbf{u}}^\mathsf{T}\mathbf{F}_{\mathbf{x}}^{-1^\mathsf{T}}\mathcal{L}_{\mathbf{xx}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{d}}\right).\end{aligned}$$

[1]The notation $\mathbf{a}_\mathbf{b} = \frac{\partial \mathbf{a}}{\partial \mathbf{b}}$ is used.

III. DESCRIPTION OF THE TECHNIQUES

All the techniques use feedback information from the plant in order to deal with plant-model mismatch. The measurement equation for the plant reads

$$\mathbf{y}_p = \mathbf{H}_p(\mathbf{x}_p), \tag{12}$$

where $\mathbf{y}_p \in \mathbb{R}^{n_y}$ are the (measured) plant outputs. Combining equations (12) and (2), the plant outputs can be expressed as a functional of $\mathbf{u}$ and $\mathbf{d}_p$, that is, $\mathbf{y}_p(\mathbf{u},\mathbf{d}_p)$. On the other hand, the outputs predicted by the model are

$$\mathbf{y} = \mathbf{H}(\mathbf{x}). \tag{13}$$

The predicted outputs $\mathbf{y} \in \mathbb{R}^{n_y}$ can be expressed as the functional $\mathbf{y}(\mathbf{u},\mathbf{d})$ by combining (13) and (4).

Neighboring-extremal control and self-optimizing control are both based on the following model accuracy assumption:

Assumption 2 (No Structural Plant-Model Mismatch) *The cost function and model equations are structurally correct, i.e., $\phi = \phi_p$, $\mathbf{F} = \mathbf{F}_p$, 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* $\mathbf{d}$.

A. Real-Time Optimization using the Classical Two-Step Approach

Let $\mathbf{u}_k$ be the input applied to the plant at the RTO iteration k. Using Assumption 1, the inequality constraints can be eliminated from the NLP problem. Hence, the parameter estimation and optimization problems at the kth RTO execution read:

Parameter Estimation Step[2]

$$\mathbf{d}_k^\star = \arg\min_{\mathbf{d}} \quad \left\|\mathbf{y}_p(\mathbf{u}_k^\star,\mathbf{d}_p) - \mathbf{y}(\mathbf{u}_k^\star,\mathbf{d})\right\|_{\mathsf{R}}^2, \tag{14}$$

where R is a positive definite weighting matrix.

Optimization Step

$$\mathbf{u}_{k+1}^\star = \arg\min_{\mathbf{u}} \quad \phi(\mathbf{x},\mathbf{u},\mathbf{d}_k^\star) \tag{15}$$

$$\text{s.t.} \quad \mathbf{F}(\mathbf{x},\mathbf{u},\mathbf{d}_k^\star) = \mathbf{0} \tag{16}$$

Input Update

$$\mathbf{u}_{k+1} = \mathbf{u}_{k+1}^\star \tag{17}$$

Remark 2 *Assumption 1 is used in* (14)-(16) *because the purpose here is to establish the relation between the two-step approach of RTO and the other two techniques. In practice, the active set can change with the disturbance values, and by incorporating inequality constraints in the optimization problem the RTO optimizer has the ability to detect these changes.*

Remark 3 *If Assumption 2 holds, and the parameters are identifiable from the output measurements, then the two-step approach can converge to the plant optimum in a single iteration [6].*

[2]The notation $\|\mathbf{v}\|_{\mathsf{M}}^2 := \mathbf{v}^\mathsf{T}\mathsf{M}\mathbf{v}$ is used henceforth.

B. Neighboring-Extremal Control with Output Feedback

Gros et al. (2009) [7] proposed a linear input update law based on the first-order variation of the NCO. Let us assume that at iteration k the process is subject to the constant parametric disturbance $\delta\mathbf{d}$. From (11), the corrective action proposed in order to bring the process back to optimality at iteration $k+1$ is:

$$\delta\mathbf{u}_{k+1} = \mathsf{K}\delta\mathbf{d}. \tag{18}$$

The linearized form of (13) at the kth iteration is:

$$\delta\mathbf{y}_k = \mathbf{H}_{\mathbf{x}}\delta\mathbf{x}_k \tag{19}$$

From (10) at the kth iteration we have

$$\delta\mathbf{x}_k = -\mathbf{F}_{\mathbf{x}}^{-1}\left(\mathbf{F}_{\mathbf{u}}\delta\mathbf{u}_k + \mathbf{F}_{\mathbf{d}}\delta\mathbf{d}\right) \tag{20}$$

Using (20) in (19),

$$\delta\mathbf{y}_k = -\mathbf{H}_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}^{-1}\left(\mathbf{F}_{\mathbf{u}}\delta\mathbf{u}_k + \mathbf{F}_{\mathbf{d}}\delta\mathbf{d}\right) \tag{21}$$

Let $\mathcal{M} = -\mathbf{H}_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{d}}$. Using (21), an estimate of the parameter deviations can be obtained from the output measurements as:

$$\delta\mathbf{d}_k = \mathcal{M}^{\dagger}\left(\delta\mathbf{y}_{p,k} + \mathbf{H}_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{u}}\delta\mathbf{u}_k\right), \tag{22}$$

where $\mathcal{M}^{\dagger}$ stands for the pseudoinverse of $\mathcal{M}$, and $\delta\mathbf{y}_{p,k} = \mathbf{y}_p(\mathbf{u}_k, \mathbf{d}_p) - \mathbf{y}(\mathbf{u}_k, \mathbf{d}_{\text{nom}})$. Replacing (22) in (18) gives the input update law [7]:

$$\delta\mathbf{u}_{k+1} = \mathsf{K}^y\delta\mathbf{y}_{p,k} + \mathsf{K}^u\delta\mathbf{u}_k \tag{23}$$

with the gains

$$\mathsf{K}^y = \mathsf{K}\mathcal{M}^{\dagger}, \qquad \mathsf{K}^u = \mathsf{K}^y\mathbf{H}_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{u}}. \tag{24}$$

Remark 4 *The correct parameter values can be uniquely identified from* (22) *if $n_y \geq n_d$, and $\mathcal{M}$ has full column rank, or equivalently, if $\mathcal{M}^{\dagger}\mathcal{M} = \mathbf{I}_{n_d}$.*

Remark 5 *Notice that the control law* (23) *is derived from a static optimization problem that uses a steady-state model. Hence, the plant measurements and the control iterations should (in principle) take place when the plant is at steady state. Therefore, NEC for static optimization problems occupies the same layer in the control hierarchy as RTO.*

Remark 6 *In [7] it is demonstrated that if: (i) Assumption 2 holds, (ii) the model equations are linear (meaning that* (10) *and* (19) *are exact), and (iii) $\mathcal{M}^{\dagger}\mathcal{M} = \mathbf{I}$, then the input update law* (23) *converges to the optimum given by* (11) *in at most two iterations. This result is correct since indeed after two iterations the approach has converged to the optimum. However, it turns out that under these conditions the approach converges to the optimum in a single iteration. The condition (ii) implies that equation* (21) *is exact. Therefore, convergence in a single iteration results from the fact that the correct parameter deviations $\delta\mathbf{d}$ can be estimated from* (22) *at any operating point $\delta\mathbf{u}_k$.*

C. Self-Optimizing Control using the Null Space Method

The idea of self-optimizing control is to select a set of n_u controlled variables $\mathbf{c} \in \mathbb{R}^{n_u}$ as functions of the output and input variables, i.e., $\mathbf{c} = \mathbf{C}(\mathbf{y}, \mathbf{u})$, such that near optimal operation is achieved by keeping $\mathbf{c}$ at the constant setpoints $\mathbf{c}_s$ [8],[9].

The null space method is based on the sensitivity matrix of the optimal outputs with respect to the disturbances [10]. Two alternative designs will be discussed.

Design A: Using (11) in (21) we obtain:

$$\delta\mathbf{y} = -\mathbf{H}_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}^{-1}\left(\mathbf{F}_{\mathbf{u}}\mathsf{K} + \mathbf{F}_{\mathbf{d}}\right)\delta\mathbf{d} = \mathsf{S}^y\delta\mathbf{d} \tag{25}$$

Let the columns in $\mathbf{N} \in \mathbb{R}^{n_y \times n_u}$ be a set of n_u orthonormal vectors that lie in the left null space of S^y. Hence $\mathbf{N}^{\mathsf{T}}\mathsf{S}^y = \mathbf{0}$, and from (25) we have,

$$\mathbf{N}^{\mathsf{T}}\delta\mathbf{y} = \mathbf{N}^{\mathsf{T}}\mathsf{S}^y\delta\mathbf{d} = \mathbf{0}. \tag{26}$$

Based on this result, the null space method consists in selecting $\mathbf{c} = \mathbf{N}^{\mathsf{T}}\mathbf{y}$ and $\mathbf{c}_s = \mathbf{N}^{\mathsf{T}}\mathbf{y}(\mathbf{u}^{\star}(\mathbf{d}_{\text{nom}}), \mathbf{d}_{\text{nom}})$ [10]. The dimension of the left null space of S^y should be greater than or equal to n_u, which requires that $n_y \geq n_u + n_d$. This way, the controlled variables $\mathbf{c}$ are selected as linear combinations of the measured output variables $\mathbf{y}$.

Design B: Alternatively, as proposed in [10], one can select $\mathbf{c}$ as linear combinations of the output and input variables. From (25) and (11) we have

$$\begin{bmatrix} \delta\mathbf{y} \\ \delta\mathbf{u} \end{bmatrix} = \begin{bmatrix} -\mathbf{H}_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}^{-1}\left(\mathbf{F}_{\mathbf{u}}\mathsf{K} + \mathbf{F}_{\mathbf{d}}\right) \\ \mathsf{K} \end{bmatrix}\delta\mathbf{d} = \mathsf{S}^{yu}\delta\mathbf{d}. \tag{27}$$

In this case, we select $\mathbf{N} \in \mathbb{R}^{(n_y+n_u)\times n_u}$ such that $\mathbf{N}^{\mathsf{T}}\mathsf{S}^{yu} = \mathbf{0}$, which requires that $n_y \geq n_d$. The controlled variables and setpoint values are:

$$\mathbf{c} = \mathbf{N}^{\mathsf{T}}\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix}, \qquad \mathbf{c}_s = \mathbf{N}^{\mathsf{T}}\begin{bmatrix} \mathbf{y}(\mathbf{u}^{\star}(\mathbf{d}_{\text{nom}}), \mathbf{d}_{\text{nom}}) \\ \mathbf{u}^{\star}(\mathbf{d}_{\text{nom}}) \end{bmatrix}.$$

Remark 7 *Notice that in [10],[17] the null space method is presented including the degrees of freedom $\mathbf{u}$ in the linear combinations, as in Design B. Nevertheless, the illustrative examples presented in these papers make use of Design A, where the controlled variables $\mathbf{c}$ are linear combinations of the output variables $\mathbf{y}$ only.*

IV. LINKS BETWEEN THE TECHNIQUES

A. Link between RTO and Neighboring-Extemal Control with Output Feedback

It will be shown that the NEC scheme (23) is a first-order approximation of the classical two-step approach of RTO.

Proposition 1 *If Assumptions 1 and 2 hold and $\mathcal{M}^{\dagger}\mathcal{M} = \mathbf{I}$, then the NEC scheme* (23) *can be viewed as a first-order approximation of the classical two-step approach of RTO.*

Proof: The NEC scheme (23) is obtained by combining equations (18) and (22). From the sensitivity analysis in Section II, we know that (18) gives the first-order variation to the solution of the optimization step (15),(16). Hence, we

only need to show that (22) can be viewed as a first-order approximation to the solution of the parameter estimation step (14). Using Assumption 2, the parameter estimation problem (14) can be rewritten as:

$$\begin{aligned} \min_{\mathbf{d}} \quad & \left\|\mathbf{H}(\mathbf{x}_{p,k}) - \mathbf{H}(\mathbf{x})\right\|_{\mathsf{R}}^2 \\ \text{s.t.} \quad & \mathbf{F}(\mathbf{x}_{p,k}, \mathbf{u}_k, \mathbf{d}_p) = \mathbf{0}, \\ & \mathbf{F}(\mathbf{x}, \mathbf{u}_k, \mathbf{d}) = \mathbf{0}, \end{aligned} \tag{28}$$

where $\mathbf{d}_p$ are the real disturbances for the plant. Assuming that $\mathbf{F}(\mathbf{x}, \mathbf{u}_k, \mathbf{d}) = \mathbf{0}$ has a unique solution $\mathbf{x}$ for $\mathbf{u}_k$ and $\mathbf{d}$ given (which is consistent with $\mathbf{F_x}$ being invertible), it is clear that $\mathbf{d} = \mathbf{d}_p$ is a minimizing solution to problem (28). This solution can be made unique by an appropriate choice of the positive definite matrix R. Hence, solving (28) is equivalent to solving the equations $\mathbf{H}(\mathbf{x}_{p,k}) = \mathbf{H}(\mathbf{x})$ and $\mathbf{F}(\mathbf{x}, \mathbf{u}_k, \mathbf{d}) = \mathbf{0}$. Equation (22) gives a first-order approximation to the solution of these equations, and therefore, (22) can be viewed as a first-order approximation to the solution of problem (28). ■

Similarly to the two-step approach of RTO, the performance of the NEC scheme strongly relies on the accuracy of the process model and the identifiability of the model parameters, with the additional limitation of being a local approximation method. Since the control iterations take place when the plant is at (pseudo) steady state, a steady-state identification scheme should be implemented, the same as with RTO.

B. Links between Neighboring-Extemal Control with Output Feedback and Self-Optimizing Control using the Null Space Method

NEC with output feedback and the null space method are both based on the same sensitivity analysis of the optimal inputs and outputs with respect to deviations in the parameter values. NEC directly determines (estimates of) the optimal input corrections, which are applied to the plant in open-loop fashion, whereas the null space method consists in designing a square multivariable controller such that the optimal input corrections are reached by the controlled plant at steady state. This link was already recognized in [17]. Here, the relation between both approaches is mathematically formalized.

Consider first SOC using Design A. The equation $\mathbf{N}^\mathsf{T}\mathsf{S}^y = \mathbf{0}$ can be rewritten as:

$$\mathbf{N}^\mathsf{T}\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u}\mathsf{K} = -\mathbf{N}^\mathsf{T}\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_d}. \tag{29}$$

Using (21) in (26) we have:

$$\mathbf{0} = \mathbf{N}^\mathsf{T}\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u}\delta\mathbf{u} + \mathbf{N}^\mathsf{T}\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_d}\delta\mathbf{d}. \tag{30}$$

Combining (29) and (30) we have

$$\mathbf{N}^\mathsf{T}[\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u}]\delta\mathbf{u} = \mathbf{N}^\mathsf{T}[\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u}]\mathsf{K}\delta\mathbf{d}. \tag{31}$$

The linear system (31) has the unique solution $\delta\mathbf{u} = \mathsf{K}\delta\mathbf{d}$ if matrix $\mathbf{N}^\mathsf{T}[\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u}]$ is nonsingular. This requires matrix $\mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u}$ to be full column rank, which requires $n_x \geq n_u$.

Next, consider SOC using Design B. In this case, after operating we arrive to the system:

$$\mathbf{N}^\mathsf{T}\mathsf{Z}\ \delta\mathbf{u} = \mathbf{N}^\mathsf{T}\mathsf{Z}\ \mathsf{K}\delta\mathbf{d}, \tag{32}$$

with

$$\mathsf{Z} = \begin{bmatrix} \mathbf{H_x}\mathbf{F_x}^{-1}\mathbf{F_u} \\ -\mathbf{I}_{n_u} \end{bmatrix}.$$

The linear system (32) has the unique solution $\delta\mathbf{u} = \mathsf{K}\delta\mathbf{d}$ if matrix $\mathbf{N}^\mathsf{T}\mathsf{Z}$ is nonsingular. Notice that Z is full column rank.

V. ILLUSTRATIVE CASE STUDY

The reactor in the Williams-Otto plant is considered [18]. It consists of an ideal CSTR in which the following reactions occur:

$$\begin{aligned} A + B &\longrightarrow C, & k_1 &= 1.6599 \times 10^6 e^{-6666.7/(T_R+273)} \\ B + C &\longrightarrow P + E, & k_2 &= 7.2117 \times 10^8 e^{-8333.3/(T_R+273)} \\ C + P &\longrightarrow G, & k_3 &= 2.6745 \times 10^{12} e^{-11111/(T_R+273)} \end{aligned}$$

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

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

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. 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:

$$\dot{X}_A = \frac{F_A}{W} - \frac{(F_A + F_B)}{W} X_A - r_1, \tag{33}$$

$$\dot{X}_B = \frac{F_B}{W} - \frac{(F_A + F_B)}{W} X_B - r_1 - r_2, \tag{34}$$

$$\dot{X}_C = -\frac{(F_A + F_B)}{W} X_C + 2r_1 - 2r_2 - r_3, \tag{35}$$

$$\dot{X}_P = -\frac{(F_A + F_B)}{W} X_P + r_2 - \frac{1}{2} r_3, \tag{36}$$

$$\dot{X}_G = -\frac{(F_A + F_B)}{W} X_G + \frac{3}{2} r_3, \tag{37}$$

$$\dot{X}_E = -\frac{(F_A + F_B)}{W} X_E + 2r_2, \tag{38}$$

$$\begin{aligned} W C_p \dot{T}_R = {} & F_B C_p T_{inB} + F_A C_p T_{inA} - \Delta H_1 W r_1 \\ & - \Delta H_2 W r_2 - \Delta H_3 W r_3 - \frac{W}{1000} \frac{A_o}{V_o} U (T_R - T_J) \\ & - (F_A + F_B) C_p T_R, \end{aligned} \tag{39}$$

$$\begin{aligned} W_J C_{pJ} \dot{T}_J = {} & F_J C_{pJ} (T_{Jin} - T_J) \\ & + \frac{W}{1000} \frac{A_o}{V_o} U (T_R - T_J). \end{aligned} \tag{40}$$

where (33)-(38) are the species mass balances; (39) is the heat balance in the reactor, and (40) 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.

TABLE I

MODEL VARIABLES AND PARAMETERS

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)

The decision variables are F_B and the setpoint of the reactor temperature, T_R^{sp}, i.e., $\mathbf{u} = [F_B \;\; T_R^{sp}]^\mathsf{T}$. The disturbance variable considered is F_A, i.e., $d_p = F_A$. The set of equations (33)-(38) is $\dot{\mathbf{x}}_p = \mathbf{F}_p(\mathbf{x}_p, \mathbf{u}, d_p)$, with $\mathbf{x}_p = [X_A \;\; X_B \;\; X_C \;\; X_P \;\; X_G \;\; X_E]^\mathsf{T}$. 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 $\mathbf{0} = \mathbf{F}_p(\mathbf{x}_p, \mathbf{u}, d_p)$. The objective is to maximize profit at steady state operation, which is expressed as the price difference between the products and the reactants:

$$\Phi = 1143.38X_PF + 25.92X_EF - 76.23F_A - 114.34F_B$$

The nominal value of F_A is $F_{A,\text{nom}} = 1.8275$ kg/s. The nominal optimum point is $\mathbf{u}^\star(F_{A,\text{nom}}) = [4.7874 \;\; 89.7]^\mathsf{T}$. At this point, the gain matrix K in (11) is computed as $\mathsf{K} = [2.3329 \;\; 6.1436]^\mathsf{T}$. In this case study, Assumptions 1 and 2 hold. The performances of NEC and SOC are evaluated for the case of a step-wise disturbance on F_A, which varies from the nominal value $F_{A,\text{nom}}$ to the perturbed value $F_A = 2.3$ kg/s at time $t = 20$ min. The measured output variables are selected as $\mathbf{y} = [X_A \;\; X_B \;\; X_P]^\mathsf{T}$.

A. Two-step Approach of RTO

The implementation is not shown. Since Assumption 2 holds, it is possible to identify the correct value of F_A (provided the measurements are taken at the new steady state reached after the disturbance in F_A takes place), and to converge to the plant optimum in a single RTO iteration.

B. Neighboring-Extremal Control

The implementation of the NEC technique requires the online identification of steady state. There are various techniques available for steady-state identification (see [3], [4] and references therein). Most of these are based on using a given statistical test on a set of data that belongs to a selected data window. Here, a simplified approach is adopted. The reactor will be considered to be at (near) steady state if, in a data window of 10 min, the data ranges of X_A and X_B are lower than or equal to 0.001, and the data range of T_w is lower than or equal to 0.5. This steady state test is performed with a frequency of 1 min, and if it is positive, the input update law (23) is applied. The gain matrices in

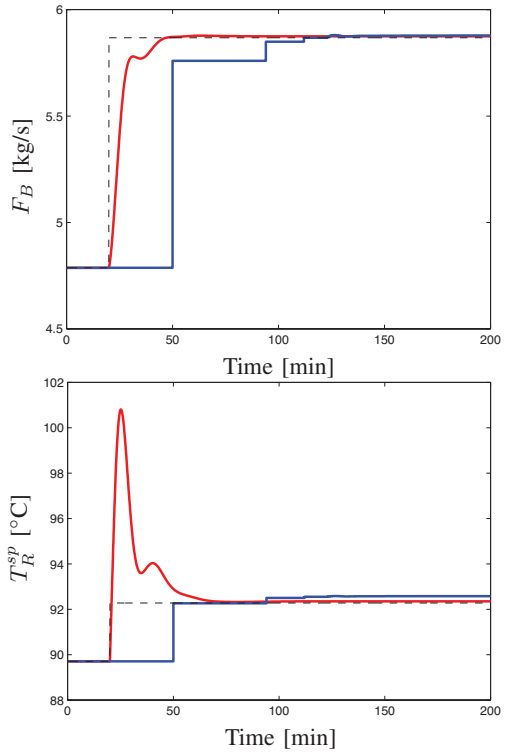

Fig. 1. Decision variables. Dashed line: steady state optimum; Blue line: NEC; Red line: SOC.

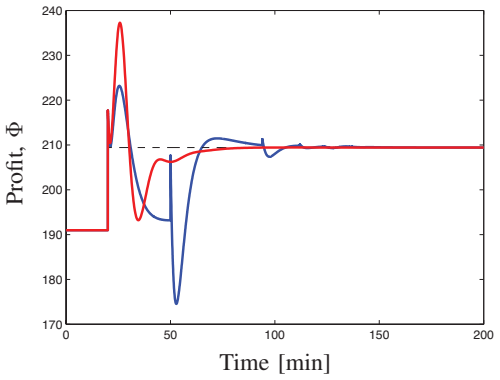

Fig. 2. Profit function. Dashed line: steady state optimal profit; Blue line: NEC; Red line: SOC.

(23) are computed as:

$$\mathsf{K}^y = \begin{bmatrix} 4.985 & -11.97 & -0.1595 \\ 13.13 & -31.53 & -0.4200 \end{bmatrix},$$

$$\mathsf{K}^u = \begin{bmatrix} 4.985 & -11.97 \\ 13.13 & -31.53 \end{bmatrix}.$$

The time responses of the decision variables F_B and T_R^{sp} are shown in Fig. 1 (blue lines). After the disturbance in F_A takes place at time $t = 20$ min, the initial response is delayed until the (near) steady state condition is satisfied. The steady state values reached upon convergence are very close to the plant optimum (dashed lines). The time response of the profit function is shown in Fig. 2 (blue line).

TABLE II
CONTROLLER TUNING - SOC

Loop	Type	Parameters
c_1-F_B	PI	$K_{pi} = 15$, $\tau^i_{pi} = 7$ (min)
c_2-T^{sp}_R	PI	$K_{pi} = -1000$, $\tau^i_{pi} = 7$ (min)

C. Self-Optimizing Control

The null space method using design A is applied here. In this case, the condition $n_y \geq n_u + n_d$ is satisfied and (25)-(26) provide the left null space of S^y as

$$\mathbf{N} = \begin{bmatrix} 0.1507 & 0.3961 \\ 0.9881 & -0.0313 \\ -0.0313 & 0.9177 \end{bmatrix}. \quad (41)$$

The above matrix allows selecting the set of controlled variables as linear combinations of the measured ones, $\mathbf{c} = \mathbf{N}^\mathsf{T}\mathbf{y}$. This way, the self-optimizing CVs are given by

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0.1507x_A + 0.9881x_B - 0.0313x_P \\ 0.3961x_A - 0.0313x_B + 0.9177x_P \end{bmatrix},$$

and they require to be controlled at fixed (nominal) setpoints via the decision variables $\mathbf{u} = [F_B\, T^{sp}_R]^\mathsf{T}$. Hence, this scenario defines a multivariable control problem with dimension 2×2. An useful tool to perform the input-output interaction analysis is the well-known "relative gain array (RGA)" [19]. Let us consider that $\mathbf{c} = \mathbf{G}\mathbf{u}$ at steady-state, then the RGA is given by $\Lambda = \mathbf{G} \otimes (\mathbf{G}^{-1})^\mathsf{T}$ with $\otimes$ the element-by-element product. In this case, these matrices are:

$$\mathbf{G} = \begin{bmatrix} 0.0752 & -0.0039 \\ -0.0109 & -0.0010 \end{bmatrix}, \; \Lambda = \begin{bmatrix} 0.6389 & 0.3611 \\ 0.3611 & 0.6389 \end{bmatrix},$$

indicating that a diagonal decentralized input-output pairing is the best choice, i.e. F_B for controlling c_1, and T^{sp}_R for c_2. On the other hand, the off-diagonal interaction in Λ is negligible, which suggests that a multivariable controller is not required here. Table II summarizes the tuning parameters for the suggested PI controllers.

The time responses of the decision variables are shown in Fig. 1 (red lines). The controller takes action immediately after the disturbance takes place. The steady state values reached are very close to the plant optimum. The time response of the profit function is shown in Fig. 2 (red line). In this implementation, the optimal steady state point is reached some minutes earlier with SOC than with NEC.

VI. DISCUSSION AND CONCLUSIONS

Both NEC with output feedback and SOC using the null space method are based on the first-order variations of the NCO. From an optimality point of view, they both represent first-order approximations to the classical two-step approach of RTO for the case when Assumptions 1 and 2 hold. The difference between both techniques is not in their optimality, but rather in their dynamic implementation. The advantage of NEC is that it uses a simple and direct control law. In contrast, it requires a steady-state identification subsystem, which represents a nontrivial design stage. On the other hand, SOC has the advantage of running at the process control layer, which may shorten the transition time between optimal steady-state operating points with respect to NEC. However, it requires solving a multivariable dynamic control problem. In other words, SOC requires the definition of: simplified dynamical models; pairing between manipulated and controlled variables; controller structure: full, sparse, or diagonal; controller technology: decentralized (e.g. via PIDs) or centralized (e.g. via MPC); and controller tuning.

Both techniques seem to be effective, and we believe the choice of one or the other will be dependent on the transient dynamical behavior of the process at hand.

REFERENCES

[1] T. E. Marlin and A. N. Hrymak, "Real-time operations optimization of continuous processes," in *AIChE Symposium Series - CPC-V*, vol. 93, 1997, pp. 156–164.

[2] M. L. Darby, M. Nikolaou, J. Jones, and D. Nicholson, "RTO: An overview and assessment of current practice," *J. Process Contr.*, vol. 21, pp. 874–884, 2011.

[3] S. Cao and R. R. Rhinehart, "An efficient method for on-line identification of steady state," *J. Process Contr.*, vol. 5, no. 6, pp. 363–374, 1995.

[4] S. A. Bhat and D. N. Saraf, "Steady-state identification, gross error detection, and data reconciliation for industrial process units," *Ind. Eng. Chem. Res.*, vol. 43, pp. 4323–4336, 2004.

[5] S.-S. Jang, B. Joseph, and H. Mukai, "On-line optimization of constrained multivariable chemical processes," *AIChE J.*, vol. 33, no. 1, pp. 26–35, 1987.

[6] B. Chachuat, B. Srinivasan, and D. Bonvin, "Adaptation strategies for real-time optimization," *Comp. Chem. Eng.*, vol. 33, pp. 1557–1567, 2009.

[7] S. Gros, B. Srinivasan, and D. Bonvin, "Optimizing control based on output feedback," *Comp. Chem. Eng.*, vol. 33, pp. 191–198, 2009.

[8] S. Skogestad, "Self-optimizing control: The missing link between steady-state optimization and control," *Comp. Chem. Eng.*, vol. 24, pp. 569–575, 2000.

[9] I. J. Halvorsen, S. Skogestad, J. C. Morud, and V. Alstad, "Optimal selection of controlled variables," *Ind. Eng. Chem. Res.*, vol. 42, pp. 3273–3284, 2003.

[10] V. Alstad and S. Skogestad, "Null space method for selecting optimal measurement combinations as controlled variables," *Ind. Eng. Chem. Res.*, vol. 46, pp. 846–853, 2007.

[11] B. Srinivasan, G. François, and D. Bonvin, "Comparison of gradient estimation methods for real-time optimization," in *Computer Aided Chemical Engineering*, vol. 29, 2011, pp. 607–611.

[12] G. François, B. Srinivasan, and D. Bonvin, "Comparison of six implicit real-time optimization schemes," *Journal Européen des Systèmes Automatisés*, vol. 46, no. 2-3, pp. 291–305, 2012.

[13] A. Maarleveld and J. E. Rijnsdorp, "Constraint control on distillation columns," *Automatica*, vol. 6, pp. 51–58, 1970.

[14] Y. Arkun and G. Stephanopoulos, "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.*, vol. 26, no. 6, pp. 975–991, 1980.

[15] C. E. Garcia and M. Morari, "Optimal operation of integrated processing systems. Part II: Closed-loop on-line optimizing control," *AIChE J.*, vol. 30, no. 2, pp. 226–234, 1984.

[16] G. François, B. Srinivasan, and D. Bonvin, "Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty," *J. Process Contr.*, vol. 15, no. 6, pp. 701–712, 2005.

[17] J. Jäschke and S. Skogestad, "NCO tracking and self-optimizing control in the context of real-time optimization," *J. Process Contr.*, vol. 21, pp. 1407–1416, 2011.

[18] T. J. Williams and R. E. Otto, "A generalized chemical processing model for the investigation of computer control," *AIEE Trans.*, vol. 79, p. 458, 1960.

[19] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control. Analysis and Design*, 2nd ed. John Wiley & Sons, 2005.