

Victor Manuel Cunha Alves

**Metamodel-based numerical techniques for  
Self-Optimizing Control**

Campina Grande, Paraíba, Brasil

2020



Victor Manuel Cunha Alves

**Metamodel-based numerical techniques for  
Self-Optimizing Control**

Dissertation submitted to the Chemical Engineering Graduate Program of the Federal University of Campina Grande as partial fulfillment of the requirements for the degree of Master of Chemical Engineering.

Universidade Federal de Campina Grande  
Unidade Acadêmica de Engenharia Química  
Programa de pós-graduação de Engenharia Química

Supervisor: Antônio Carlos Brandão de Araújo

Campina Grande, Paraíba, Brasil

2020

A732m Alves, Victor Manuel Cunha.  
Metamodel-based numerical techniques for self-optimizing control /  
Victor Manuel Cunha Alves. – Campina Grande, 2020.  
98 f. : il. color.

Dissertação (Mestrado em Engenharia Química) – Universidade  
Federal de Campina Grande, Centro de Ciências e Tecnologia, 2020.  
"Orientação: Prof. Dr. Antônio Carlos Brandão Araújo".  
Referências.

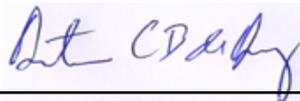
1. Engenharia Química. 2. Controle Auto-Otimizante. 3. Kriging.  
4. Modelos Substitutos. 5. Método Exato Local. 6. Método do Espaço  
Nulo. I. Araújo, Antônio Carlos Brandão. II. Título.

CDU 66.01(043)

Victor Manuel Cunha Alves

## **Metamodel-based numerical techniques for Self-Optimizing Control**

Dissertation submitted to the Chemical Engineering Graduate Program of the Federal University of Campina Grande as partial fulfillment of the requirements for the degree of Master of Chemical Engineering.



---

**Antônio Carlos Brandão de Araújo**  
Orientador



---

**Avaliador Interno**  
Heleno Bispo da Silva Júnior



---

**Avaliador Externo**  
Sidinei Kleber da Silva

Campina Grande, Paraíba, Brasil  
2020



*To my parents.*



# Acknowledgements

Firstly, I would like to thank Professor Antônio Araújo for his supervision over the last four years and also the support provided during the execution of this study. His suggestions and supervision indisputably helped the execution of this dissertation, and my formation as an engineer and academic, generally speaking.

To my dearest friends Felipe Lima and Paloma Lins, for all of the support and partnership over the last 8 years. One could not ask for better, most honest and true friends than both of you. It was a pleasure to have you by my side giving support and (a lot of) good laughs.

I would also like to thank in special my family for all support and love over my entire life - Ana, Beatriz and Jailson. Without them, I would never be able to write these words today.

To Luiza, I would like to express my endless gratitude and my most sincere acknowledgment for all the love, patience and support that she has always been providing me.



*“Mystery creates wonder and wonder is the basis of man’s desire to understand.”*  
*(Neil Armstrong)*



# Resumo

A tecnologia de controle auto-otimizante (*Self-optimizing control*) é um campo de estudo bem-conhecido da grande área de seleção de estruturas de controle, tendo uma robusta fundamentação matemática. Com o auxílio de simuladores de processo comerciais e pacotes numéricos, a modelagem de processos tornou-se uma tarefa mais fácil. Entretanto, abordar sistemas complexos ainda é uma tarefa tediosa, ou até mesmo impraticável, mesmo com as ferramentas inovadoras supracitadas. Modelos substitutos, também chamados metamodelos, podem ser usados para substituir parcial ou totalmente os modelos originais, para fins de predição e otimização, reduzindo a complexidade da avaliação de processos de larga escala e altamente não-lineares. Este trabalho tem como objetivo a aplicação de técnicas recentes de control auto-otimizante à superfícies de resposta (metamodelos) utilizando o kriging como técnica de construção dos metamodelos. Um procedimento para aplicação de controle auto-otimizante à modelos substitutos é descrito em detalhes, junto com como a otimização pode ser efetuada. Estudos de caso conhecidos da literatura tiveram metamodelos construídos e estes foram analisados para gerar, utilizando as técnicas citadas, estruturas de controle ótimas que minimizam a pior-perda, e os mesmos resultados foram encontrados se comparados com a implementação utilizando os modelos originais de autores anteriores. Os resultados indicam a eficácia dos modelos substitutos quando aplicados ao design de estruturas de controle auto-otimizantes, simplificando toda a metodologia.

**Palavras-chave:** Controle auto-otimizante. Kriging. Modelos substitutos. Método exato local. Método do espaço nulo.



# Abstract

Self-optimizing control technologies are a well-known study field of control structure design, having a robust mathematical background. With the aid of commercial process simulators and numerical packages, process modelling became an easier task. However, dealing with extremely large and complex systems still is a tedious task, and sometimes not feasible, even with these innovative tools. Surrogate models, also called metamodels, can be used to substitute partially or totally the original mathematical models for prediction and optimization purposes, reducing the complexity of evaluating large-scale and highly non-linear processes. This work aims at applying recent self-optimizing control techniques to surface responses of processes using kriging method as reduced model builder. A procedure to apply Self-Optimizing control to surrogate responses was described in detail, together with how the optimization can be done. Well-known case studies had their surface responses successfully built and analyzed to generate using the techniques cited, the optimal selection of controlled variables that minimizes the worst-case loss, and the same results were found when compared with the implementation in the original models from previous authors. The results indicate the effectiveness of the reduced models when applied to design self-optimizing control structures, simplifying the task.

**Keywords:** Self-optimizing control. Kriging. Surrogate modelling. Exact local method. Null space method.



# List of Figures

Figure 1 – Flowchart for the algorithm developed. . . . .	43
Figure 2 – Depropanizer column. . . . .	54
Figure 3 – Kriging hold-out validation for the depropanizer case study. . . . .	56
Figure 4 – Temperature profile for depropanizer column. . . . .	61
Figure 5 – Valve opening and closing (5% amplitude) - Depropanizer case study, control structure 1. . . . .	64
Figure 6 – 2.5% increase in vapor fraction - Depropanizer case study, control structure 1. . . . .	64
Figure 7 – $\pm 5\%$ disturbance of propene feed molar fraction. - Depropanizer case study, control structure 1. . . . .	65
Figure 8 – Valve opening and closing (5% amplitude) - Depropanizer case study, control structure 2. . . . .	65
Figure 9 – 2.5% increase in vapor fraction - Depropanizer case study, control structure 2. . . . .	66
Figure 10 – $\pm 5\%$ disturbance of propene feed molar fraction. - Depropanizer case study, control structure 2. . . . .	66
Figure 11 – Evaporation system flowsheet. . . . .	67
Figure 12 – $\pm 5\%$ disturbance in feed molar fraction (5.25% and 4.75%, respectively) - Evaporation system case study. . . . .	72
Figure 13 – $\pm 20\%$ disturbance in feed temperature. (48°C and 32°C, respectively) - Evaporation system case study. . . . .	73
Figure 14 – $\pm 20\%$ disturbance in steam temperature (26.25°C and 23.75°C, respec- tively). - Evaporation system case study. . . . .	73
Figure 15 – Cumene process production flowsheet. . . . .	74



# List of Tables

Table 1 – Steady-State Degrees of Freedom for Main Process Units. . . . .	34
Table 2 – Brute-force infill criteria - Depropanizer case study. . . . .	56
Table 3 – Optimization Results of Surrogates and Original Model. . . . .	56
Table 4 – Single-measurement policy: Results for worst and average-case losses for depropanizer case study. . . . .	58
Table 5 – Loss evaluation: linear combination of measurements of different subsets' sizes for depropanizer case study. . . . .	59
Table 6 – Loss evaluation: sub-optimal solution with extended null space method for depropanizer case study. . . . .	60
Table 7 – Temperatures Single measurements as CV candidates: Worst and average-case losses for depropanizer case study. . . . .	62
Table 8 – Linear Combinations of Stage Temperatures - Depropanizer case study. . . . .	63
Table 9 – Optimization Results of Surrogate responses and using the Original nonlinear Model (Implemented in MATLAB) for Comparison Purposes - Evaporation system case study. . . . .	68
Table 10 – Decision/State Variables and its optimal values: Original implementation and kriging response - Evaporation system case study. . . . .	69
Table 11 – Matrix Mean-Squared error between gradients and hesssians obtained by Kariwala, Cao, and Janardhanan (2008) and this work - Evaporation system case study. . . . .	70
Table 12 – Linear combinations of measurements as CV candidates for Self-Optimizing control and its losses - Order sequence comparison: Kariwala, Cao, and Janardhanan (2008) <i>vs.</i> this study - Evaporation system case study. . . . .	72
Table 13 – Cumene process production optimization problem summary. . . . .	75
Table 14 – Optimization Results of Surrogates and Original Model - Cumene process production case study. . . . .	76
Table 15 – Cumene process production CV Candidates. . . . .	77
Table 16 – High-order data obtainment: <i>Aspen Plus vs Kriging</i> - Cumene Process case study. . . . .	78
Table 17 – Matrix Mean-Squared error between gradients and hesssians obtained using the process simulator and metamodels - Cumene process production case study. . . . .	78
Table 18 – Worst and Average-case losses evaluation - Cumene Process Production Case Study. . . . .	79



# List of abbreviations and acronyms

BLUP Best linear unbiased predictor

CV Controlled Variable

DACE Design and Analysis of Computer Experiments

DIPB Di-isopropyl benzene

DOF Degree of Freedom

LHS Latin Hypercube Sampling

MSV Minimum Singular Value

NLP Nonlinear programming

RSM Response surface methodology

RTO Real-time optimization



# List of symbols

## Subsection 2.1

$\mathcal{N}$	Null space
$\sigma$	Singular value
$c$	Selected controlled variables
$d$	Process disturbance
$g$	Process constraint (equality/inequality)
$G^y$	Gain matrix with respect to the measurements
$G_d^y$	Gain matrix with respect to the disturbances
$H$	Optimal measurement matrix
$J_0$	Cost Function
$J_0$	Cost Function as function of disturbances/measurements at given point
$J_{opt}$	Optimized Cost Function
$J_{ud}$	Hessian matrix with respect to the manipulated variables and disturbances
$J_{uu}$	Hessian matrix with respect to the manipulated variables
$L$	Loss
$n_d$	Number of disturbances
$n_y$	Number of measurements
$u$	Degrees of freedom for reduced-space problem (remaining “unconstrained” variables)
$u_0$	Original independent variables (“constrained” and “unconstrained” variables)
$W_d$	Diagonal matrix of the disturbances magnitudes
$W_n^y$	Diagonal matrix of the disturbances magnitudes
$x$	Process state
$y$	Process measurement

## Subsection 2.2

$\beta$	Regression parameter matrix
---------	-----------------------------

$\hat{y}''$	Predictor second derivative
$\hat{y}'$	Predictor first derivative
$\hat{y}_l(x)$	Metamodel of y
$\mathcal{F}$	Regression model function
$\mathcal{R}$	Spatial correlation function
$\sigma_t^2$	Process variance
$\theta$	Hyperparameter value (activity)
$H_f$	Correlation model Hessian
$H_r$	Regression model Hessian
$J_f$	Correlation model Jacobian
$J_r$	Regression model Jacobian
$l$	$l^{th}$ dimension
$m$	Number of experiments
$n$	Dimension of input x
$p$	Smoothness parameter (Hyperparameter of $\mathcal{R}$ )
$R$	Spatial correlation function
$s$	Element of S
$y(x)$	Deterministic function
$z_l(x)$	Stochastic process

# Contents

<b>I</b>	<b>INTRODUCTION</b>	<b>25</b>
<b>1</b>	<b>INTRODUCTION</b> . . . . .	<b>27</b>
<b>1.1</b>	<b>Publications</b> . . . . .	<b>29</b>
<b>II</b>	<b>METHODOLOGY</b>	<b>31</b>
<b>2</b>	<b>METHODOLOGY</b> . . . . .	<b>33</b>
<b>2.1</b>	<b>Self-Optimizing Control Technology: State-of-the-art</b> . . . . .	<b>33</b>
2.1.1	Degrees of Freedom Analysis . . . . .	34
2.1.2	Selection of Controlled Variables: Main Methods . . . . .	35
2.1.2.1	Null Space, Extended Null-Space, and Exact Local Method with Explicit Solution Methods. . . . .	37
2.1.3	Branch-and-Bound Methods for Subset Selection . . . . .	39
<b>2.2</b>	<b>Response Surface Methodology</b> . . . . .	<b>41</b>
2.2.1	Aspects of Kriging . . . . .	41
2.2.2	The Optimization Algorithm . . . . .	42
2.2.3	Usage of Jacobian and Hessians evaluated by Kriging equations . . . . .	46
<b>III</b>	<b>CASE STUDIES</b>	<b>49</b>
<b>3</b>	<b>INTRODUCTION</b> . . . . .	<b>51</b>
<b>4</b>	<b>CASE STUDIES</b> . . . . .	<b>53</b>
<b>4.1</b>	<b>Case I - Economic Self-Optimizing Control of a Depropanizer column</b> <b>53</b>	
4.1.1	Problem description . . . . .	53
4.1.2	Surrogate Model Generation . . . . .	54
4.1.2.1	Initial Sample . . . . .	55
4.1.3	Optimization of the Surrogate Response . . . . .	55
4.1.4	Reduced Space Surrogate Model. . . . .	57
4.1.5	High-Order Data Obtainment . . . . .	57
4.1.6	Expected Disturbances and Implementation Errors . . . . .	58
4.1.7	Loss Evaluation . . . . .	58
4.1.8	Dynamic Simulations . . . . .	63
<b>4.2</b>	<b>Case II - Economic Self-Optimizing Control of a Evaporation process</b> <b>67</b>	
4.2.1	Problem Description . . . . .	67

4.2.2	Surrogate model generation and optimization . . . . .	68
4.2.3	Reduced Space surrogate model . . . . .	68
4.2.4	High-order Data Obtainment . . . . .	69
4.2.5	Expected Disturbances and Implementation Errors . . . . .	71
4.2.6	Loss evaluation . . . . .	71
4.2.7	Dynamic simulations . . . . .	72
<b>4.3</b>	<b>Case III - Economic Self-Optimizing Control of Cumene Process production . . . . .</b>	<b>73</b>
4.3.1	Problem Description . . . . .	73
4.3.2	Surrogate model generation and optimization . . . . .	75
4.3.3	Reduced Space Surrogate Model . . . . .	76
4.3.4	High-order data obtainment . . . . .	78
4.3.5	Expected Disturbances and Implementation Errors . . . . .	78
4.3.6	Loss Evaluation . . . . .	79
<b>IV</b>	<b>CONCLUSIONS</b>	<b>81</b>
<b>5</b>	<b>CONCLUSIONS . . . . .</b>	<b>83</b>
	<b>References . . . . .</b>	<b>85</b>
	<b>APPENDIX</b>	<b>93</b>
	<b>APPENDIX A – KRIGING PROOFS: PREDICTOR, GRADIENT, AND HESSIAN . . . . .</b>	<b>95</b>

# Part I

## Introduction



# 1 Introduction

One of most challenging activities for process engineers in an industrial environment is to operate processes optimally (or as near as possible to optimality) from an economical point of view, ensuring compliance with safety and environmental regulations. As a matter of fact, many methods in process control area claim to ensure optimal operation, such as Real Time Optimization (RTO) based-methods for instance. However, these methods fail in having a practical implementation, by becoming too complex to be applied in real-life systems. Differently from RTO-based approaches, Self-Optimizing control considers a constant setpoint policy. As result of this, there is an intrinsic (positive) loss that must be acceptable to the designer of the control structure and plant operators. Thus, a trade-off relationship between the loss and simpler implementation can be found. The advantage of using Self-Optimizing control relies on the lack of necessity in reoptimizing the process every time that there is a disturbance in the system and the technology guarantees that the best set of controlled variables is chosen among all possible, minimizing the loss in all disturbance region. However, the loss is not an exclusive result of the disturbances, but also due to implementation errors in the control structure and measurements errors. Therefore, the best candidates set should also be chosen considering these.

Several contributions over the years were made to the technology, and the very first ideas regarding Self-Optimizing control must be credited to [Morari, Arkun, and Stephanopoulos \(1980\)](#), who wrote:

“in attempting to synthesize a feedback optimizing control structure, our main objective is to translate the economic objectives into process control objectives. In other words, *we want to find a function  $c$  of the process variables which when held constant, leads automatically to the optimal adjustments of the manipulated variables, and with it, the optimal operating conditions. [. . .]*”

From this idea, [Skogestad \(2000\)](#) developed a systematical procedure to find the best controlled variables for Self-Optimizing control, evaluating the loss directly from the model of the process, denominated by [Umar et al. \(2012\)](#) as a “brute-force” approach. From this work to the late 00’s there was a significant amount of contributions. [Halvorsen et al. \(2003\)](#) made their contribution based on local methods, linearizing the objective function around the nominal optimal point, allowing to evaluate the worst-case loss using the minimum singular value rule and the exact local method as criterions to choose the best subsets of controlled variables. Later, [Hori, Skogestad, and Alstad \(2005\)](#) based on the same assumptions, derived an expression for the case of “perfect indirect control”, in attempt to find the best set of measurements to be used as controlled variables to

indirectly control primary variables. [Alstad and Skogestad \(2007\)](#) introduced the null space method to the case where the objective is to select the optimal measurement matrix  $H$  prioritizing the disturbance rejection and not considering the loss due to the implementation errors, for a system with the number of measurements being equal to the number of degrees of freedom and disturbances summed, on the reduced-space problem. [Alstad, Skogestad, and Hori \(2009\)](#) extended the null space method for the case when there are extra measurements and not enough measurements available and derived an explicit expression for the exact local method, guaranteeing the evaluation of the optimal matrix  $H$  that is a linear combination of the available measurements, that minimizes the worst-case loss generated by the disturbances and the implementation errors simultaneously. More recently, [Jacobsen and Skogestad \(2011\)](#) explored methods to determine regions where active constraints become inactive and [Jäschke and Skogestad \(2012\)](#) derived a method to find optimal controlled variables that are polynomial combinations of measurements. The last two contributions were not considered in this work, and only optimal selection using linear combinations of measurements using surrogate responses from the original model were considered.

The contributions to the Self-Optimizing Control technology cited above, that indeed are efficient, were exhaustively tested and proven to work. ([ARAÚJO; SHANG, 2009](#); [ARAÚJO; SKOGESTAD, 2008](#); [JENSEN; SKOGESTAD, 2007](#); [HORI, E.; SKOGESTAD, S., 2007b,a](#); [LERSBAMRUNGSUK et al., 2008](#); [JAGTAP; KAISTHA; SKOGESTAD, 2011](#); [PANAHI; SKOGESTAD, 2012](#); [GERA; KAISTHA, et al., 2011](#); [JÄSCHKE; SKOGESTAD, 2014](#); [KHANAM; SHAMSUZZOHA; SKOGESTAD, 2014](#)). However, at complex systems, the implementation of the methods derived can become a challenging task and even not feasible in some cases. For instance, in a large simulation of a chemical process in a process simulator, evaluation of Hessians and gradients can become inaccurate, jeopardizing the optimal selection of linear combination of measurements and even giving the wrong order of candidates, generating an erroneous analysis. For this reason, the use of surrogate models to simplify the evaluation of such variables can become a powerful tool to find the best subset of controlled variables when designing a control structure. Another advantage of using surrogate models (also called metamodels) in Self-Optimizing control is the evaluation time that may be drastically reduced.

Seeing that we are dealing with a technology that relies heavily on gradients and Hessians, it is important that the calculation of these to be precise, or at least satisfactory enough. However, when dealing with “black-box”<sup>1</sup> models such as process simulators, this task of estimation is seldom simple or without hassle (e.g. how small a perturbation step

---

<sup>1</sup> We follow the definition of a “black-box” model as the same of [Caballero and Grossmann \(2008\)](#), [Forrester, Sobester, and Keane \(2008\)](#) and [Quirante, Javaloyes, and Caballero \(2015\)](#) where it is a model which a user has limited access of its “innards”. In this case, the model may be a commercial process modeler such as Aspen Plus, HYSYS, PROII, etc.

should be in order to get reliable results for a hessian or gradient?), particularly when dealing with separation process such as distillation columns. Therefore, it might be possible to circumvent such challenges using surrogate models to get estimates of Hessians/gradients. In addition, the optimization of these “black-box” models can be exploited with the use of metamodels when dealing with complex cases where the original model can’t be optimized directly, albeit used as source of design experiments. With the data obtained from these experiments, we can build a simpler model (surrogate) and perform the optimization on it. (JONES; SCHONLAU; WELCH, 1998; KUSHNER, 1964; SACKS et al., 1989; JONES, 2001)

This work seeks to link the well-known and developed theory of Self-Optimizing Control to the engineering design using surrogate models, specially using kriging interpolators, as a way of simplifying the determination of the optimal linear combination of measurements, in order to generate a control structure that minimizes the loss resulted from disturbances and implementation errors.

## 1.1 Publications

The proposed methodology of this dissertation and two of the case studies described in this work generated the following paper:

ALVES, Victor M. C. et al. Metamodel-Based Numerical Techniques for Self-Optimizing Control. **Industrial & Engineering Chemistry Research**, v. 57, n. 49, p. 16817–16840, 2018. DOI: [10.1021/acs.iecr.8b04337](https://doi.org/10.1021/acs.iecr.8b04337). eprint: <https://doi.org/10.1021/acs.iecr.8b04337>. Available from: <<https://doi.org/10.1021/acs.iecr.8b04337>>



## Part II

### Methodology



## 2 Methodology

### 2.1 Self-Optimizing Control Technology: State-of-the-art

A brief analysis of the self-optimizing control technology will be done in this work in order to keep the reader up to date. For an extensive and exhaustive review of the subject, refer mainly to [Skogestad \(2000, 2004\)](#) and [Umar et al. \(2012\)](#).

Self-optimizing control consists in achieving a control structure based on a constant set point policy that leads to near-optimal operation. From [Skogestad \(2004\)](#):

***“Self-optimizing control is when one can achieve an acceptable loss with constant set point values for the controlled variables without the need to re-optimize when disturbances occur.”***

The process economics are assumed to be only influenced by its steady-state operation. Ergo, a steady-state model of the process studied is used to evaluate the selection of the self optimizing control structures (when it is possible to use a steady-state model, e.g., continuous process).

The optimal operation it is assumed to be quantified using a scalar cost function that needs to be minimized with respect to the degrees of freedom available,  $u_0$ :

$$\min_{u_0} J_0(x, u_0, d) \quad (2.1)$$

subject to the (often highly nonlinear) constraints:

$$g_1(x, u_0, d) = 0 \quad (2.2)$$

$$g_2(x, u_0, d) \leq 0 \quad (2.3)$$

where  $x$  and  $d$  represent the states and the disturbances of the system, respectively. The latter can have an exogenous nature (e.g., change in feed conditions), changes in specifications (constraints), parameters of the cost function (prices), and changes in the model (typically changes in  $g_1$ ). Using numerical optimization, the cost function can be evaluated directly with the model, for the expected disturbances and implementation errors. According to [Skogestad \(2000\)](#) the main steps are:

1. Definition of the number of degrees of freedom.

2. Definition of optimal operation (Cost function and constraints).
3. Identification of important disturbances.
4. Numerical optimization of the problem.
5. Identification of variables candidates to be controlled.
6. Evaluation of loss for alternative combinations of controlled variables (loss imposed by keeping constant setpoints when there are disturbances or implementation errors), including feasibility investigation.
7. Final evaluation and selection (controllability analysis).

The loss, intrinsic to the technology, the result of a constant set point-based policy, is given as

$$L = J_0(d, n) - J_{\text{opt}}(d) \quad (2.4)$$

### 2.1.1 Degrees of Freedom Analysis

To determine the degrees of freedom in self-optimizing control technology, it is the first, and mandatory, step to ensure the success of its implementation and analysis. In fact, several methodologies and approaches to sum up these degrees were developed during the years (DIXON, 1972; PHAM, 1994; PONTON, 1994; LUYBEN, 1996; SKOGESTAD, 2002; STEPHANOPOULOS, 2003; KONDA; RANGAIAH; KRISHNASWAMY, 2006; AL., 2016) and the reader should refer to those for a deep understating on the subject. Considering individual units, Table 1 yields, from Skogestad (2002)

Table 1 – Steady-State Degrees of Freedom for Main Process Units.

Process Unit	DOF
Each external feed stream	1 (feedrate)
Splitter	n - 1 split fractions (n=number of exit streams)
Mixer	0
Compressor, turbine, and pump	1 (work)
Adiabatic flash tank	0 <sup>a</sup>
Liquid phase reactor	1 (holdup)
Gas phase reactor	0 <sup>a</sup>
Heat exchanger	1 (duty or net area)
Columns (e.g., distillation) excluding heat exchangers	0 <sup>a</sup> + number of side streams
a = Add 1 degree of freedom for each extra pressure that is set.	

### 2.1.2 Selection of Controlled Variables: Main Methods

Among the methods used in self-optimizing control technology, the most used ones that can be cited are the “brute-force” approach and the local methods, with the latter having different approaches, such as the minimum singular value rule, the exact local method, and the null-space and extended nullspace methods. The “brute-force” approach is the earliest method to evaluate CVs with the self-optimizing control technology, described mainly by Skogestad (2000). The loss is evaluated by using the process model directly, becoming in complex cases a tedious and even infeasible task because the optimization problem described in Equation 2.1-Equation 2.3 can be a large nonconvex problem, as stated by Umar et al. (2012).

To overcome the intrinsic complexity of evaluating the loss directly on the process model, local methods were developed based on the quadratic approximation of the loss function (UMAR et al., 2012). Because of the simplicity, these methods are an excellent way to prescreen the best sets of CVs or its linear combinations. One drawback of this methodology is the assumption that the set of active constraints does not change with respect to the disturbance allowable region. The model can be linearized with respect to the measurements as follows:

$$\Delta y = G^y \Delta u + G_d^y \Delta d \quad (2.5)$$

with

$$\Delta y = y - y^* \quad (2.6)$$

$$\Delta u = u - u^* \quad (2.7)$$

$$\Delta d = d - d^* \quad (2.8)$$

and  $G^y$ ,  $G_d^y$  being the gain matrices with respect to the measurements and disturbances. For the CVs, linearization yields

$$\Delta c = H \Delta y = G \Delta u + G_d \Delta d \quad (2.9)$$

with

$$G = H G^y \quad (2.10)$$

$$G_d = HG_d^y \quad (2.11)$$

Linearization of the loss function yields

$$L = J(u, d) - J_{\text{opt}}(d) = \frac{1}{2} \|z\|_2^2 \quad (2.12)$$

$$z = J_{uu}^{1/2} (u - u_{\text{opt}}) = J_{uu}^{1/2} G^{-1} (c - c_{\text{opt}}) \quad (2.13)$$

From [Equation 2.11](#), the importance of the evaluation of H can be elucidated. This matrix can be simply a selection index (consisting of elements being exclusively 1 and 0) or a linear combination of measurements. If the first case is considered, individual measurements will be selected to form the control structure, and in the latter, the elements inside H are arbitrary, generating CVs as linear combinations of available measurements.

One local method, developed initially by [Skogestad and Postlethwaite \(1996\)](#) and later formally discussed and derived by [Halvorsen et al. \(2003\)](#) namely the minimum singular value (MSV) rule, considered the maximization of the minimum singular values of the system scaled gain matrix as a metric of selecting subsets that have good self-optimizing control properties. Later, to overcome limitations involving the MSV rule regarding the assumption of independency in set points shown in some examples ([HORI; SKOGESTAD; ALSTAD, 2005; HORI, E. S.; SKOGESTAD, Sigurd, 2008](#)) the exact local method developed by [Halvorsen et al. \(2003\)](#) was proposed. Around the optimal operating point, it was shown that the loss is given as

$$z = J_{uu}^{1/2} \left[ \left( J_{uu}^{-1} J_{ud} - G^{-1} G_d \right) \Delta d + G^{-1} n \right] \quad (2.14)$$

With  $J_{uu}$  and  $J_{ud}$  being the Hessians with respect to the manipulated variables and both manipulated variables and disturbances, respectively. Assuming that  $W_d$  represents the diagonal matrix of the disturbances,  $W_n^y$  the magnitude of the implementation error, and the combined disturbances and implementation errors being 2-norm-bounded (with a discussion and justification for the latter at [Halvorsen et al. \(2003\)](#)):

$$d - d^* = W_d d' \quad (2.15)$$

$$n = HW_n^y n^{y'} = W_n n^y \quad (2.16)$$

$$\left\| \begin{pmatrix} d' \\ n^{y'} \end{pmatrix} \right\| \leq 1 \quad (2.17)$$

Considering these, the worst-case loss is given by [Equation 2.17](#)

Thus, to minimize the worst-case loss with respect to the combined effect of the disturbances and implementation errors is equivalent to maximize the singular values of  $M$  ( $\sigma(M)$ ).

$$\left\| \begin{pmatrix} d' \\ n' \end{pmatrix} \right\|_2 \leq 1 \quad L_{worst} = \frac{\bar{\sigma}(M)^2}{2} \quad (2.18)$$

The average-case loss was derived by [Kariwala, Cao, and Janardhanan \(2008\)](#) with the justification of being more realistic to analyze if compared to the worst-case one, which is clearly a conservative approach, [Equation 2.19](#)

$$L_{average} = \frac{1}{6(n_y + n_d)} \left\| J_{uu}^{1/2} (HG^y)^{-1} H\tilde{F} \right\|_F^2 \quad (2.19)$$

The definitions of [Equation 2.20-Equation 2.22](#) are introduced, and their derivation are discussed in [Halvorsen et al. \(2003\)](#) and [Alstad, Skogestad, and Hori \(2009\)](#): They correspond to a more appropriate way to represent the uncertainty variables regarding the contribution of the disturbances and measurement errors on the incurred loss:

$$M = [M_d M_n^y] \quad (2.20)$$

$$M_d = -J_{uu}^{1/2} (HG^y)^{-1} HFW_d \quad (2.21)$$

$$M_n^y = -J_{uu}^{1/2} (HG^y)^{-1} HW_{n^y} \quad (2.22)$$

$$\tilde{F} = [FW_d W_n^y]; \quad F = G_d^y - G^y J_{uu}^{-1} J_{ud} \quad (2.23)$$

depends on the evaluation of the optimal measurement matrix (H). Therefore, the evaluation of this matrix is mandatory to the success of the implementation of the mathematical formulations described so far.

### 2.1.2.1 Null Space, Extended Null-Space, and Exact Local Method with Explicit Solution Methods.

[Alstad and Skogestad \(2007\)](#) derived the null space method, and [Alstad, Skogestad, and Hori \(2009\)](#) presented the extended null-space method. Both methods have the approach of minimizing the loss prioritizing the effect of disturbances first, and if there

are any degrees of freedom left, to minimize the loss with respect to the measurement errors. Considering the definition of  $F$  as being the optimal sensitivity matrix, the optimal matrix  $H$  that minimizes the loss (neglecting the effect of the implementation errors and considering only the disturbances effect) and generates a set of linear combinations of  $y$  is given through the selection of  $H$  at the left null space of  $F$ . Mathematically:

$$H \in \mathcal{N}(F^T) \quad (2.24)$$

$$HF = 0 \quad (2.25)$$

However, evaluating  $H$  using the original null space method forces the designer of the control structure to have a number of measurements that is at least equal to the number of disturbances and degrees of freedom (one limitation imposed by [Alstad and Skogestad \(2007\)](#) to make them able to derive [Equation 2.25](#)). If this number is greater, which criteria should be used to select the best set to square the system? If it is lower, which expression may be used to evaluate  $H$ ? To answer the first question, efficient algorithms called “Branch-and-Bound” were developed and will be discussed later in this work. For the second, [Alstad, Skogestad, and Hori \(2009\)](#) improved the evaluation of  $H$  by extending the null space method, deriving an explicit expression to evaluate  $H$  at any dimension with respect to the number of measurements, disturbances, and degrees of freedom of the unconstrained problem. Another useful contribution was to derive an explicit expression for  $H$  for the exact local method, which minimizes the loss with respect to disturbances and implementation errors simultaneously. For the extended null space and exact local method,  $H$  evaluations are given as, respectively:

$$H = M_n^{-1} \tilde{J} \left( W_n^{y-1} \tilde{G}^y \right)^\dagger W_n^{y-1} \quad (2.26)$$

$$H = \left( \tilde{F} \tilde{F}^T \right)^{-1} G_y \left( G_y^T \left( \tilde{F} \tilde{F}^T \right)^{-1} G_y \right)^{-1} J_{uu}^{1/2} \quad (2.27)$$

[Equation 2.26](#) is enough to deal with any dimension of the problem purposed for the extended null space method, and [Equation 2.27](#) to the exact local method. However, [Alstad, Skogestad, and Hori \(2009\)](#) made an extensive analysis to cases where the number of measurements available is greater than the sum of the number of disturbances and degrees of freedom and vice versa. For the first case, [Equation 2.26](#) can be simplified to

$$H = M_n^{-1} J \left( \tilde{G}^y \right)^{-1} \quad (2.28)$$

where  $\tilde{G}^y = \begin{bmatrix} G_y & G_d^y \end{bmatrix}$  corresponds to the augmented plant and  $\tilde{J} = \begin{bmatrix} J_{uu}^{1/2} & J_{uu}^{1/2} & J_{uu}^{-1} & J_{ud} \end{bmatrix}$  by definition (see [Alstad, Skogestad, and Hori \(2009\)](#) for further details). For the case

when there are “too few” measurements ( $n^y < n^u + n^d$ ) the following expression for H is derived from Equation 2.26:

$$H = M_n^{-1} \tilde{J} (\tilde{G}^y)^\dagger \quad (2.29)$$

### 2.1.3 Branch-and-Bound Methods for Subset Selection

It is common that in large processes, the number of measurements available is greater than the number of unconstrained degrees of freedom. As stated before, to select the best subset with a criterion was an unsolved problem, at least in a practical way. Considering for example, a modest process plant with 10 unconstrained degrees of freedom and 50 measurements available, there are

$$\binom{50}{10} = \frac{50!}{10!40!} = 1.027 \times 10^{10} \quad (2.30)$$

possible control structures. Therefore, an analysis of all of them to select the best (and with which criteria?) would be impractical. Luckily, methods of variable selection have been developed and are quite successful in overcoming this problem. Branchand- Bound algorithms are one of those and were used in this work. The very first ideas regarding this category of methods, are present in the work of Lawler and Wood (1966), being improved later by Narendra and Fukunaga (1977). Saha and Cao (2003), Cao and Saha (2005), Cao and Kariwala (2008) and Kariwala and Cao (2009) enhanced the method, guaranteeing faster runtime and fewer iterations. To understand how the algorithm works, suppose that  $X_s = x_1, x_2, \dots, x_s$  is a set of S elements and a subset  $X_n$  with  $n$  elements are selected from  $X_s$  ( $X_n \subset X_s$ ). Therefore, there are

$$\frac{S!}{n!(S-n)!} \quad (2.31)$$

ways of selecting  $X_n$  subsets from  $X_s$ . Considering  $\Gamma$  as the criterion function used during the selection procedure, there is a subset with  $n$  elements ( $X_n^*$ ) that satisfies the equality:

$$\Gamma(X_n^*) = \max_{X_n \subset X_s} \Gamma(X_n) \quad (2.32)$$

which is a mathematical way of saying that  $X_n^*$  is the best subset that attends the criterion  $\Gamma$  among all possible  $X_n$  subsets. It is also clear that the criterion must attend a monotonic property, given by

$$\Gamma(X_n) \leq \Gamma(X_s), \text{ if } X_n \subseteq X_s \quad (2.33)$$

With this property, any subset of a given dimension cannot be better than any larger set containing the first one, as observed by [Saha and Cao \(2003\)](#).

Ordinary branch-and-bound algorithms are unidirectional, with the subsets being sliced one by one until reaching the target size or being appended until reaching the desired size. Considering a descending branch-and-bound algorithm, supposing  $m > n$  and  $\Gamma_n(X_m)$  as an upper limit on  $\Gamma$  descending on all subsets of  $m$  elements,  $X_m$ :

$$\Gamma_n(X_m) \geq \max_{X_n \subseteq X_m} \Gamma(X_n) \quad (2.34)$$

and defining  $B$  as the lower limit of  $\Gamma(X_n^*)$

$$B \leq \Gamma(X_n^*) \quad (2.35)$$

the following is true:

$$\text{if } \bar{\Gamma}_n(X_m) < B, \Gamma(X_n) < \Gamma(X_n^*) \forall X_n \subseteq X_m \quad (2.36)$$

which implies that no subset from  $X_m$  can be optimal, allowing the exclusion of them immediately, this exclusion is referred as ‘‘pruning’’. In the reverse order (upward search), an analogous analysis can be done, as shown by [Cao and Kariwala \(2008\)](#) Assuming  $B$  as the lower limit of  $\Gamma(X_n^*)$  exactly as before, but considering  $\bar{\Gamma}_n$  (with  $m < n$ ) as an upper limit upward, the following is also true:

$$\Gamma_n(X_m) \geq \max_{X_n \supseteq X_m} \Gamma(X_n) \quad (2.37)$$

therefore,

$$\text{if } \bar{\Gamma}_n(X_m) < B, \Gamma(X_n) < \Gamma(X_n^*) \forall X_n \supseteq X_m \quad (2.38)$$

Guaranteeing that no sets upward of  $X_m$  are globally optimal. As a result, all subsets on this direction including  $X_m$  can be ‘‘pruned’’ immediately. In this work, similarly as in [Silva et al. \(2017\)](#) was used the bidirectional branch-and-bound proposed by [Cao and Kariwala \(2008\)](#) The search for the best sets of controlled variables is carried out in both directions, which generates a higher efficiency if compared with unidirectional branch-and-bound algorithms. The criterion function for the branch-and-bound used was the minimum singular value, which fortunately has a monotonic property. For a detailed proof of it, refer to [Silva et al. \(2017\)](#) and [Cao, Rossiter, and Owens \(1998\)](#).

## 2.2 Response Surface Methodology

### 2.2.1 Aspects of Kriging

The brief surface response methodology (RSM) description given here can be seen in depth from the famous article published by [Sacks et al. \(1989\)](#) and from the widely known DACE toolbox developed by [Lophaven, Nielsen, and Søndergaard \(2002\)](#). However, for a more straightforward and intuitive explanation of the RSM, specially using the Kriging as a spatial correlation function (SCF), the reader is referred to the works of [Forrester, Sobester, and Keane \(2008\)](#) and [Jones, Schonlau, and Welch \(1998\)](#). For historical purposes, the reader can refer to the work of [Kriging \(1951\)](#) the original author of the Kriging method (which was named after him). The reason we chose to present the RSM in the formalism of [Sacks et al. \(1989\)](#) and [Lophaven, Nielsen, and Søndergaard \(2002\)](#) is because it allows us to easily develop relations for the gradient and Hessian of the kriging predictor.

Given a set of  $m$  experiments  $S = [s_1 \dots s_m]^T$  with  $s_i \in \mathbb{R}_n$  and responses  $Y = [y_1 \dots y_m]^T$  with  $y_i \in \mathbb{R}_q$ . A model  $\hat{y}(x)$  that expresses the response of a deterministic function  $y_x \in \mathbb{R}_q$  for a  $n$  dimensional input  $x \in \mathcal{D} \subseteq \mathbb{R}^n$ , as a realization of a regression model  $\mathcal{F}$  and a random function (stochastic process)  $z$  is given by [Equation 2.39](#)

$$\hat{y}_l(x) = \mathcal{F}(\beta_{:,l}, x) + z_l(x), \quad l = 1, \dots, q \quad (2.39)$$

A regression model which is a linear combination of  $t$  chosen functions  $f_j : \mathbb{R}^n \rightarrow \mathbb{R}$  is used:

$$\mathcal{F}(\beta_{:,l}, x) \equiv f(x)^T \beta_{:,l} \quad (2.40)$$

Typically,  $f(x)$  is chosen as a polynomial of order ranging from zero to two ( $h \in [0, 2]$ ).  $\beta_{k,l}$  are regression parameters and  $z$  is assumed to have mean zero and covariance between two points  $w$  and  $x$ :

$$\text{Cov}[z_l(w), z_l(x)] = \sigma_l^2 \mathcal{R}(\theta_l, w, x), \quad l = 1, \dots, q \quad (2.41)$$

where  $\sigma_l^2$  is the process variance for the  $l$ th component of the response and  $\mathcal{R}(\theta, w, x)$  is the correlation model. In the case of Kriging as a correlation function, the value of the hyperparameter  $\theta$  can be interpreted as measuring the importance or “activity” of the variable  $x$  (i.e., a low value of  $\theta$  means that all points have a high correlation) or how fast the correlation goes to zero as we move in the  $l$ th coordinate direction ([CABALLERO](#);

(GROSSMANN, 2008) In this work, we use the correlation model in the form of:

$$\mathcal{R}(\theta_l, w, x) = \exp\left(-\sum_{i=1}^m \theta_l (w - x_i)^p\right), \quad (2.42)$$

$$(\theta_l \geq 0, p_l \in [0, 2])$$

The hyperparameter  $p$  represents the “smoothness” of the correlation. In other words, by reducing its value, the rate at which the correlation initially drops as the distance between  $w$  and  $x_i$  increases. With a  $p \approx 0$ , it is tantamount to say that there is no immediate correlation between these points and there is a near discontinuity between  $Y(w)$  and  $Y(x_i)$  (FORRESTER; SOBESTER; KEANE, 2008).

For gradient and Hessian evaluation, being derived by Lophaven, Nielsen, and Søndergaard (2002) and by the authors on this study, respectively, the following equations yields

$$\hat{y}'(x) = J_f(x)^T \beta^* + J_r(x)^T \gamma^* \quad (2.43)$$

$$\hat{y}''(x) = H_f(x) \beta^* + H_r(x) \gamma^* \quad (2.44)$$

Detailed information and proofs regarding the Kriging predictor, the gradient, and the Hessian approximations are given in Appendix A.

## 2.2.2 The Optimization Algorithm

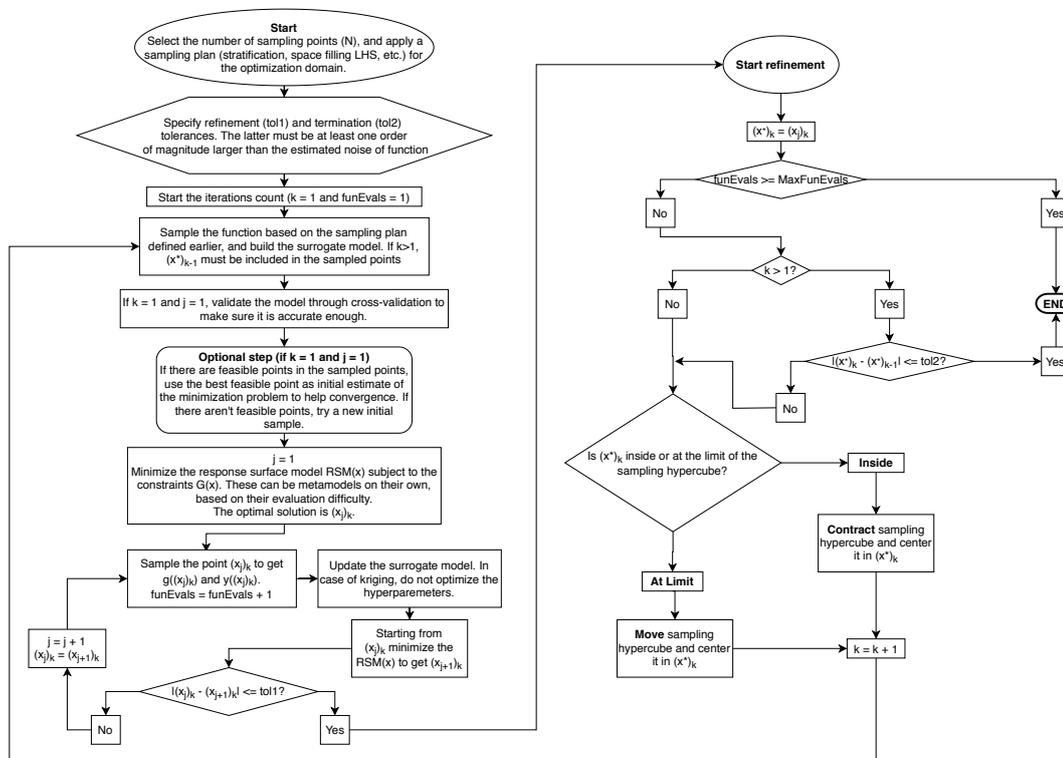
Jones (2001) made a review of several methods of global and local optimization using response surfaces, in which with a limited number of starting samples, the surface model is adjusted based on a “infill sampling criteria” (i.e., minimizing the surface directly, maximizing expected improvement or probability of improvement, lower bound of functions, etc.). Precisely, the “Method 2” presented by Jones (2001) is the one used in this work. It consists on minimizing the response surface, sampling at that point, and updating the surface model in each iteration. Once there is no improvement in two successive iterations, an optimality test is done. In addition, Jones (2001) refers to a “gradient matching” technique, where the gradient of the response surface is forced to agree with the gradient of the true function, combined with a trust region approach in order to ensure local convergence.

Alexandrov et al. (2000) developed such an algorithm to force gradient matching that uses a correction factor. To get the next iteration, the surface is optimized within a trust region around the incumbent solution. The optimum is then sampled, if the true function value at that location fails to decrease, the trust region is contracted, and a new

iteration is done. In their work, this approach is proved to converge to a critical point of the function.

Although Alexandrov’s approach of contracting the feasible region is quite robust, it was preferred to use the methodology of Caballero and Grossmann (2008) to ensure local convergence, where successive contractions and movements of the search region are done. There were four main reasons for this choice: (1) Caballero and Grossmann’s algorithm is easier to implement because there is no need to build a surrogate of Lagrangian functions like is proposed by Alexandrov et al. (2000)(2) no need of perturbations in the black box functions to obtain numerical differential data, and (3) fewer restrictions regarding the dimension of the model. Alexandrov et al. (2000) algorithm is appropriated for problems where the dimensionality (i.e., number of independent variables) is no greater than four. However, as defined by Forrester, Sobester, and Keane (2008) the “curse of dimensionality” still looms over us: the higher the dimensionality, the higher the number of sampling points required to achieve reasonably accurate predictions. (4) Lower chance of premature convergence in Caballero and Grossmann’s algorithm when dealing with highly constrained problems. The optimization algorithm is summarized in Figure 1. Its description is as follows:

Figure 1 – Flowchart for the algorithm developed.



1. Define the bounds of independent variables and the number of points to be sampled. The amount of points to be sampled is a rather heuristic choice. In addition, a prime

(or at least odd) number of points is preferred to avoid symmetrical distributions of sampling points (CABALLERO; GROSSMANN, 2008). The type of sampling plan technique is quite important. Space filling methods such as Latin hypercube sampling (LHS) are favored to guarantee that the entire domain space is covered (uniformity of distribution is more important than its randomness), and the numerical noise introduced by black-box models does not significantly affect the surrogate (because it is a variance reduction technique through the maximin criterion which selects the best sampling plan that maximizes the minimum distance between the points (FORRESTER; SOBESTER; KEANE, 2008)). This numerical noise (which we will just call for the sake of brevity as noise) is defined as the uncertainty introduced by the black-box model due to its internal stopping criteria (e.g., the tolerance of the solver used to solve MESH equations in a distillation column.), truncation, or rounding of values, in this case, by the process simulator. We also include  $2^n$  points that correspond to the bounds of the independent variables; this is done to ensure the Kriging does not perform “extrapolations” near the corners of the optimization space, and the remaining  $m - 2^n$  points are distributed by the sampling plan chosen (QUIRANTE; JAVALOYES; CABALLERO, 2015). However, this procedure of inserting the hypercube vertices is limited for cases where the number of input variables is low (e.g., less than 6) due to the amount of samples required (i.e., in the case of 8 input variables, we would need to sample 256 points excluding the ones from the LHS) and the “extrapolation” inconvenience is present and needs to be fixed so that meaningful results can be obtained. Still regarding the bounds, the better performance of the optimization algorithm (i.e., a lesser number of true function samplings) is attained, the narrower they are

2. Estimate the noise so that refinement and termination tolerances can be set. One way to do it: simulate the flowsheet process for a fixed set of independent variables starting from different initial points (CABALLERO; GROSSMANN, 2008). Moreover, it is recommended to tighten the model convergence criteria (e.g., for distillation columns, set the equation solver convergence as low as possible) in order to get consistent results with reduced noise variance. Then the termination criteria should be at least 1 order of magnitude greater than the noise.
3. Build the surrogate model using the sampled data from the black box model. In the Kriging case, it is important that the surrogate models be univariate because Kriging does not incorporate cross-correlation between different simulation outputs (QUIRANTE; JAVALOYES; CABALLERO, 2015; KLEIJNEN, 2009).
4. Check if the surrogate is accurate enough. This step is executed only once, before the optimization step 5 begins. Here, this is done by using cross-validation and plotting the “standardized cross-validated residuals” (JONES; SCHONLAU; WELCH, 1998)

or the relative error between the prediction and actual values. If the number of sampled points is low, say  $m < 100$ , use the “leave-one-out” (subset size  $q = 1$ ) type of cross-validation. The model is considered valid if the number of standard errors that the actual value is above or below the predicted value lies in the interval  $[-3, +3]$  (i.e., 99.73% of confidence) or there are not too many outliers. For larger number of sampled points, use subsets size  $q = 5$  or  $10$  and check if the cross-validation error is low enough (FORRESTER; SOBESTER; KEANE, 2008) to be considered a valid model. The reason for larger subsets  $q$  is because the computation time becomes prohibitive in the case where  $q = 1$  and  $m > 100$  because  $m$  surrogate models are built, although this can be mitigated using parallel processing capabilities when available.

5. Perform the optimization of the black-box metamodel:

$$\begin{aligned} & \min_x \hat{y}(x) \\ & \text{subject to:} \\ & \hat{y}(x) \leq 0 \end{aligned} \tag{2.45}$$

**Remark:** An optional step is proposed to facilitate convergence. Only before in the first optimization, a screening of the initial sampling data is done for the best feasible point found. If such a point exists, use it as the initial estimative of a standard constrained nonlinear programming (NLP) solver. In this work, the MATLAB *fmincon* routine was used coupled with gradient information on objective and constraint metamodels functions from Equation 2.43 as a solver of the NLP problem Equation 2.45, simply to aid the convergence performance.

6. Sample the optimum point obtained from Equation 2.45, add it to the set of sampled points, and update the Kriging model without reoptimizing its hyperparameters. Because Kriging is an interpolating procedure, all new points introduced are exact (zero error) (CABALLERO; GROSSMANN, 2008). However, a reoptimization of the hyperparameters might be done if the user needs to be sure that the surrogate model is adequate.
7. If the optimum found in two consecutive iterations does not change based on the refinement tolerance (tol1), start the refinement procedure (proceed to step 8). Even using the simplest “infill criteria” (i.e., minimizing the surface response, sampling the optimum, and updating the Kriging until there is no significant change between the points sampled) does not guarantee a local optimum as aforementioned. To address this issue, it is necessary to use the gradient matching combined with a trust-region approach (JONES; SCHONLAU; WELCH, 1998; CABALLERO; GROSSMANN, 2008; QUIRANTE; JAVALOYES; CABALLERO, 2015; ALEXANDROV et al., 2000).

8. The refinement procedure is done as follows:

- If the optimal solution obtained in step 7 is an internal point of the original hypercube, select a contraction factor and reduce the hypercube limits. This reduced hypercube must be centered at the last optimal solution obtained. The contraction factor is a tunable parameter. We use the same default values as Caballero and Grossmann's implementation: 40% for the first contraction and 20% for the subsequent ones.
- If the optimal solution is at the limit of the hypercube, then do not contract it, simply move the hypercube and center it at the optimal solution.
- The limits of the hypercube can go further than the bound of the variables, although the sampling is always performed inside those bounds. The reader is referred to the work of [Caballero and Grossmann \(2008\)](#) for a better understanding of the actions taken in the refinement procedure.

9. Go back to step 3 and repeat until the termination criteria or maximum number of samplings is met. If the maximum number of sampling is achieved, probably it is because the domain bounds are too wide or the Kriging model is not accurate enough; to remedy this, one could simply narrow the bounds or try to increase the number of points in the initial sampling plan.

After the algorithm successfully converges, we proceed to the next phase, which is the active constraint checking and obtainment of the gradient and Hessian through Equations 2.43-2.44.

### 2.2.3 Usage of Jacobian and Hessians evaluated by Kriging equations

One could ask oneself the reason why Kriging was chosen as the surrogate-builder method, among a wide variety of reduced models available (i.e., polynomials, neural networks, etc.). As could be seen at the discussion so far, Kriging interpolators have an analytic expression for both gradient and Hessian evaluation. The absence of using numerical perturbations to generate the high-order data (gradient/Hessian) removes a source of uncertainty derived from heuristics. For instance, [Araújo, Govatsmark, and Skogestad \(2007\)](#) used the first-order forward difference approximation to evaluate each element of the gradients, which can be quite sensitive to noise:

$$\frac{\partial c_i(i)}{\partial u_j} = \lim_{h \rightarrow 0} \frac{c(u + e_j h) - c(u)}{h_j} \quad (2.46)$$

For hessian approximation:

$$\frac{\partial J(u)}{\partial u_i \partial u_j} = \lim_{h \rightarrow 0} \frac{J(u + E_{ii}h + E_{jj}h) - J(u + E_{ii}h) - J(u + E_{jj}h) + J(u)}{[hh^T]_{ij}} \quad (2.47)$$

[Silva et al. \(2017\)](#) in order to avoid the problem stated before, used the Akima cubic spline and the bicubic spline to evaluate the gradients and Hessians, respectively. However, as stated in their work, one disadvantage of using cubic splines is when there is an outlier in the region studied, jeopardizing the evaluation in some cases. Therefore, it can be seen that the approaches from past contributors could fail in evaluating gradients and Hessians.

We argue that, along with the fact where the expressions derived for gradient and Hessian approximations using Kriging interpolators are explicit and analytical, the nature of Kriging itself, being the best linear unbiased predictor (BLUP) , as can be seen in the work of [Sacks et al. \(1989\)](#) are essential characteristics for a reduced model when high-order data approximation are paramount to any methodology, as it is in self-optimizing control.



## Part III

### Case studies



## 3 Introduction

At this section, a systematic approach of the proposed methodology will be done. Three case studies with increasing complexity will be discussed in detail. The main idea of this section consists in:

1. Display the methodology through the case studies.
2. Compare the results obtained using surrogates with the traditional approach (i.e.: Optimization with the original nonlinear model - when possible - and comparing the high-order data obtained from the original models, also when available)
3. Propose metamodel-based control structures (Single measurement policy or linear combinations) and compare them with the ones found in the literature using nonlinear models

The three case studies are:

- Economic Self-Optimizing Control of a Depropanizer column.
- Economic Self-Optimizing Control of a Evaporation process.
- Economic Self-Optimizing Control of Cumene Process production.



## 4 Case studies

### 4.1 Case I - Economic Self-Optimizing Control of a Depropanizer column

#### 4.1.1 Problem description

The first problem to be studied is a depropanizer column (described in [Figure 2](#)), similar to the one considered in [Skogestad \(2000\)](#). The feed is composed by a propene/propane mixture with the mole fraction of the light component nominally at 65 mol %. The column has 146 stages, a feed composed of saturated liquid at 12 atm and a molar flow rate of 420 kmol/h. Nominally, the tower is operated at a molar reflux ratio and a distillate to feed ratio of 11 and 0.75, respectively. The feed enters at stage 108 (with the top stage being numbered as the first one), and the tower has a pressure drop of 1 atm across itself, and a top stage of 9.7 atm. The overhead product must be at least 99.5% pure in molar units ( $X_D \geq 0.995$ ). The only physical limitation imposed to the system, besides non-negativity of flow rates and concentrations, is regarding the reboiler, unable of adding more duty than  $80GJ/h$  to the process. The process has five degrees of freedom in a control point of view, but two of these are used to stabilize the holdups, and another one is used in a “floating pressure” control approach using the maximum cooling capacity available, as suggested by [Shinskey \(1984\)](#) improving the relative volatility. Therefore, two degrees of freedom are available for optimization.

The degrees of freedom chosen (arbitrarily) were

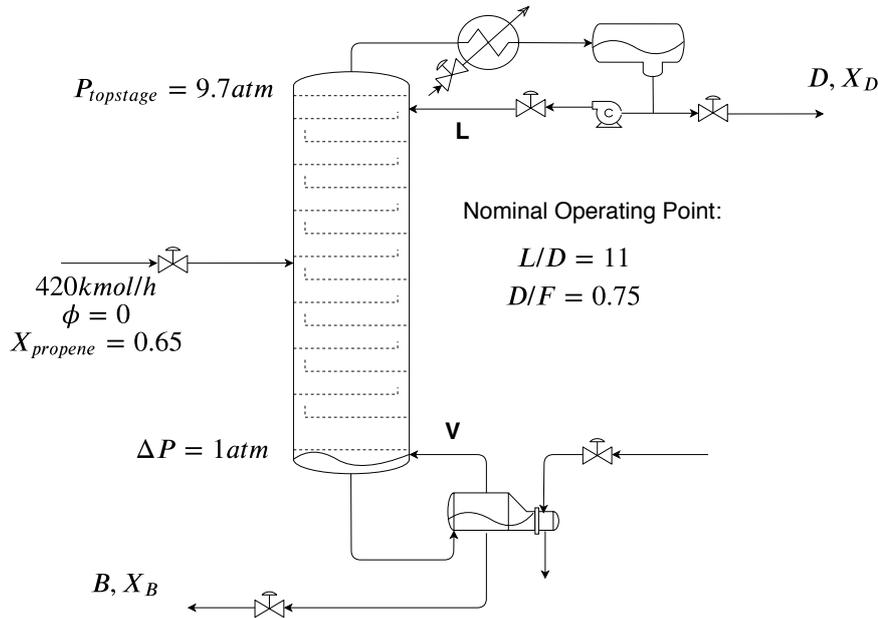
$$u = \begin{pmatrix} L/D \\ D/F \end{pmatrix} \quad (4.1)$$

The objective function is given as the profit of the unit (\$/h):

$$J = 20D + (10 - 20x_B) B - 70Q_R \quad (4.2)$$

which accounts for the production of propene, the cost of the steam used at the reboiler, the propene lost at the bottoms stream, and the production of propane.

Figure 2 – Depropanizer column.



Therefore, the optimization problem becomes

$$\begin{aligned}
 \max_u J &= 20D + (10 - 20x_B) B - 70Q_R \\
 \text{s.t.} & \\
 x_D &\geq 99.5\% \\
 Q_R &\leq 80\text{GJ/h}
 \end{aligned}
 \tag{4.3}$$

In addition, the degrees of freedom available (manipulated variables) are bounded due to physical/operational limitations:

$$\begin{aligned}
 L/D &\leq 50 \\
 D/F &\leq 0.9
 \end{aligned}
 \tag{4.4}$$

The disturbances considered in this case were the flow rate for each component at the feed and the feed vapor fraction, varying up to  $\pm 10\%$  of their nominal values.

Initially, The candidates as controlled variables considered were the same as Skogestad (2000) ( $x_D$ ,  $x_B$ ,  $D/F$ ,  $L/F$ ,  $V/F$ , and  $L/D$ ), among others possible.

#### 4.1.2 Surrogate Model Generation

Specially in this case study for illustration purposes, the infill criteria was manually performed in order to show to the user the adjustment of the objective function and the nonlinear constraints, a phenomena described in detail by Forrester, Sobester, and

Keane (2008). However, it is clear that the methodology proposed uses the algorithm from Caballero and Grossmann (2008), due to the fact that there is no user intervention on the latter.

#### 4.1.2.1 Initial Sample

Using the bounds for the manipulated variables and disturbances, 999 cases were generated randomly using the Latin hypercube sampling method (this is achieved by scaling the results from the MATLAB's `lhsdesign` routine with a maximin criterion to the input limits given because the routine results are scaled between 0 and 1), and a sensitivity analysis was performed at the simulator (Aspen Plus) to obtain the outputs for each measurement (CV candidates), the cost function, and the nonlinear constraints (Reboiler duty and Propene tops composition). With the input/output data, a Kriging response was generated using DACE toolbox, from Lophaven, Nielsen, and Søndergaard (2002) using the Gaussian correlation and linear regression models. There were 100 cases left out of the construction of the surrogate response for validation purposes, as suggested by different authors (FORRESTER; SOBESTER; KEANE, 2008; QUIRANTE; JAVALOYES; CABALLERO, 2015). Figure 3 shows the validation between those cases from the surrogate response and the original ones evaluated at the process simulator, corroborating the effectiveness of the reduced model generated. Another sample of merely 33 points was used as the initial sample to be provided to the algorithm of Caballero and Grossmann (2008). This was done to investigate if the algorithm is capable of finding an optimal operating point with a drastically reduced initial sample through its operations of movement and contraction of the hyperspace.

#### 4.1.3 Optimization of the Surrogate Response

The optimization performed was done with genetic algorithms available in MATLAB and Microsoft Excel. The objective function of the problem and its nonlinear constraints were the Kriging responses generated earlier. Both infill criteria methods described here before, brute-force and the automated algorithm from Caballero and Grossmann (2008) gave similar results for the optimal decision variables. The brute-force infill was performed 12 times, until the stopping criteria was achieved, in this case this criteria was the variation between the infills on decision variables and objective function values became irrelevant (cents magnitude in cost function). Therefore, the final first kriging response was composed of 1014 for the brute-force method: 999 points, followed by the edges of the hypercube, and the 12 infills cited above. Table 2 shows the importance of the infill criteria, adjusting the shape of the kriging prediction closer to the real function, especially for the response of the propene composition at the top stream of the tower, which tend to have a noisy response (QUIRANTE; JAVALOYES; CABALLERO, 2015).

Figure 3 – Kriging hold-out validation for the depropanizer case study.

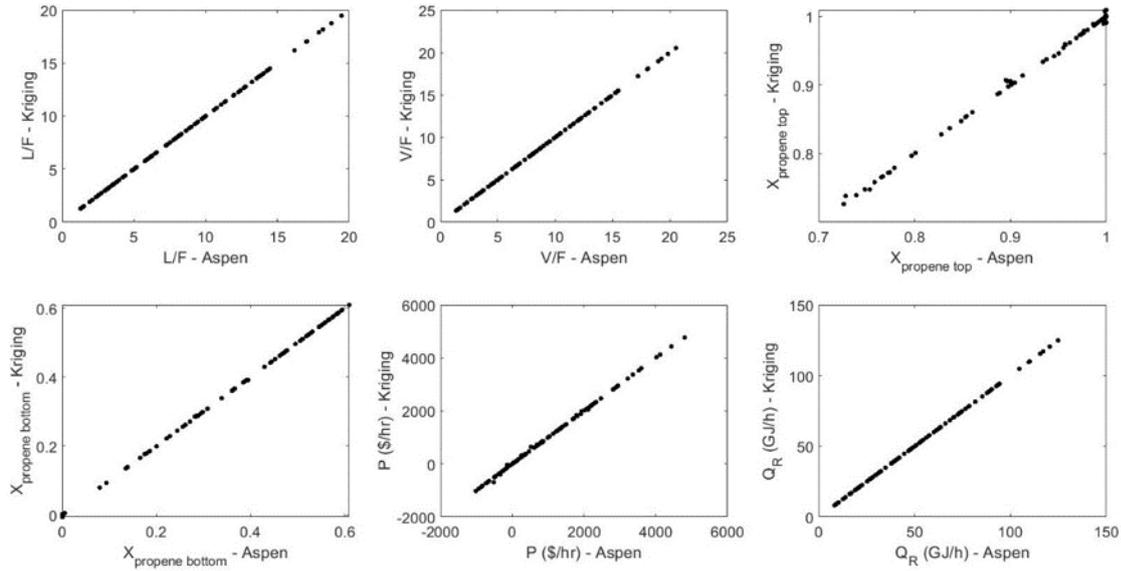


Table 2 – Brute-force infill criteria - Depropanizer case study.

Case	L/D	D/F	L/F	V/F	$X_D$	$X_B$	P [\$/h]	$Q_R$ [GJ/h]
1	12.5002	0.6388	7.9846	8.6918	0.991818	0.045590	3045.450	52.843
2	13.5982	0.6399	8.7015	9.4197	0.994949	0.037024	2766.524	57.272
3	13.6265	0.6405	8.7276	9.4467	0.994976	0.035416	2762.509	57.437
4	13.6194	0.6398	8.7139	9.4323	0.994998	0.037159	2760.410	57.348
5	13.6212	0.6399	8.7157	9.4341	0.994999	0.037029	2760.226	57.360
6	13.6204	0.6398	8.7147	9.4330	0.995000	0.037139	2760.180	57.353
7	13.6155	0.6396	8.7087	9.4269	0.995000	0.037675	2760.294	57.315
8	13.5773	0.6378	8.6602	9.3762	0.995001	0.042359	2759.732	57.005
9	13.5675	0.6375	8.6486	9.3642	0.994999	0.043403	2759.934	56.931
10	13.5686	0.6375	8.6499	9.3655	0.995000	0.043295	2759.886	56.939
11	13.5601	0.6371	8.6388	9.3540	0.994999	0.044386	2759.597	56.869
12	13.5641	0.6373	8.6438	9.3591	0.995000	0.043916	2759.633	56.900

The results for both optimization techniques with infill criteria, compared to the optimization computed on the original rigorous model in Aspen Plus, are described in [Table 3](#):

Table 3 – Optimization Results of Surrogates and Original Model.

	Profit [\$/h]	D/F	L/D	$X_D$	$Q_R$ [GJ/h]
Aspen Plus ®	2759.608	0.6372	13.5629	0.9950	56.8907
Brute-force infill	2759.699	0.6372	13.5623	0.9950	56.8875
Caballero's algorithm	2760.600	0.6389	13.5982	0.9950	57.1829

#### 4.1.4 Reduced Space Surrogate Model.

The composition of propene at the top stream is active for all the range of the manipulated variables and disturbances, as can be seen on the previous subsection. Thus, the control of this variable is mandatory [Skogestad \(2000\)](#) consuming one degree of freedom. The active-constraint control is implemented at the process simulator, and a second design of experiments is carried out (with only 50 points) with the the amplitude of  $\pm 0.1\%$  of the nominal values of the remaining degree of freedom and of the disturbances, to ensure that the gradients and the Hessian with respect to the remaining degree of freedom and disturbances are following the assumption of linearization of the cost function around the nominal optimal point, necessary to evaluate the worst-case loss using the exact local method and the extended nullspace method by [Alstad, Skogestad, and Hori \(2009\)](#) and to evaluate the average-case loss developed by [Kariwala, Cao, and Janardhanan \(2008\)](#).

The reduced space problem remains with one degree of freedom left. Without loss of generality,  $D/F$  was chosen to close the loop with the composition of propene at top stream. Because the model and its surrogate are in steady-state, the choice of using either of the DOFs available does not influence the methodology. Consequently, the measurements left as candidates for self-optimizing control are

$$y = [x_B, L/F, V/F, D/F, L/D] \quad (4.5)$$

#### 4.1.5 High-Order Data Obtainment

The DACE toolbox from [Lophaven, Nielsen, and Søndergaard \(2002\)](#) evaluates the gradient of the surrogate response natively. For the Hessian, the expression developed in this study was capable of evaluating the second-order derivatives. The gradients and Hessians with respect to the remaining degree of freedom and disturbances are

$$G_y = \begin{bmatrix} 1.3560 \\ 1.4169 \\ -0.1389 \\ 0.0523 \\ 1 \end{bmatrix} \quad G_y^d = \begin{bmatrix} 0.0170 & -0.0316 & 0.1511 \\ 0.0184 & -0.0342 & -1.1291 \\ -0.00099 & 0.0018 & 0.0290 \\ 0.0013 & -0.0023 & -0.0111 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.6)$$

$$\begin{aligned} J_{ud} &= [15.1472 \quad -28.0727 \quad -485.7841] \\ J_{uu} &= 1921.6795 \end{aligned} \quad (4.7)$$

### 4.1.6 Expected Disturbances and Implementation Errors

The weighting matrix for the disturbances  $W_d$  is up to  $\pm 10\%$  of its nominal values, the allowable magnitude for them in the process is therefore

$$W_d = \text{diag}(27.3, 14.7, 0.1) \quad (4.8)$$

For the weighting matrix of implementation errors, a magnitude of 0.1% over the optimal nominal values for flow ratios and composition was considered. Then  $W_n^y$  becomes

$$W_n^y = \text{diag}\left(0.0086427, 0.009358, 4.4020 \times 10^{-5}, 6.372 \times 10^{-5}, 0.0135632\right) \quad (4.9)$$

### 4.1.7 Loss Evaluation

Evaluating the worst-case loss and average-case loss for single measurements as candidates, the best candidate is the composition of propene at bottoms. Therefore, the set of CVs becomes

$$c_1 = \begin{pmatrix} x_D \\ x_B \end{pmatrix} \quad (4.10)$$

As found previously by Skogestad (2000) This set, however, needs some engineering insight because dual composition control is a well-known problem for distillation columns, given the unreliability and expensiveness of online analyzers and due to strong interactions between the components of this control structure (SHINSKEY, 1984; SKOGESTAD; MORARI, 1987). Therefore, alternatives with other single measurements and combinations will be discussed.

Table 4 shows the worst-case and average-case losses for single measurements policy:

Table 4 – Single-measurement policy: Results for worst and average-case losses for depropanizer case study.

Measurement	Worst-case loss [\$/h]	Average-case loss [\$/h]
$X_B$	0.8244	0.0687
$L/F$	31.4245	2.6187
$V/F$	40.4240	3.3687
$L/D$	89.5913	7.4659
$D/F$	356.9744	29.7479

From Table 4, it is possible to conclude that controlling the composition of bottom stream indeed yield the lowest loss in both worst and average cases, followed by  $L/F$  and  $V/F$ . Open loop policies, as expected ( $D/F$ ,  $L/D$ ) yield unacceptable losses. This result

can be related to the similar example developed by Skogestad (2000) showing the capability of the surrogate model generated of being used for self-optimizing control purposes. A good trade-off between complexity of implementation the control structure and operating it can be achieved by using  $L/F$  as the last CV to square the problem, due to the fact that is easier to keep a liquid flow constant rather than a vapor one (SKOGESTAD, 2000). Given this fact, another possible set of single measurements could be

$$c_2 = \begin{pmatrix} x_D \\ L/F \end{pmatrix} \quad (4.11)$$

Considering all measurements available as linear combinations, using Equation 2.27 (with explicit solution from Alstad, Skogestad, and Hori (2009)), the worst-case and average losses decrease drastically to 0.000486 \$/h and 0.000020 \$/h, respectively. Using the extended null space method, these losses are 1.6384 \$/h and 0.0683 \$/h. These results were expected due to the fact that the increasing number of measurements available reduces significantly the loss as seen in previous works (ALSTAD; SKOGESTAD; HORI, 2009; KARIWALA; CAO; JANARDHANAN, 2008). However, implementing such structure can be a challenging task.

Other possible subsets of measurements combinations were also evaluated and are available in Table 5. As also noted by Kariwala, Cao, and Janardhanan (2008) there is a clear trade-off between complexity and the incurred loss.

Table 5 – Loss evaluation: linear combination of measurements of different subsets' sizes for depropanizer case study.

$n_y$	Measurements	Worst-case loss [\$/h]	Average-case loss [\$/h]
2	$x_B, L/D$	0.0070	0.0005
	$x_B, V/F$	0.0095	0.0006
	$x_B, D/F$	0.0192	0.0013
	$x_B, L/F$	0.0328	0.0022
	$L/F, D/F$	0.4240	0.0283
	$V/F, L/D$	0.5398	0.0360
	$V/F, D/F$	6.5368	0.4358
	$L/F, V/F$	30.5602	2.0373
3	$x_B, D/F, V/F$	0.0005046	0.0000280
	$x_B, L/F, V/F$	0.0005935	0.0000330
	$x_B, L/D, V/F$	0.0009989	0.0000555
	$V/F, D/F, L/D$	0.0601630	0.0033424
	$L/F, V/F, D/F$	0.0813750	0.0045208
	$L/F, D/F, L/D$	0.3605520	0.0200306
4	$x_B, D/F, V/F, L/D$	0.0004890	0.0000230
	$x_B, L/F, V/F, D/F$	0.0004990	0.0000240
	$x_B, L/F, V/F, L/D$	0.0006000	0.0000263
	$x_B, D/F, L/D, L/F$	0.0040000	0.0002000
	$L/F, V/F, D/F, L/D$	0.0309000	0.0015000

Another possibility of implementation is to consider a “just enough” subset of measurements for the null space method (prioritizing the reduction of disturbances effects

on the loss) when the number of these are equal to the sum of the number of disturbances and degrees of freedom ( $n_y \geq n_u + n_d$ ). Then, using Equation 2.26 to evaluate H, the losses can be calculated and some of the possible subsets are present in Table 6. However, as observed by Alstad, Skogestad, and Hori (2009) this requires the evaluation of all possible subsets of measurements, which in larger problems may be impractical. To overcome this, a suboptimal selection can be done, maximizing  $\underline{\sigma}(\tilde{G}^Y)$  using the branch-and-bound algorithm proposed by Cao and Kariwala (2008). Because the column has a small number of combinations for this subset size, given as

$$\binom{5}{4} = \frac{5!}{4!(5-4)!} = 5 \quad (4.12)$$

Both methodologies were computed to show the difference between the optimal and suboptimal solutions. From Table 6, it can be clearly shown that for this example, prioritizing the reduction of loss generated by disturbances and neglecting the implementation errors to deal with then later (if there are remaining degrees of freedom left), as suggested by Alstad, Skogestad, and Hori (2009) may not be a good choice.

Table 6 – Loss evaluation: sub-optimal solution with extended null space method for depropanizer case study.

$n_y$	Measurements	Worst-case loss [\$/h]
	$x_{\bar{B}}, L/F, V/F, L/D$	1.8349
	$x_B, D/F, V/F, L/D$	3.2075
	$x_B, L/F, V/F, D/F$	4.4891
	$L/F, V/F, D/F, L/D$	$1.1925 \times 10^6$
	$x_B, D/F, L/F, L/D$	$1.7051 \times 10^6$

Using the suboptimal rule of maximizing  $\underline{\sigma}(\tilde{G}^Y)$ , the subset found by the branch and bound method was  $y = [L/F, V/F, x_B, L/D]$  which yields worst-cases losses using the exact local method and the extended null space method of 0.0006 \$/h and 1.8349 \$/h, evidencing the sensitivity to implementation errors in some cases. The sub-optimality of this approach is also evidenced because for a subset of four measurements, as can be seen in Table 5, using the exact local method the subset  $y = [x_B, D/F, V/F, L/D]$  yielded worst-case and average-case losses of 0.000489 \$/h and 0.000023\$/h, respectively, differently from the result cited before. In fact, the subset found by the suboptimal approach is the third that yielded the lowest loss through the exact local method, as can be seen in Table 5.

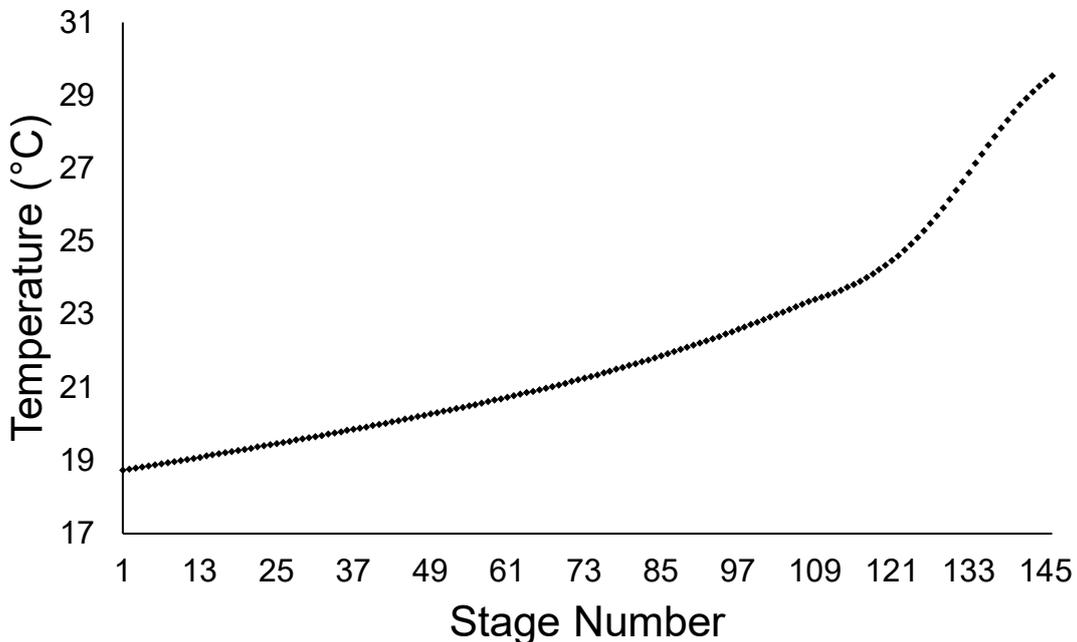
Control structures considering tray temperatures as controlled variables in distillation processes is a common industrial practice, and self-optimizing control can be used to evaluate the loss of implying a structure that uses these variables, with a first-principles model or a surrogate-based one (actually, if the surrogate is accurate enough, any variable

that can be measured can be considered to be tested as a CV candidate, which is the main point of this work). To illustrate that, the example was extended to the case where some stage temperatures are available measurements.

It can be seen in Figure 4 that the region between stages 120-140 has a steeper “slope” (which actually is the change in temperature from tray to tray) if compared with the rest of the stages along the column. In fact, this is one of the “classical” available criteria to prescreen optimal stage temperatures to be used as controlled variables, as reported by Luyben (2006, 2013). Therefore, it was considered as available measurements for some promising candidates (and some that with the criteria shown above, will be poor candidates, namely stages 10, 11 and 12) to check if the self-optimizing control methodology will corroborate the proposed methodology by Luyben (2006, 2013):

$$y = [T_{10}, T_{11}, T_{12}, T_{120}, T_{125}, T_{130}, T_{133}, T_{134}, T_{135}, T_{136}] \quad (4.13)$$

Figure 4 – Temperature profile for depropanizer column.



The expected disturbances were considered as the same as the first part of this case study, up to  $\pm 10\%$  of nominal conditions:

$$W_d = \text{diag}(27.3, 14.7, 0.1) \quad (4.14)$$

and the implementation error for the temperatures was considered as  $\pm 1C$ :

$$W_n^y = \text{diag}(1, 1, 1, 1, 1, 1, 1, 1, 1, 1) \quad (4.15)$$

Considering the exact local method with explicit solution [Alstad, Skogestad, and Hori \(2009\)](#) using single temperature measurements, the following losses were obtained ([Table 7](#)):

Table 7 – Temperatures Single measurements as CV candidates: Worst and average-case losses for depropanizer case study.

Measurement	Worst-case loss [\$/h]	Average-case loss [\$/h]
$T_{133}$	43.5694	3.6308
$T_{130}$	44.7776	3.7315
$T_{134}$	45.2288	3.7691
$T_{135}$	48.1316	4.0110
$T_{136}$	52.4720	4.3727
$T_{125}$	69.9778	5.8315
$T_{120}$	156.7512	1.0626
$T_{11}$	$1.4810 \times 10^9$	$1.2342 \times 10^8$
$T_{11}$	$2.3186 \times 10^9$	$1.9322 \times 10^8$
$T_{10}$	$3.8192 \times 10^9$	$3.1827 \times 10^8$

It can be clearly seen from [Table 7](#) that, in fact, one of the temperatures that is in the region with steepest slopes between stages, which is stage 133, have the best self-optimizing control properties, generating the lowest worst and average-case losses among the all stages considered (45.2288 \$/h and 3.6308 \$/h, respectively), followed also by temperatures that are in the same region. However, stages 10, 11, and 12, which are in a region with “flatter” slopes (checking the column temperature profile, from [Figure 4](#)), generated the worst possible losses for this case study. Both results were expected because considerable changes in temperature between trays indicate a region where compositions of the components are changing more abruptly, and therefore, keeping a constant tray temperature in this region should hold the composition profile ([LUYBEN, 2006, 2013](#)), if compared with a column region that has irrelevant temperature change between stages (such as the region where stages 10,11 and 12 are located).

Considering linear combinations of the stage temperatures available as CVs candidates, different subsets were generated and the best and worst of them (to display all of them is impracticable due to combinatorial explosion) are available in loss ascending order ([Table 8](#)).

From [Table 8](#) it can be shown that the linear combinations that uses temperatures from the column bottom region generates the subsets that yields the lowest losses. On the other hand, subsets generated with temperatures of stages 10, 11, and 12 produce unacceptable losses.

Table 8 – Linear Combinations of Stage Temperatures - Depropanizer case study.

	Measurements	Worst-case loss [\$/h]	Average-case loss [\$/h]
2	$T_{133}, T_{134}$	23.0352	1.5357
	$T_{130}, T_{133}$	23.0895	1.5393
	$T_{130}, T_{134}$	23.4697	1.5646
	$T_{133}, T_{135}$	23.6773	1.5785
	$T_{11}, T_{12}$	$8.9655 \times 10^8$	$5.9770 \times 10^7$
	$T_{10}, T_{12}$	$1.0562 \times 10^9$	$7.0415 \times 10^7$
	$T_{10}, T_{11}$	$1.4280 \times 10^9$	$9.5199 \times 10^7$
3	$T_{130}, T_{133}, T_{134}$	16.0981	0.8943
	$T_{133}, T_{134}, T_{135}$	16.2733	0.9041
	$T_{130}, T_{133}, T_{135}$	16.3679	0.9093
	$T_{10}, T_{11}, T_{12}$	$7.2400 \times 10^8$	$4.0222 \times 10^7$
4	$T_{130}, T_{133}, T_{134}, T_{135}$	12.6869	0.6041
	$T_{130}, T_{133}, T_{134}, T_{136}$	12.8910	0.6139
	$T_{133}, T_{134}, T_{135}, T_{136}$	12.9549	0.6169
	$T_{110}, T_{111}, T_{112}, T_{125}$	69.9623	3.3315
	$T_{110}, T_{111}, T_{112}, T_{120}$	156.8485	7.4690

#### 4.1.8 Dynamic Simulations

Dynamic performance of two control structures were considered: controlling  $L/F$  with the unconstrained degree of freedom along with the active constraint found (distillate propene molar fraction), and controlling a linear combination of two temperatures ( $T_{133}$ ,  $T_{134}$ ) to show the robustness of the subsets found using “surrogate modelling-aided” self-optimizing control.

The regulatory layer was implemented with a classical L-V structure, using the distillate flow rate to control the reflux drum level and using the bottom flow rate to keep the sump level at its set point, with proportional-only controllers, tuned with the Ziegler-Nichols method. Regarding the first control structure, on the supervisory layer (also tuned with Ziegler- Nichols method), an RGA analysis suggests the pairing of the active constraint (propene top stream molar fraction) with the boilup flow rate and to keep the  $L/F$  ratio manipulating the reflux flow rate ( $L$ )

$$c_2 = \begin{pmatrix} x_D \\ L/F \end{pmatrix} \quad (4.16)$$

The first and second lines of the following matrix correspond to propene composition at the top stream and  $L/F$ , respectively. The first and second columns corresponds to the reflux flow rate and boil up flow rate:

$$\Lambda_{\text{structurel}} = \begin{bmatrix} -0.052 & 1.052 \\ 1.052 & -0.052 \end{bmatrix} \quad (4.17)$$

Disturbances on feed flow rate, composition, and vapor fraction were considered to evaluate the performance of the control structure.

Figure 5 – Valve opening and closing (5% amplitude) - Depropanizer case study, control structure 1.

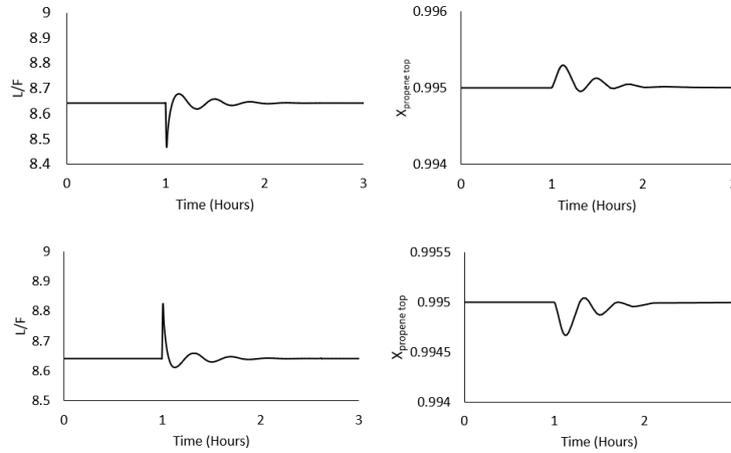
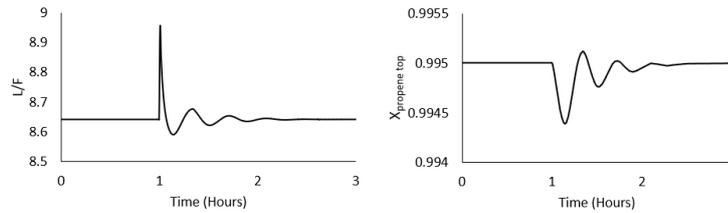


Figure 6 – 2.5% increase in vapor fraction - Depropanizer case study, control structure 1.

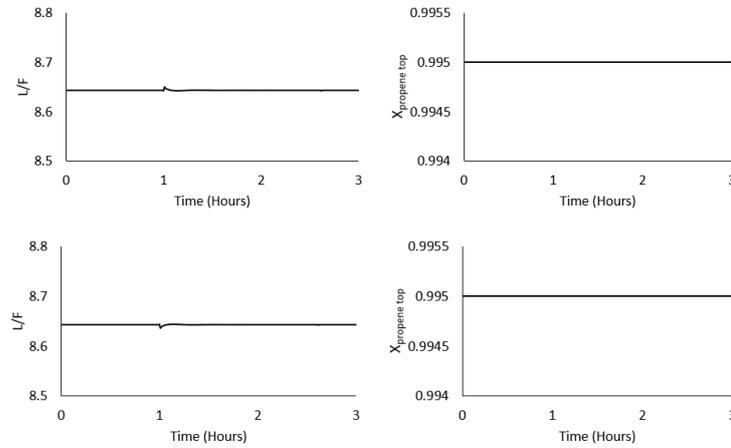


As can be seen in Figures 5-7, the control structure is robust and capable to deal with the expected disturbances of the process. As stated before, another control structure, considering a linear combination of two temperatures ( $T_{133}, T_{134}$ ) was also implemented because its worst-case loss is promising (23.0532 \$/h). Coefficients for the linear combinations are from the H matrix obtained by Equation 2.27, and the control structure becomes

$$c_3 = \begin{pmatrix} 0.71343T_{133} + 0.70070T_{134} \\ x_D \end{pmatrix} \quad (4.18)$$

A second RGA analysis was performed, and it suggested the pairings  $x_D - V$  and  $T_{133} + T_{134} - L$ . Similarly, as before, the first and second lines of the following matrix correspond to propene composition at the top stream and the combination of temperatures, respectively. The first and second columns correspond to the reflux flow rate and boilup

Figure 7 –  $\pm 5\%$  disturbance of propene feed molar fraction. - Depropanizer case study, control structure 1.

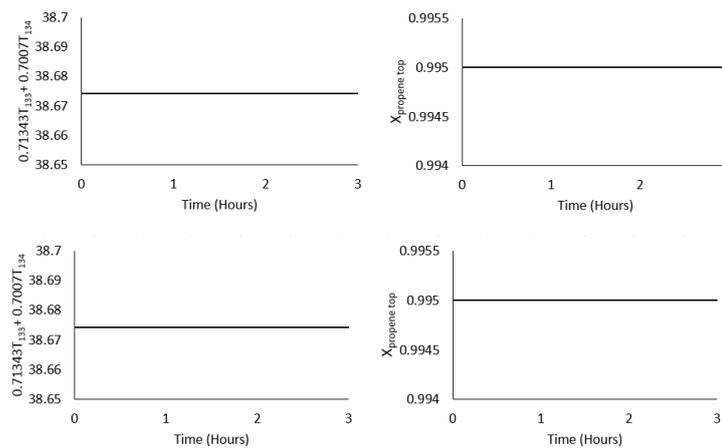


flow rate:

$$\Lambda_{\text{structure 2}} = \begin{bmatrix} -0.0328 & 1.0328 \\ 1.0328 & -0.0328 \end{bmatrix} \quad (4.19)$$

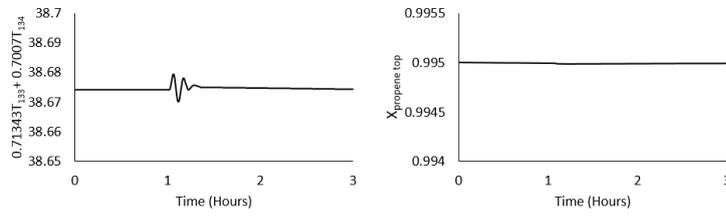
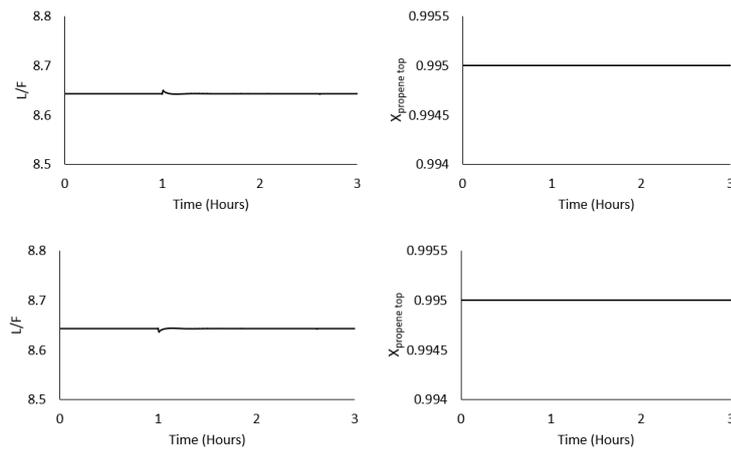
Similarly, as in the first proposed control structure, disturbances in feed flow rate, composition and vapor fraction were performed:

Figure 8 – Valve opening and closing (5% amplitude) - Depropanizer case study, control structure 2.



It can be seen in Figures 8-10 that the second control structure also has a robust performance, being able to deal with the disturbances expected, while the worst-case loss (23.0352 \$/h) is smaller if compared with the first control structure (31.4245 \$/h). Therefore, we recommend the use of the second one.

Figure 9 – 2.5% increase in vapor fraction - Depropanizer case study, control structure 2.

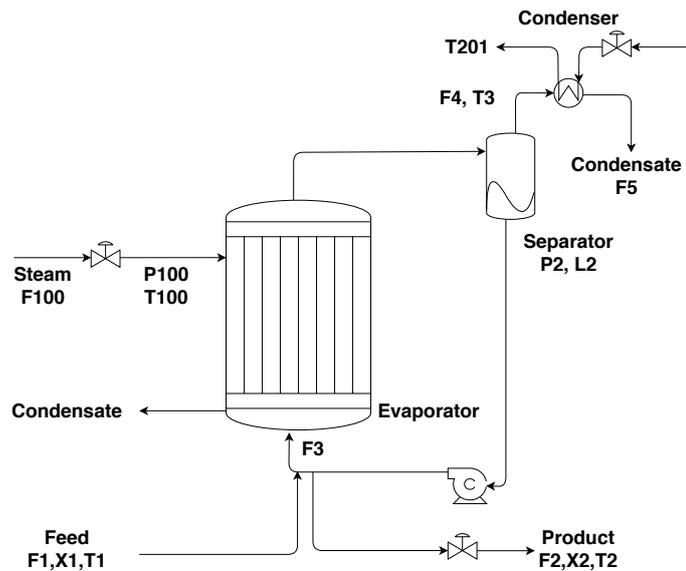
Figure 10 –  $\pm 5\%$  disturbance of propene feed molar fraction. - Depropanizer case study, control structure 2.

## 4.2 Case II - Economic Self-Optimizing Control of a Evaporation process

### 4.2.1 Problem Description

This case consists in an evaporation process slightly modified [Kariwala, Cao, and Janardhanan \(2008\)](#) from the work of [Newell and Lee \(1989\)](#) a “forced circulation” evaporator designed to increase the concentration of dilute liquor by evaporating solvent from the feed using a vertical heat exchanger. This process was also studied by [Govatsmark and Skogestad \(2001\)](#) The system can be seen in [Figure 11](#).

Figure 11 – Evaporation system flowsheet.



The objective function (of economic nature) is the same as the one depicted in the work of [Kariwala, Cao, and Janardhanan \(2008\)](#)

$$J = 600F_{100} + 0.6F_{200} + 1.009(F_2 + F_3) + 0.2F_1 - 4800F_2 \quad (4.20)$$

The first three terms correspond to operational costs of steam, water and pumping, respectively, and the fourth is the value of product generated. The process has limitations

related to design, product specification and safety:

$$\begin{aligned}
 X_2 &\geq 35.5\% \\
 40\text{kPa} &\leq P_2 \leq 80\text{kPa} \\
 0\text{kg/min} &\leq F_{200} \leq 400\text{kg/min} \\
 8.5\text{kg/min} &\leq F_1 \leq 20\text{kg/min} \\
 0\text{kg/min} &\leq F_3 \leq 100\text{kg/min}
 \end{aligned} \tag{4.21}$$

The lower bound of  $F_1$  was modified from 0 to  $8.5\text{kg/min}$  following [Govatsmark and Skogestad \(2001\)](#) suggestion, to avoid operational instability. There is a backoff of 0.5% in  $X_2$  in order to ensure feasibility for all disturbances.

The process has five control degrees of freedom ( $F_1, F_2, P_{100}, F_3, F_{200}$ ), and three disturbances ( $X_1, T_1, T_{200}$ ). From the five available degrees of freedom, one is consumed to stabilize the separator level. Nominally, the evaporator works with  $X_1 = 5\%$ ,  $T_1 = 40^\circ\text{C}$  and  $T_{200} = 25^\circ\text{C}$ . The disturbances allowable range are the same as [Kariwala, Cao, and Janardhanan \(2008\)](#).

## 4.2.2 Surrogate model generation and optimization

To test the effectiveness of [Caballero and Grossmann \(2008\)](#) algorithm when dealing with complex problems, an initial Latin hypercube sampling of only 53 points was performed and the algorithm could reach the values from the original non-linear model with robust precision as can be seen in [Table 9](#), after adding 27 best points from 148 evaluated. The brute-force infill technique, however, was not able to reach the desired precision. This phenomenon was well-investigated in the past by [Forrester, Sobester, and Keane \(2008\)](#). [Table 10](#) compares the values found for all variables between the kriging prediction and the original model.

Table 9 – Optimization Results of Surrogate responses and using the Original nonlinear Model (Implemented in MATLAB) for Comparison Purposes - Evaporation system case study.

	$F_1$ [kg/min]	$F_3$ [kg/min]	$P_{100}$ [kPa]	$F_{200}$ [kg/min]	$X_2$ [%]	$J$ [\$/h]
Rigorous model	9.469	24.721	400.000	217.738	35.5	-582.233
Kriging	9.469	24.722	400.000	217.785	35.5	-582.233

## 4.2.3 Reduced Space surrogate model

Exactly as the original model, there were two active constraints for the whole disturbance region:  $X_2 = 35.5\%$  and  $P_{100} = 400\text{kPa}$ . From the five degrees of freedom available, the separator level and the active constraints consume three of them, and as

Table 10 – Decision/State Variables and its optimal values: Original implementation and kriging response - Evaporation system case study.

Variable	Description	Units	MATLAB model	Kriging
$F_1$	feed flow rate	kg/min	9.46900	9.46940
$F_2$	product flow rate	kg/min	1.33400	1.33370
$F_3$	circulating flow rate	kg/min	24.7210	24.7225
$F_4$	vapor flow rate	kg/min	8.13500	8.13570
$F_5$	condensate flow rate	kg/min	8.13500	8.13570
$X_1$	feed composition	%	5.00000	5.00000
$X_2$	product composition	%	35.5000	35.5000
$T_1$	feed temperature	°C	40.0000	40.0000
$T_2$	product temperature	°C	88.4000	88.4006
$T_3$	vapor temperature	°C	81.06600	81.0662
$P_2$	operating pressure	kPa	51.41200	51.4126
$F_{100}$	steam flow rate	kg/min	9.43400	9.43460
$T_{100}$	steam temperature	°C	151.5200	151.5200
$P_{100}$	steam pressure	kPa	400.0000	400.0000
$Q_{100}$	heat duty	kW	345.2920	345.3079
$F_{200}$	cooling water flow rate	kg/min	217.73800	217.78500
$T_{200}$	inlet temperature of cooling water	°C	25.00000	25.00000
$T_{201}$	outlet temperature of cooling water	°C	45.55000	45.54610
$Q_{200}$	condenser duty	kW	313.21000	313.22500

in the original model, in reduced space there are two remaining degrees of freedom. The active constraints were implemented in the model, and the reduced surrogate response was built with only a hundred cases. These remaining degrees of freedom can be chosen freely, but to compare the results with the ones from [Kariwala, Cao, and Janardhanan \(2008\)](#) the following were selected:

$$u = [F_{200} F_1]^T \quad (4.22)$$

$$y = [P_2 T_2 T_3 F_2 F_{100} T_{201} F_3 F_5 F_{200} F_1]^T \quad (4.23)$$

#### 4.2.4 High-order Data Obtainment

Similar to case 1, the gradients and Hessians with respect to remaining degrees of freedom and disturbances were found using the DACE toolbox and the expression for the Hessian developed in this study:

$$J_{uu} = \begin{bmatrix} 0.0052 & -0.1201 \\ -0.1201 & 15.0545 \end{bmatrix} \quad J_{ud} = \begin{bmatrix} 0.0185 & 3 \times 10^{-6} & -6.8 \times 10^{-6} \\ -157.6693 & -1.1483 & 1.2717 \end{bmatrix} \quad (4.24)$$

$$G_y = \begin{bmatrix} -0.0932 & 11.6796 \\ -0.0523 & 6.5592 \\ -0.0472 & 5.9215 \\ 0 & 0,1408 \\ -0.0009 & 1.1152 \\ -0.0945 & 2.1713 \\ -0.0318 & 6.5993 \\ 0 & 0,8592 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad G_y^d = \begin{bmatrix} -3.6269 & 0 & 1.9724 \\ -2.0369 & 0 & 1.1077 \\ -1.8388 & 0 & 1 \\ 0.2668 & 0 & 0 \\ -0.3175 & -0.0181 & 0 \\ -0.3175 & 0 & 1 \\ -2.2556 & -0.0657 & 0.6733 \\ -0.2668 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.25)$$

Comparing with the gradients and Hessians obtained by [Kariwala, Cao, and Janardhanan \(2008\)](#) from the original non-linear model:

$$J_{uu} = \begin{bmatrix} 0.006 & -0.133 \\ -0.133 & 16.737 \end{bmatrix} \quad J_{ud} = \begin{bmatrix} 0.023 & 0 & -0.001 \\ -158.373 & -1.161 & 1.484 \end{bmatrix} \quad (4.26)$$

$$G_y = \begin{bmatrix} -0.093 & 11.678 \\ -0.052 & 6.559 \\ -0.047 & 5.921 \\ 0 & 0.141 \\ -0.001 & 1.115 \\ -0.094 & 2.170 \\ -0.031 & 6.594 \\ 0 & 0.859 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad G_y^d = \begin{bmatrix} -3.626 & 0 & 1.972 \\ -2.036 & 0 & 1.108 \\ -1.838 & 0 & 1 \\ 0.267 & 0 & 0 \\ -0.317 & -0.018 & 0.0201 \\ -0.317 & 0 & 1 \\ -2.253 & -0.066 & 0.673 \\ -0.267 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (4.27)$$

It can be clearly seen that the approximation for the matrices is precise, as shown in [Table 11](#):

Table 11 – Matrix Mean-Squared error between gradients and Hessians obtained by [Kariwala, Cao, and Janardhanan \(2008\)](#) and this work - Evaporation system case study.

Matrices: <a href="#">Kariwala, Cao, and Janardhanan (2008)</a> and this work	Matrix Mean-Squared Error (MSE)
$G_y$	$1.6585 \times 10^{-6}$
$G_y^d$	$3.3047 \times 10^{-7}$
$J_{uu}$	0.7060
$J_{ud}$	0.0861

### 4.2.5 Expected Disturbances and Implementation Errors

The weighting matrices for allowable disturbance set and implementation errors were the same as [Kariwala, Cao, and Janardhanan \(2008\)](#):

$$\begin{aligned} W_d &= \text{diag}(0.25, 8, 5) \\ W_n^y &= \text{diag}(1.285, 1, 1, 0.027, 0.189, 1, 0.494, 0.163, 4.355, 0.189) \end{aligned} \quad (4.28)$$

### 4.2.6 Loss evaluation

Using the surrogate response generated and the matrices obtained through DACE and hessian approximation, the best individual measurements (yielding the lowest worst-case and average-case losses) for the two remaining degrees of freedom were:

$$c = [F_3 F_{200}]^T \quad (4.29)$$

With a worst-case and average-case losses of 62.3165 \$/h and 4.1819 \$/h respectively. Using all 10 measurements available, the losses for worst and average cases using the exact local method decrease drastically to 8.3591 \$/h and 0.2160 \$/h, and using the extended null space method of 9.7102 \$/h and 0.2508 \$/h. [Table 12](#) shows some possible subsets of different sizes using the exact local method approach to minimize the worst-case loss and can be compared to the findings of [Kariwala, Cao, and Janardhanan \(2008\)](#) It is important to emphasize that despite the fact that the values found for the losses in each case are slightly numerically higher from the ones evaluated through the rigorous model (as a result of gradient and hessian approximations), the quality of the analysis was not compromised in any way, since the exact same sequence of subsets possible (from lowest to highest losses order) that were found in the work of [Kariwala, Cao, and Janardhanan \(2008\)](#) were also on this study, as can be checked in [Table 12](#).

Considering a “just-enough” measurement subset, using the sub-optimal rule from [Alstad, Skogestad, and Hori \(2009\)](#) finding the subset with size equal to  $n_y = n_u + n_d$  using the branch and bound method with the criterion of maximizing the minimum singular value of  $\widetilde{G}^Y$ , the sub-optimal set found was

$$y = [P_2, F_2, T_{201}, F_3, F_{200}] \quad (4.30)$$

With a worst-case loss of 12.603 \$/h, which can be an option considering the reduced complexity if compared to the approach of using all 10 measurements.

Table 12 – Linear combinations of measurements as CV candidates for Self-Optimizing control and its losses - Order sequence comparison: [Kariwala, Cao, and Janardhanan \(2008\)](#) vs. this study - Evaporation system case study.

	Measurements	Worst-Case Loss		Average-case loss	
		Original Model	Kriging	Original Model	Kriging
2	$F_3, F_{200}$	56.713	62.316	3.808	4.182
	$T_{201}, F_3$	57.140	62.616	4.330	4.677
	$P_2, T_{201}$	57.862	63.447	4.388	4.737
	$F_{100}, F_{200}$	58.370	63.550	3.900	4.744
	$P_2, F_{200}$	58.386	63.598	3.964	4.747
3	$F_2, F_{100}, F_{200}$	11.636	12.700	1.238	0.711
	$F_2, F_{100}, T_{201}$	13.327	13.895	1.124	1.161
	$F_2, T_{201}, F_3$	16.619	18.421	1.143	1.214
	$F_2, F_{100}, T_{201}, F_3$	17.797	18.868	1.565	1.052
4	$F_2, T_{201}, F_3, F_{200}$	9.195	10.359	0.793	0.635
	$F_2, F_{100}, F_5, F_{200}$	9.427	10.501	0.701	0.504
	$F_2, F_{100}, F_5, F_{200}$	9.879	10.734	0.845	0.515
	$F_2, F_{100}, F_3, F_{200}$	10.547	11.833	0.799	0.568

#### 4.2.7 Dynamic simulations

In addition, the best individual measurements policy control structure  $c = [F_3 \ F_{200}]^T$  was implemented in the process and its dynamic performance was evaluated for the expected disturbances: Feed composition ( $X_1$ ), Feed temperature ( $T_1$ ) and steam temperature ( $T_2$ ). The regulatory layer (separator level) and the active constraint X2 had their control loops tuned using the built-in MATLAB/Simulink PID Tune tool. The MV-CV pairing for the regulatory layer and for X2 follows from the work of [Govatsmark and Skogestad \(2001\)](#). The remaining active constraint ( $P_{100}$ ) and both self-optimizing control variables ( $F_3, F_{200}$ ) correspond to manipulated variables and, therefore, were just kept at their nominally optimal values. Important process variables ( $T_2, X_2, P_2$ ) associated with the product were observed:

Figure 12 –  $\pm 5\%$  disturbance in feed molar fraction (5.25% and 4.75%, respectively) - Evaporation system case study.

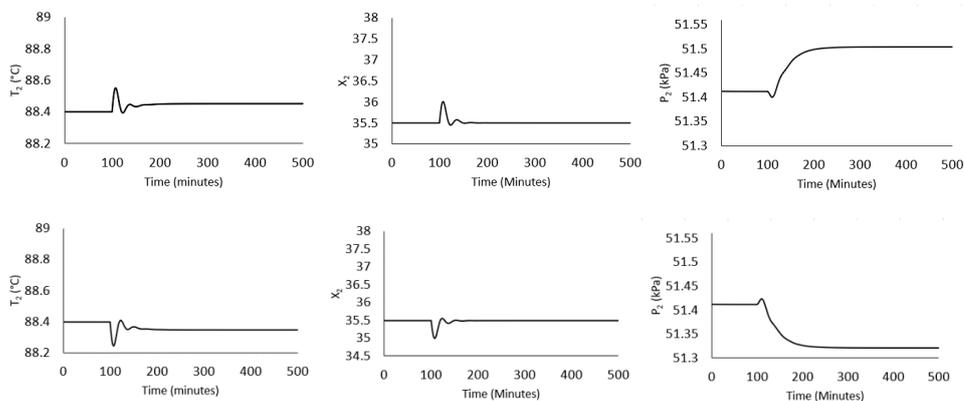


Figure 13 –  $\pm 20\%$  disturbance in feed temperature. ( $48^\circ\text{C}$  and  $32^\circ\text{C}$ , respectively) - Evaporation system case study.

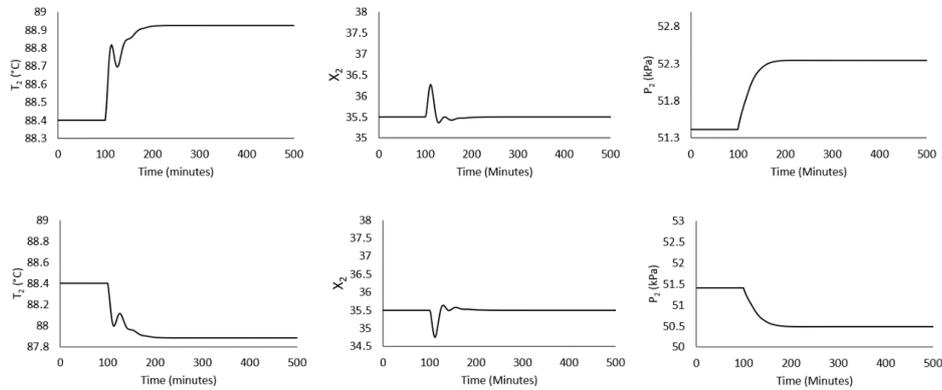
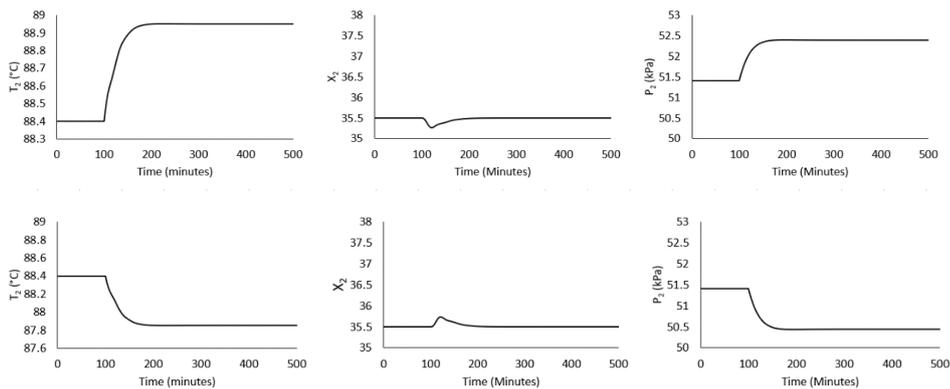


Figure 14 –  $\pm 20\%$  disturbance in steam temperature ( $26.25^\circ\text{C}$  and  $23.75^\circ\text{C}$ , respectively). - Evaporation system case study.



From figures 12-14, it can be seen that the proposed metamodel and SOC-based control structure is capable of keeping the process operating robustly with (near) optimal operation.

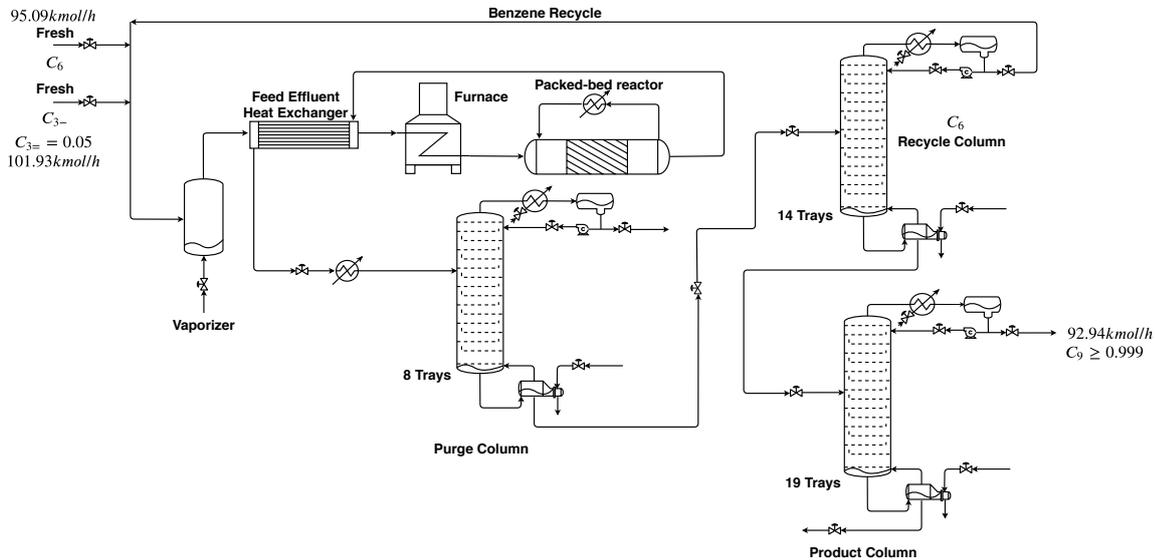
## 4.3 Case III - Economic Self-Optimizing Control of Cumene Process production

### 4.3.1 Problem Description

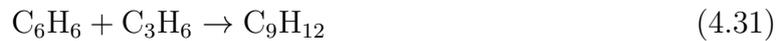
The last problem to be addressed with the proposed methodology of this study consists in the cumene process production, a large-scale case study. The process has been studied in the past by Luyben (2010) and Gera, Panahi, et al. (2013). Fresh/recycled benzene and propylene are mixed and vaporized and sent to a feed effluent heat exchanger. Afterwards, the vapor stream is heated in a furnace to the reaction temperature, and sent

to a PBR reactor. In this reactor, propylene reacts with benzene in vapor phase to produce cumene ( $C_9$ ). However, there is a sequential reaction, producing undesirable di-isopropyl benzene (DIPB/ $C_{12}$ ). The reactor effluent is cooled down in the FEHE and partially condensed in an aftercooler. This stream is then sent to three distillation columns: The purge column separates propane (inert) and unreacted propylene at the top stream, being used as fuel gas. The bottoms is sent to the 2nd column (recycle column): The unreacted benzene from the distillate is recycled and the bottoms of this column is sent to the last (product) column where cumene and DIPB are distilled. The discharge of  $C_{12}$  is also used as fuel. The process flowsheet is available in Figure 15. All simulation conditions were based on the work of Gera, Panahi, et al. (2013). The main process disturbances are the propylene feed and amount of impurity (inert propane) on the propylene feed.

Figure 15 – Cumene process production flowsheet.



The reactions that occur in the PBR reactor (main and side ones) are the following:



For this process there are 11 DOFs. However, some simplifications can be made in order to avoid a problem with poor convergence, as did before by different authors (JAGTAP; KAISTHA, 2012; ARAÚJO; SKOGESTAD, 2008; GERA; PANAHI, et al., 2013). Here we followed the same simplifications from Gera, Panahi, et al. (2013): The reactor effluent aftercooler is maintained at this nominal setpoint ( $100^\circ C$ ), since it has almost no impact in the objective function and will guarantee that the reactor products

will be condensed. The vent stream of the purge column and the compositions of propane at the bottoms of the first column and the composition of cumene leaking at the distillate of the recycle column were also kept fixed by the same reason. Therefore, the problem is reduced to the most impactful variables, remaining with 7 DOFs. They were chosen, without loss of generality, to maximize the number of cases converged at the process simulator, and are described in Table 13.

The objective function for this case study consists in the profit of the unit. All of the information regarding the optimization problem is present Table 13, also based in the work of Gera, Panahi, et al. (2013). There was a slight modification regarding the lower bound for the furnace temperature setpoint: In order to avoid its shutdown, a lower limit of  $350^{\circ}\text{C}$  was imposed.

Table 13 – Cumene process production optimization problem summary.

Objective Function: Profit [ $\$/h$ ]
$J = 150 \times FC_9 - 34.3 \times FC_3 - 68.5 \times FC_6$ $- 9.83 \times (Q_{column1}^{reboiler} + Q_{column2}^{reboiler} + Q_{column3}^{reboiler} + Q_{Vaporizer}) - 16.8 \times Q_{furnace}$ $- 0.16 \times (Q_{column1}^{condenser} + Q_{column2}^{condenser} + Q_{column3}^{condenser} + Q_{aftercooler})$ $+ 36.14 \times F_{DIPB} + 11.1874 \times F_{Vent} + 6.67 \times Q_{reactor}$
Process constraints
$0 \leq$ Material (liquid) flows $\leq 2$ (base case)
$0 \leq V_1, V_2, V_3 \leq 1.5$ (base case)
Vent temperature = $32^{\circ}\text{C}$
$0 \leq$ energy flows $\leq 1.7$ (base case)
$5.5 \text{ bar} \leq P_{Rxt} \leq 25 \text{ bar}$
Cumene product purity $\geq 0.999$ molar fraction
Degrees of freedom
$200 \leq$ Benzene total mole flow $\leq 300$ ( $kmol/h$ )
$350 \leq$ Furnace Temperature $\leq 400$ $^{\circ}\text{C}$
$350 \leq$ Reactor coolant temperature $\leq$ $^{\circ}\text{C}$
$5.5 \leq$ Reactor pressure $\leq 25$ ( $bar$ )
$1.5 \times 10^{-4} \leq$ Recycle column $C_6$ Bottoms composition $\leq 15 \times 10^{-4}$
$0.5 \leq$ Product column bottoms flowrate $\leq 4$ ( $kmol/h$ )
$80 \leq$ Product column boilup rate ( $V_3$ ) $\leq 160$ ( $kmol/h$ )

### 4.3.2 Surrogate model generation and optimization

Considering the process constraints imposed regarding the liquid flows, the boilup limitation on the first two columns (the product column constraint was incorporated as a decision variable), the energy flow constraints and the product specification, there are 19 nonlinear constraints that must have kriging responses built, and the objective function. This totalized 20 surrogate responses that had an initial sample of 200 points generated and provided to the infill algorithm from Caballero and Grossmann (2008).

The nominally optimal operating point was calculated and its results were compared with an optimization performed within the process simulator (Aspen Plus) in Equation-

Oriented Mode. One can see that the results are extremely close in a qualitative (active constraints found) and quantitative point of views.

The cumene product specification and the reactor pressure were found as active constraints, a result that actually could be anticipated due to the economic nature of the objective function, as well discussed in the past by Skogestad (2000) and Minasidis, Skogestad, and Kaistha (2015). However, we chose to keep them as decision variables to observe how the proposed methodology of this work would perform in a case that has a high dimensionality and it is extremely nonlinear/large-scale. Another active constraint was found was the boilup of the recycle column (V2): A result that has been found in the past by Gera, Panahi, et al. (2013), and that will maximize the production of the valuable product. The last active constraint found (Furnace Temperature) aims to energy consumption reduction, and could also be expected to happen. A summary of the optimization results and the comparison with an optimization run performed using the process simulator directly is found in Table 14.

Table 14 – Optimization Results of Surrogates and Original Model - Cumene process production case study.

Variable	Aspen Plus ®	Kriging
Profit [\$/h]	3991.1	3997.4
Benzene total mole flow ( <i>kmol/h</i> )	292.5200	288.2961
Furnace Temperature °C (active constraint)	350	350
Reactor coolant temperature °C	360.364	356.6269
Reactor pressure ( <i>bar</i> ) (active constraint)	25	25
Recycle column $C_6$ Bottoms composition	$6 \times 10^{-4}$	$5 \times 10^{-4}$
Product column bottoms flowrate ( <i>kmol/h</i> )	0.88051	0.83620
Product column boilup rate (V3) ( <i>kmol/h</i> )	95.5514	93.2532
Recycle column boilup rate (V2) ( <i>kmol/h</i> ) (active constraint)	211.38	211.38
Product column cumene purity (active constraint)	0.999	0.999

### 4.3.3 Reduced Space Surrogate Model

For this case study, the reduced-space problem boils down to 3 unconstrained degrees of freedom. The active constraints were implemented in the process simulator and another design of experiments was performed, now including the disturbances as independent variables of the kriging responses. 200 cases with an amplitude of  $\pm 0.5\%$  were sampled. The unconstrained degrees of freedom chosen were the remaining from the original decision variables from the optimization problem: Total benzene molar flowrate, Reactor coolant temperature and product column boilup flowrate.

The measurements considered as CV candidates are available in Table 15. They were chosen based on the work of Gera, Panahi, et al. (2013) and considering variables that are typically available in industrial processes.

Table 15 – Cumene process production CV Candidates.

Variable	Description
$Y_1$	Benzene Total Flowrate ( $kmol/h$ )
$Y_2$	Reactor Coolant Temperature $^{\circ}C$
$Y_3$	Product Column Benzene Distillate molar fraction
$Y_4$	Product Column DIPB Distillate molar fraction
$Y_5$	Product Column Stage 16 Temperature $^{\circ}C$
$Y_6$	Product Column Stage 17 Temperature $^{\circ}C$
$Y_7$	Product Column Boilup flowrate $^{\circ}C$
$Y_8$	Product Column Reflux Ratio
$Y_9$	Reactor Benzene to Propylene ratio
$Y_{10}$	Product Column Reflux flowrate ( $kmol/h$ )
$Y_{11}$	Product Column Reflux to feed ratio
$Y_{12}$	Product Column Boilup to feed ratio
$Y_{13}$	Product Column Cumene Distillate molar fraction
$Y_{14}$	Furnace Temperature $^{\circ}C$
$Y_{15}$	Recycle Column Boilup flowrate ( $kmol/h$ )
$Y_{16}$	Reactor pressure ( $bar$ )

### 4.3.4 High-order data obtainment

The gradients and Hessians with respect to the degrees of freedom of the unconstrained problem and expected disturbances were extracted. Using the process simulator in Equation-Oriented mode, it was possible to evaluate the gradients with respect to the unconstrained degrees of freedom and disturbances, for comparison purposes with the high-order data obtained using the metamodel built, as can be seen on tables 16-17. Once again, there is an excellent agreement between the gradients evaluated using the process simulator and using metamodels.

Table 16 – High-order data obtainment: *Aspen Plus vs Kriging* - Cumene Process case study.

	$G^y$			$G_d^y$	
<i>Kriging</i>	$\begin{bmatrix} 78.2356 & 0.0726 & 9.7728 \times 10^{-5} \\ 4.7267 \times 10^{-10} & 1.0000 & -4.8129 \times 10^{-10} \\ 2.0092 \times 10^{-5} & -3.8512 \times 10^{-7} & -1.2307 \times 10^{-8} \\ -2.0092 \times 10^{-5} & 3.8512 \times 10^{-7} & 1.2307 \times 10^{-8} \\ -1.9317 & -0.9156 & 7.0635 \\ -1.5111 & -0.7203 & 5.5333 \\ 5.9048 \times 10^{-9} & -9.9608 \times 10^{-10} & 1.0000 \\ 0.0014 & 0.0006 & 0.0064 \\ 0.0103 & 8.7878 \times 10^{-5} & -2.0154 \times 10^{-8} \\ 0.1304 & 0.0563 & 0.6114 \\ 0.0014 & 0.0005 & 0.0064 \\ -3.0034 \times 10^{-5} & -0.0002 & 0.0104 \end{bmatrix}$	$\begin{bmatrix} -0.1570 & 0.0678 \\ 1.4908 \times 10^{-9} & -1.7712 \times 10^{-8} \\ -1.5660 \times 10^{-6} & -6.2690 \times 10^{-7} \\ 1.5660 \times 10^{-6} & 6.2690 \times 10^{-7} \\ -7.4864 & 0.4184 \\ -5.8679 & 0.3126 \\ -1.6463 \times 10^{-9} & 9.7379 \times 10^{-8} \\ -0.0054 & -0.0002 \\ -0.0306 & 7.4593 \times 10^{-5} \\ -0.2116 & -0.0197 \\ -0.0054 & -0.0002 \\ -0.0099 & 0.0001 \end{bmatrix}$			
<i>Aspen Plus</i>	$\begin{bmatrix} 78.2337 & 0.0856 & 0 \\ 0 & 1 & 0 \\ 2.0136 \times 10^{-5} & -4.2035 \times 10^{-7} & -1.3748 \times 10^{-8} \\ -2.0136 \times 10^{-5} & 4.2035 \times 10^{-7} & 1.3748 \times 10^{-8} \\ -1.9338 & -0.9706 & 7.0767 \\ -1.4933 & -0.7533 & 5.4699 \\ 0 & 0 & 1 \\ 0.0013 & 0.0005 & 0.0065 \\ 0.0103 & 0.0001 & 0 \\ 0.1282 & 0.0535 & 0.6208 \\ 0.0013 & 0.0005 & 0.0065 \\ -2.2983 \times 10^{-5} & -0.0003 & 0.0105 \end{bmatrix}$	$\begin{bmatrix} -0.1575 & 0.0869 \\ 0 & 0 \\ -1.5721 \times 10^{-6} & -6.8428 \times 10^{-7} \\ 1.5721 \times 10^{-6} & 6.8428 \times 10^{-7} \\ -7.4895 & 0.4390 \\ -5.7926 & 0.3402 \\ 0 & 0 \\ -0.0055 & -0.0002 \\ -0.0306 & 0.0001 \\ -0.2222 & -0.0195 \\ -0.0055 & -0.0001 \\ -0.0099 & 0.0002 \end{bmatrix}$			

Table 17 – Matrix Mean-Squared error between gradients and Hessians obtained using the process simulator and metamodels - Cumene process production case study.

Matrices: Aspen Plus and metamodels	Matrix Mean-Squared Error (MSE)
$G_y$	$2.4736 \times 10^{-4}$
$G_d^y$	$3.0605 \times 10^{-4}$

### 4.3.5 Expected Disturbances and Implementation Errors

For temperature measurements, it was considered a  $0.5^\circ C$  error, associated with typical process instrumentation. For flows and ratios, a value of  $1 \times 10^{-3}$  to represent

the error of flow sensors was used. For composition measurements, a value of  $1 \times 10^{-6}$  (1ppm) was considered. For the expected disturbances, up to 5% of the nominal individual flowrates of propylene and propane at the fresh propylene feed were considered. Therefore, the weighting matrices are

$$W_d = \text{diag}(4.8417, 0.2548) \quad (4.33)$$

$$W_n^y = \text{diag}(0.0010, 0.5000, 1 \times 10^{-6}, 1 \times 10^{-3}, 0.5000, \\ 0.5000, 1 \times 10^{-3}, 1 \times 10^{-3}, 1 \times 10^{-3}, \\ 1 \times 10^{-3}, 1 \times 10^{-3}, 1 \times 10^{-3}) \quad (4.34)$$

### 4.3.6 Loss Evaluation

The loss evaluation for this case study was performed for linear combinations using 3 and 4 measurements, and the five best structures for each subset size can be inspected in [Table 18](#):

Table 18 – Worst and Average-case losses evaluation - Cumene Process Production Case Study.

	Measurements	Worst-case loss [\$/h]	Average-case loss [\$/h]
3	$Y_2, Y_3, Y_6$	0.7604	0.0521
	$Y_2, Y_3, Y_5$	0.7607	0.0521
	$Y_1, Y_2, Y_6$	0.8490	0.0580
	$Y_1, Y_2, Y_5$	0.8494	0.0580
	$Y_2, Y_3, Y_{12}$	0.8807	0.0602
4	$Y_1, Y_7, Y_9, Y_{10}$	0.0040	0.0002
	$Y_3, Y_7, Y_9, Y_{10}$	0.0043	0.0003
	$Y_1, Y_5, Y_7, Y_{10}$	0.0072	0.0004
	$Y_3, Y_5, Y_7, Y_{10}$	0.0087	0.0005
	$Y_1, Y_6, Y_7, Y_{10}$	0.0117	0.0007

As can be seen in [Table 18](#), Control structures that use sensitive tray temperatures associated with compositions, such as the best control structure for the subset of size 3, are promising self-optimizing control structures. It was expected that control structures with compositions measurements would be among the best structures due to the economic impact of directly maintaining such variables at optimal setpoints. However, there are interesting control structures that could be applied in a decentralized fashion, such as the third or fourth best ones

$$Y_a = (Y_1, Y_2, Y_6) \quad (4.35)$$

$$Y_b = (Y_1, Y_2, Y_5) \quad (4.36)$$

That keeps constant the total amount of benzene recycled back to the reaction section, the reactor coolant temperature and a sensitive temperature from the product column, and both have a small incurred loss.

## Part IV

### Conclusions



## 5 Conclusions

This study aimed at developing a systematical procedure to the selection of control structures (SOC-Based) of chemical processes using surrogate responses of them. The final objective is that the procedure described is capable of simplifying and reducing the complexity of evaluating control structures in these processes. The use of the surrogate responses, as can be seen in the three case studies, guaranteed the reduction of the complexity while the precision was not heavily compromised. This can be a powerful tool unlocking the SOC technology for processes that were not feasible to be analyzed in the past, when the evaluation of Hessians and gradients, mandatory to implement the exact local method (From Halvorsen et al. (2003)) and the extended null space method (From Alstad, Skogestad, and Hori (2009)), are cumbersome and/or even infeasible using the original (also called rigorous) models.

In chapter 2 the main aspects of the self-optimizing control technology were presented, and the problem of evaluating high-order data in order to use the local methods presented was raised to the reader. In addition, the main aspects of the response surface methodology (the usage of kriging interpolators and the infill criteria to optimize surrogate responses) were formally discussed and derived as a proposition to solve the aforementioned problem.

In chapters 3-4, three case-studies that were studied before by several authors were tested using the proposed methodology. For the first case study, a depropanizer column proposed by Skogestad (2000), it was shown that the same order of best control structures was found using the surrogate responses to optimize the process and to extract the high-order data. The infill criteria was executed both manually (“brute-force” approach) and using the algorithm proposed by Caballero and Grossmann (2008), and an important result could be found: That a large number of points will not be always necessary to find the optimal operating point with precision. This result is of great importance due to the fact that there are processes that have large dimensionality and are extremely non-linear (case study 3 it is a good example, for instance), and they can be time consuming at the sampling step. Therefore, there is a good approach of providing an initial sample not extremely large and let the algorithm of Caballero and Grossmann (2008) interactively add more sampled points based on the optimization of the response surface of the objective function.

The second case study of an evaporation process it is of particular interest because it was built by Kariwala, Cao, and Janardhanan (2008) using the symbolic math toolbox from MATLAB, and they could extract the high-order data (Gradients and Hessians)

directly from it. Therefore, it is a great case study to directly compare if the high-order data generated by the analytical expressions for the first and second derivatives of the kriging predictor are precise. It was shown that they were not only extremely close, but also as a result of that, the exact same order of control structures for single measurement policy and linear combinations of measurements were found, compared with the original study from [Kariwala, Cao, and Janardhanan \(2008\)](#). This result also corroborates the promising results of the proposed methodology.

The last case study, a large-scale cumene process production, was built to show how the proposed methodology was going to perform in a large scale example, and the results were satisfactory. The variables that were expected to be active constraints were found (due to the economic nature of the objective function), and the best CV candidates found were in compliance with results found previously in the literature. In addition, we used the process simulator to compare the gradients obtained by the kriging interpolator, with the results generated by the former, and once again the results were extremely close: An additional result in favor of the proposed methodology.

# References

- AL., Warren D. Seider et. **Product and Process Design Principles: Synthesis, Analysis and Evaluation**. 4. ed. [S.l.]: Wiley, 2016. ISBN 978-1119355243, 1119355249.
- ALEXANDROV, N et al. Optimization with variable-fidelity models applied to wing design. In: 38TH aerospace sciences meeting and exhibit. [S.l.: s.n.], 2000. P. 841.
- ALSTAD, Vidar; SKOGESTAD, Sigurd. Null Space Method for Selecting Optimal Measurement Combinations as Controlled Variables. **Industrial & Engineering Chemistry Research**, v. 46, n. 3, p. 846–853, 2007. DOI: [10.1021/ie060285+](https://doi.org/10.1021/ie060285+). eprint: <https://doi.org/10.1021/ie060285+>. Available from: <https://doi.org/10.1021/ie060285+>.
- ALSTAD, Vidar; SKOGESTAD, Sigurd; HORI, Eduardo S. Optimal measurement combinations as controlled variables. **Journal of Process Control**, v. 19, n. 1, p. 138–148, 2009. ISSN 0959-1524. DOI: <https://doi.org/10.1016/j.jprocont.2008.01.002>. Available from: <http://www.sciencedirect.com/science/article/pii/S0959152408000073>.
- ALVES, Victor M. C. et al. Metamodel-Based Numerical Techniques for Self-Optimizing Control. **Industrial & Engineering Chemistry Research**, v. 57, n. 49, p. 16817–16840, 2018. DOI: [10.1021/acs.iecr.8b04337](https://doi.org/10.1021/acs.iecr.8b04337). eprint: <https://doi.org/10.1021/acs.iecr.8b04337>. Available from: <https://doi.org/10.1021/acs.iecr.8b04337>.
- ARAUJO, Antonio C. B. de; SHANG, Helen. Enhancing a Smelter Off-Gas System Using a Plant-Wide Control Design. **Industrial & Engineering Chemistry Research**, v. 48, n. 6, p. 3004–3013, 2009. DOI: [10.1021/ie800254s](https://doi.org/10.1021/ie800254s). eprint: <https://doi.org/10.1021/ie800254s>. Available from: <https://doi.org/10.1021/ie800254s>.
- ARAUJO, Antonio; SKOGESTAD, Sigurd. Control structure design for the ammonia synthesis process. **Computers & Chemical Engineering**, v. 32, n. 12, p. 2920–2932, 2008. ISSN 0098-1354. DOI: <https://doi.org/10.1016/j.compchemeng.2008.03.001>. Available from: <http://www.sciencedirect.com/science/article/pii/S0098135408000392>.
- ARAUJO, Antonio C.B. de; GOVATSMARK, Marius; SKOGESTAD, Sigurd. Application of plantwide control to the HDA process. I—steady-state optimization and self-optimizing control. **Control Engineering Practice**, v. 15, n. 10, p. 1222–1237, 2007. Special Issue - International Symposium on Advanced Control of Chemical Processes (ADCHEM). ISSN 0967-0661. DOI: <https://doi.org/10.1016/j.conengprac.2006.10.014>.

Available from:

<<http://www.sciencedirect.com/science/article/pii/S0967066106001997>>.

CABALLERO, José A.; GROSSMANN, Ignacio E. An algorithm for the use of surrogate models in modular flowsheet optimization. **AICHE Journal**, v. 54, n. 10, p. 2633–2650, 2008. DOI: [10.1002/aic.11579](https://doi.org/10.1002/aic.11579). eprint:

<https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.11579>. Available from: <<https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.11579>>.

CAO, Yi; KARIWALA, Vinay. Bidirectional branch and bound for controlled variable selection: Part I. Principles and minimum singular value criterion. **Computers & Chemical Engineering**, v. 32, n. 10, p. 2306–2319, 2008. ISSN 0098-1354. DOI:

<https://doi.org/10.1016/j.compchemeng.2007.11.011>. Available from:

<<http://www.sciencedirect.com/science/article/pii/S0098135407002906>>.

CAO, Yi; ROSSITER, Diane; OWENS, David H. Globally Optimal Control Structure Selection Using Branch and Bound Method. **IFAC Proceedings Volumes**, v. 31, n. 11, p. 185–190, 1998. 5th IFAC Symposium on Dynamics and Control of Process Systems 1998 (DYCOPS 5), Corfu, Greece, 8–10 June 1998. ISSN 1474-6670. DOI:

[https://doi.org/10.1016/S1474-6670\(17\)44926-3](https://doi.org/10.1016/S1474-6670(17)44926-3). Available from:

<<http://www.sciencedirect.com/science/article/pii/S1474667017449263>>.

CAO, Yi; SAHA, Prabirkumar. Improved branch and bound method for control structure screening. **Chemical Engineering Science**, v. 60, n. 6, p. 1555–1564, 2005. ISSN 0009-2509. DOI: <https://doi.org/10.1016/j.ces.2004.10.025>. Available from:

<<http://www.sciencedirect.com/science/article/pii/S0009250904008462>>.

DIXON, D. Christopher. Degrees of Freedom in Dynamic and Static Systems. **Industrial & Engineering Chemistry Fundamentals**, v. 11, n. 2, p. 198–205, 1972. DOI:

[10.1021/i160042a009](https://doi.org/10.1021/i160042a009). eprint: <https://doi.org/10.1021/i160042a009>. Available

from: <<https://doi.org/10.1021/i160042a009>>.

FORRESTER, Alexander; SOBESTER, Andras; KEANE, Andy. **Engineering design via surrogate modelling: a practical guide**. [S.l.]: John Wiley & Sons, 2008.

GERA, Vivek; KAISTHA, Nitin, et al. Plantwide Control of a Cumene Manufacture Process. In: PISTIKOPOULOS, E.N.; GEORGIADIS, M.C.; KOKOSSIS, A.C. (Eds.).

**21st European Symposium on Computer Aided Process Engineering**. [S.l.]:

Elsevier, 2011. v. 29. (Computer Aided Chemical Engineering). P. 522–526. DOI:

<https://doi.org/10.1016/B978-0-444-53711-9.50105-X>. Available from:

<<http://www.sciencedirect.com/science/article/pii/B978044453711950105X>>.

GERA, Vivek; PANAH, Mehdi, et al. Economic Plantwide Control of the Cumene Process. **Industrial & Engineering Chemistry Research**, v. 52, n. 2, p. 830–846,

2013. DOI: [10.1021/ie301386h](https://doi.org/10.1021/ie301386h). eprint: <https://doi.org/10.1021/ie301386h>. Available from: <https://doi.org/10.1021/ie301386h>.
- GOVATSMARK, Marius S.; SKOGESTAD, Sigurd. Control structure selection for an evaporation process. In: GANI, Rafiqul; JØRGENSEN, Sten Bay (Eds.). **European Symposium on Computer Aided Process Engineering - 11**. [S.l.]: Elsevier, 2001. v. 9. (Computer Aided Chemical Engineering). P. 657–662. DOI: [https://doi.org/10.1016/S1570-7946\(01\)80104-8](https://doi.org/10.1016/S1570-7946(01)80104-8). Available from: <http://www.sciencedirect.com/science/article/pii/S1570794601801048>.
- HALVORSEN, Ivar J. et al. Optimal Selection of Controlled Variables. **Industrial & Engineering Chemistry Research**, v. 42, n. 14, p. 3273–3284, 2003. DOI: [10.1021/ie020833t](https://doi.org/10.1021/ie020833t). eprint: <https://doi.org/10.1021/ie020833t>. Available from: <https://doi.org/10.1021/ie020833t>.
- HORI, E.S.; SKOGESTAD, S. Control structure selection of a deethanizer column with partial condenser. In: EUROPEAN Congress of Chemical Engineering (ECCE-6). [S.l.: s.n.], 2007. P. 16–20.
- HORI, E.S.; SKOGESTAD, S. Selection of Control Structure and Temperature Location for Two-Product Distillation Columns. **Chemical Engineering Research and Design**, v. 85, n. 3, p. 293–306, 2007. ISSN 0263-8762. DOI: <https://doi.org/10.1205/cherd06115>. Available from: <http://www.sciencedirect.com/science/article/pii/S0263876207730514>.
- HORI, Eduardo S.; SKOGESTAD, Sigurd; ALSTAD, Vidar. Perfect Steady-State Indirect Control. **Industrial & Engineering Chemistry Research**, v. 44, n. 4, p. 863–867, 2005. DOI: [10.1021/ie0497361](https://doi.org/10.1021/ie0497361). eprint: <https://doi.org/10.1021/ie0497361>. Available from: <https://doi.org/10.1021/ie0497361>.
- HORI, Eduardo Shigueo; SKOGESTAD, Sigurd. Selection of Controlled Variables: Maximum Gain Rule and Combination of Measurements. **Industrial & Engineering Chemistry Research**, v. 47, n. 23, p. 9465–9471, 2008. DOI: [10.1021/ie0711978](https://doi.org/10.1021/ie0711978). eprint: <https://doi.org/10.1021/ie0711978>. Available from: <https://doi.org/10.1021/ie0711978>.
- JACOBSEN, Magnus G.; SKOGESTAD, Sigurd. Active Constraint Regions for Optimal Operation of Chemical Processes. **Industrial & Engineering Chemistry Research**, v. 50, n. 19, p. 11226–11236, 2011. DOI: [10.1021/ie2012196](https://doi.org/10.1021/ie2012196). eprint: <https://doi.org/10.1021/ie2012196>. Available from: <https://doi.org/10.1021/ie2012196>.
- JAGTAP, Rahul; KAISTHA, Nitin. Economic Plantwide Control of a C4 Isomerization Process. **Industrial & Engineering Chemistry Research**, v. 51, n. 36, p. 11731–11743, 2012. DOI: [10.1021/ie3001293](https://doi.org/10.1021/ie3001293). eprint:

<https://doi.org/10.1021/ie3001293>. Available from:

<<https://doi.org/10.1021/ie3001293>>.

JAGTAP, Rahul; KAISTHA, Nitin; SKOGESTAD, Sigurd. Plantwide Control for Economic Optimum Operation of a Recycle Process with Side Reaction. **Industrial & Engineering Chemistry Research**, v. 50, n. 14, p. 8571–8584, 2011. DOI:

[10.1021/ie1024358](https://doi.org/10.1021/ie1024358). eprint: <https://doi.org/10.1021/ie1024358>. Available from:

<<https://doi.org/10.1021/ie1024358>>.

JÄSCHKE, Johannes; SKOGESTAD, Sigurd. A Self-Optimizing Strategy for Optimal Operation of a Preheating Train for a Crude Oil Unit. In: KLEMEŠ, Jiří Jaromír; VARBANOV, Petar Sabev; LIEW, Peng Yen (Eds.). **24th European Symposium on Computer Aided Process Engineering**. [S.l.]: Elsevier, 2014. v. 33. (Computer Aided Chemical Engineering). P. 607–612. DOI:

<https://doi.org/10.1016/B978-0-444-63456-6.50102-2>. Available from:

<<http://www.sciencedirect.com/science/article/pii/B9780444634566501022>>.

JÄSCHKE, Johannes; SKOGESTAD, Sigurd. Optimal controlled variables for polynomial systems. **Journal of Process Control**, v. 22, n. 1, p. 167–179, 2012. ISSN 0959-1524.

DOI: <https://doi.org/10.1016/j.jprocont.2011.09.005>. Available from:

<<http://www.sciencedirect.com/science/article/pii/S0959152411001909>>.

JENSEN, Jørgen Bauck; SKOGESTAD, Sigurd. Optimal operation of simple refrigeration cycles: Part II: Selection of controlled variables. **Computers & Chemical Engineering**, v. 31, n. 12, p. 1590–1601, 2007. ISSN 0098-1354. DOI:

<https://doi.org/10.1016/j.compchemeng.2007.01.008>. Available from:

<<http://www.sciencedirect.com/science/article/pii/S0098135407000233>>.

JONES, Donald R. A taxonomy of global optimization methods based on response surfaces. **Journal of global optimization**, Springer, v. 21, n. 4, p. 345–383, 2001.

JONES, Donald R.; SCHONLAU, Matthias; WELCH, William J. Efficient Global Optimization of Expensive Black-Box Functions. **Journal of Global Optimization**, v. 13, n. 4, p. 455–492, Dec. 1998. ISSN 1573-2916. DOI: [10.1023/A:1008306431147](https://doi.org/10.1023/A:1008306431147).

Available from: <<https://doi.org/10.1023/A:1008306431147>>.

KARIWALA, Vinay; CAO, Yi. Bidirectional branch and bound for controlled variable selection. Part II: Exact local method for self-optimizing control. **Computers & chemical engineering**, Elsevier, v. 33, n. 8, p. 1402–1412, 2009.

KARIWALA, Vinay; CAO, Yi; JANARDHANAN, S. Local Self-Optimizing Control with Average Loss Minimization. **Industrial & Engineering Chemistry Research**, v. 47, n. 4, p. 1150–1158, 2008. DOI: [10.1021/ie070897+](https://doi.org/10.1021/ie070897+). eprint:

<https://doi.org/10.1021/ie070897+>. Available from:

<<https://doi.org/10.1021/ie070897+>>.

- KHANAM, Ambari; SHAMSUZZOHA, Mohammad; SKOGESTAD, Sigurd. Optimal Operation and Control of Divided Wall Column. In: KLEMEŠ, Jiří Jaromír; VARBANOV, Petar Sabev; LIEW, Peng Yen (Eds.). **24th European Symposium on Computer Aided Process Engineering**. [S.l.]: Elsevier, 2014. v. 33. (Computer Aided Chemical Engineering). P. 673–678. DOI: <https://doi.org/10.1016/B978-0-444-63456-6.50113-7>. Available from: <http://www.sciencedirect.com/science/article/pii/B9780444634566501137>.
- KLEIJNEN, Jack P.C. Kriging metamodeling in simulation: A review. **European Journal of Operational Research**, v. 192, n. 3, p. 707–716, 2009. ISSN 0377-2217. DOI: <https://doi.org/10.1016/j.ejor.2007.10.013>. Available from: <http://www.sciencedirect.com/science/article/pii/S0377221707010090>.
- KONDA, N.V.S.N. Murthy; RANGAIAH, G.P.; KRISHNASWAMY, P.R. A simple and effective procedure for control degrees of freedom. **Chemical Engineering Science**, v. 61, n. 4, p. 1184–1194, 2006. ISSN 0009-2509. DOI: <https://doi.org/10.1016/j.ces.2005.08.026>. Available from: <http://www.sciencedirect.com/science/article/pii/S0009250905006998>.
- KRIGE, Daniel G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Southern African Institute of Mining and Metallurgy**, Southern African Institute of Mining and Metallurgy, v. 52, n. 6, p. 119–139, 1951.
- KUSHNER, H. J. A New Method of Locating the Maximum Point of an Arbitrary Multipipeak Curve in the Presence of Noise. **Journal of Basic Engineering**, v. 86, n. 1, p. 97–106, Mar. 1964. ISSN 0021-9223. DOI: [10.1115/1.3653121](https://doi.org/10.1115/1.3653121). eprint: [https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/86/1/97/5763745/97\\_1.pdf](https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/86/1/97/5763745/97_1.pdf). Available from: <https://doi.org/10.1115/1.3653121>.
- LAWLER, E. L.; WOOD, D. E. Branch-and-Bound Methods: A Survey. **Operations Research**, v. 14, n. 4, p. 699–719, 1966. DOI: [10.1287/opre.14.4.699](https://doi.org/10.1287/opre.14.4.699). eprint: <https://doi.org/10.1287/opre.14.4.699>. Available from: <https://doi.org/10.1287/opre.14.4.699>.
- LERSBAMRUNGSUK, Veerayut et al. Control structure design for optimal operation of heat exchanger networks. **AIChE Journal**, v. 54, p. 150–162, Jan. 2008. DOI: [10.1002/aic.11366](https://doi.org/10.1002/aic.11366).
- LOPHAVEN, Søren Nymand; NIELSEN, Hans Bruun; SØNDERGAARD, Jacob. **DACE - A Matlab Kriging Toolbox, Version 2.0**. [S.l.: s.n.], 2002.

- LUYBEN, William L. Design and Control Degrees of Freedom. **Industrial & Engineering Chemistry Research**, v. 35, n. 7, p. 2204–2214, 1996. DOI: [10.1021/ie960038d](https://doi.org/10.1021/ie960038d). eprint: <https://doi.org/10.1021/ie960038d>. Available from: <https://doi.org/10.1021/ie960038d>.
- LUYBEN, William L. Design and Control of the Cumene Process. **Industrial & Engineering Chemistry Research**, v. 49, n. 2, p. 719–734, 2010. DOI: [10.1021/ie9011535](https://doi.org/10.1021/ie9011535). eprint: <https://doi.org/10.1021/ie9011535>. Available from: <https://doi.org/10.1021/ie9011535>.
- LUYBEN, William L. **Distillation Design and Control Using Aspen Simulation**. 2. ed. [S.l.]: Wiley-AIChE, 2013. ISBN 1118411439,9781118411438.
- LUYBEN, William L. Evaluation of criteria for selecting temperature control trays in distillation columns. **Journal of Process Control**, v. 16, n. 2, p. 115–134, 2006. ISSN 0959-1524. DOI: <https://doi.org/10.1016/j.jprocont.2005.05.004>. Available from: <http://www.sciencedirect.com/science/article/pii/S0959152405000569>.
- MINASIDIS, Vladimiro; SKOGESTAD, Sigurd; KAISTHA, Nitin. Simple Rules for Economic Plantwide Control. In: GERNAEY, Krist V.; HUUSOM, Jakob K.; GANI, Rafiqul (Eds.). **12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering**. [S.l.]: Elsevier, 2015. v. 37. (Computer Aided Chemical Engineering). P. 101–108. DOI: <https://doi.org/10.1016/B978-0-444-63578-5.50013-X>. Available from: <http://www.sciencedirect.com/science/article/pii/B978044463578550013X>.
- MORARI, Manfred; ARKUN, Yaman; STEPHANOPOULOS, George. Studies in the synthesis of control structures for chemical processes: Part I: Formulation of the problem. Process decomposition and the classification of the control tasks. Analysis of the optimizing control structures. **AIChE Journal**, Wiley Online Library, v. 26, n. 2, p. 220–232, 1980.
- NARENDRA; FUKUNAGA. A Branch and Bound Algorithm for Feature Subset Selection. **IEEE Transactions on Computers**, v. C-26, n. 9, p. 917–922, Sept. 1977. ISSN 2326-3814. DOI: [10.1109/TC.1977.1674939](https://doi.org/10.1109/TC.1977.1674939).
- NEWELL, R.B.; LEE, P.L. **Applied Process Control: A Case Study**. [S.l.]: Prentice-Hall of Australia, 1989. ISBN 9780130409652.
- NOCEDAL, J.; WRIGHT, S. **Numerical Optimization**. [S.l.]: Springer New York, 2006. (Springer Series in Operations Research and Financial Engineering). ISBN 9780387227429. Available from: <https://books.google.com.br/books?id=7wDpBwAAQBAJ>.

- PANAHI, Mehdi; SKOGESTAD, Sigurd. Selection of Controlled Variables for a Natural Gas to Liquids Process. **Industrial & Engineering Chemistry Research**, v. 51, n. 30, p. 10179–10190, 2012. DOI: [10.1021/ie202678h](https://doi.org/10.1021/ie202678h). eprint: <https://doi.org/10.1021/ie202678h>. Available from: <https://doi.org/10.1021/ie202678h>.
- PHAM, Q.T. Degrees of freedom of equipment and processes. **Chemical Engineering Science**, v. 49, n. 15, p. 2507–2512, 1994. ISSN 0009-2509. DOI: [https://doi.org/10.1016/0009-2509\(94\)E0056-V](https://doi.org/10.1016/0009-2509(94)E0056-V). Available from: <http://www.sciencedirect.com/science/article/pii/0009250994E0056V>.
- PONTON, Jack W. Degrees of freedom analysis in process control. **Chemical Engineering Science**, v. 49, n. 13, p. 2089–2095, 1994. ISSN 0009-2509. DOI: [https://doi.org/10.1016/0009-2509\(94\)E0033-M](https://doi.org/10.1016/0009-2509(94)E0033-M). Available from: <http://www.sciencedirect.com/science/article/pii/0009250994E0033M>.
- QUIRANTE, Natalia; JAVALOYES, Juan; CABALLERO, José A. Rigorous design of distillation columns using surrogate models based on Kriging interpolation. **AIChE Journal**, v. 61, n. 7, p. 2169–2187, 2015. DOI: [10.1002/aic.14798](https://doi.org/10.1002/aic.14798). eprint: <https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.14798>. Available from: <https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.14798>.
- SACKS, Jerome et al. Design and analysis of computer experiments. **Statistical science**, JSTOR, p. 409–423, 1989.
- SAHA, Prabirkumar; CAO, Yi. Globally optimal control structure selection using Hankel singular value through branch and bound method. In: CHEN, Bingzhen; WESTERBERG, Arthur W. (Eds.). **Process Systems Engineering 2003, 8th International Symposium on Process Systems Engineering**. [S.l.]: Elsevier, 2003. v. 15. (Computer Aided Chemical Engineering). P. 1014–1019. DOI: [https://doi.org/10.1016/S1570-7946\(03\)80441-8](https://doi.org/10.1016/S1570-7946(03)80441-8). Available from: <http://www.sciencedirect.com/science/article/pii/S1570794603804418>.
- SHINSKEY, F.G. **Distillation control: for productivity and energy conservation**. [S.l.]: McGraw-Hill, 1984. ISBN 9780070568945.
- SILVA, S. K. et al. Development and application of an automatic tool for the selection of control variables based on the self-optimizing control methodology. en. **Brazilian Journal of Chemical Engineering**, scielo, v. 34, p. 851–871, July 2017. ISSN 0104-6632. Available from: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0104-66322017000300851&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-66322017000300851&nrm=iso).

- SKOGESTAD, Sigurd. Control structure design for complete chemical plants. **Computers & Chemical Engineering**, v. 28, n. 1, p. 219–234, 2004. Escape 12. ISSN 0098-1354. DOI: <https://doi.org/10.1016/j.compchemeng.2003.08.002>. Available from:  
<<http://www.sciencedirect.com/science/article/pii/S0098135403001984>>.
- SKOGESTAD, Sigurd. Plantwide control: the search for the self-optimizing control structure. **Journal of Process Control**, v. 10, n. 5, p. 487–507, 2000. ISSN 0959-1524. DOI: [https://doi.org/10.1016/S0959-1524\(00\)00023-8](https://doi.org/10.1016/S0959-1524(00)00023-8). Available from:  
<<http://www.sciencedirect.com/science/article/pii/S0959152400000238>>.
- SKOGESTAD, Sigurd. Plantwide control: Towards a systematic procedure. In: GRIEVINK, Johan; SCHIJNDEL, Jan van (Eds.). **European Symposium on Computer Aided Process Engineering-12**. [S.l.]: Elsevier, 2002. v. 10. (Computer Aided Chemical Engineering). P. 57–69. DOI:  
[https://doi.org/10.1016/S1570-7946\(02\)80039-6](https://doi.org/10.1016/S1570-7946(02)80039-6). Available from:  
<<http://www.sciencedirect.com/science/article/pii/S1570794602800396>>.
- SKOGESTAD, Sigurd; MORARI, Manfred. Control configuration selection for distillation columns. **AIChE Journal**, v. 33, n. 10, p. 1620–1635, 1987. DOI:  
[10.1002/aic.690331006](https://doi.org/10.1002/aic.690331006). eprint:  
<https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.690331006>.  
Available from:  
<<https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690331006>>.
- SKOGESTAD, Sigurd; POSTLETHWAITE, Ian. **Multivariable Feedback Control: Analysis and Design**. 1. ed. [S.l.]: Wiley-Interscience, 1996.
- STEPHANOPOULOS, George. **Chemical Process Control: An Introduction to Theory and Practice**. 1. ed. [S.l.]: Prentice-Hall International Ltd, 2003.
- UMAR, L M et al. Plantwide Control: Recent Developments and Applications. In: ed. by G P Rangaiah and V Kariwala. [S.l.]: John Wiley & Sons, 2012. Selection of Controlled Variables using Self-optimizing Control Method.

# Appendix



# APPENDIX A – Kriging Proofs: Predictor, Gradient, and Hessian

To present the Kriging predictor, we assume that  $q = 1$  (single response), thus the linear predictor takes form of:

$$\hat{y}(x) = c^T Y \quad (\text{A.1})$$

where  $c = c(x) \in \mathbb{R}_m$ . We treat [Equation A.1](#) as random and compute the mean squared error (MSE) of this predictor averaged over the random process. Therefore, the best unbiased linear predictor is obtained by choosing  $c$  in order to minimize:

$$\varphi(x) = \text{MSE}(\hat{y}(x)) = E \left[ (\hat{y}(x) - y(x))^2 \right] \quad (\text{A.2})$$

The  $E[\cdot]$  is the expectation operator. The prediction error is defined by

$$\begin{aligned} \hat{y}(x) - y(x) &= c^T Y - y(x) \\ &= c^T (F\beta + Z) - (f(x)^T \beta + z) \\ &= c^T Z - z + (F^T c - f(x))^T \beta \end{aligned} \quad (\text{A.3})$$

where  $F$  is the expanded design matrix:

$$F_{ij} = f_j(s_i), i = 1, \dots, m \text{ and } j = 1, \dots, h \quad (\text{A.4})$$

with  $f(x)$  defined in [Equation 2.40](#), and  $Z = [z_1, \dots, z_m]^T$  being the error at the design sites.

[Equation A.3](#) is restricted to the unbiasedness constraint:

$$F^T c - f(x) = 0 \quad (\text{A.5})$$

Consequently, combining [Eqs. A.1-A.5](#):

$$\begin{aligned} \varphi(x) &= E \left[ (c^T Z - z)^2 \right] \\ &= E \left[ z^2 + c^T Z Z^T c - 2c^T Z z \right] \\ &= \sigma^2 \left( 1 + c^T R c - 2c^T r \right) \end{aligned} \quad (\text{A.6})$$

where  $\sigma^2$  is the process variance, and  $R$  and  $r$  are defined as

$$\begin{aligned} R_{ij} &= \mathcal{R}(\theta_l, s_i, s_j), \quad i, j = 1, \dots, m \\ r &= r(x) = [\mathcal{R}(\theta_l, s_1, x) \dots \mathcal{R}(\theta_l, s_m, x)]^T \end{aligned} \quad (\text{A.7})$$

To minimize [Equation A.6](#) with respect to  $c$ , we make use of the Lagrangian multipliers in the objective function:

$$L(c, \lambda) = \sigma^2 (1 + c^T R c - 2c^T r) - \lambda^T (F^T c - f(x)) \quad (\text{A.8})$$

The gradient of this function is

$$\nabla_c L(c, \lambda) = \sigma^2 (R c - 2r) - F \lambda \quad (\text{A.9})$$

and the optimality conditions for [Equation A.8](#) are that at the solution  $(c^*, \lambda^*)$  ([NOCEDAL; WRIGHT, 2006](#)):

$$\begin{aligned} \nabla_c L(c^*, \lambda^*) &= 0 \\ \lambda^* (F^T c^* - f(x)) &= 0 \end{aligned} \quad (\text{A.10})$$

From [Equation A.9](#), the following set of equations is obtained:

$$\begin{aligned} \begin{bmatrix} R & F \\ F^T & 0 \end{bmatrix} \begin{bmatrix} c \\ \tilde{\lambda} \end{bmatrix} &= \begin{bmatrix} r \\ f(x) \end{bmatrix} \\ \tilde{\lambda} &= -\frac{\lambda}{2\sigma^2} \end{aligned} \quad (\text{A.11})$$

As a solution to [Equation A.11](#) by calculating its partitioned inverse:

$$\begin{aligned} c &= R^{-1}(r - F\tilde{\lambda}) \\ \tilde{\lambda} &= (F^T R^{-1} F)^{-1} (F^T R^{-1} r - f(x)) \end{aligned} \quad (\text{A.12})$$

By applying [Equation A.12](#) to [Equation A.1](#), the best unbiased linear predictor becomes

$$\hat{y}(x) = r^T R^{-1} Y - (F^T R^{-1} r - f(x))^T (F^T R^{-1} F)^{-1} F^T R^{-1} Y \quad (\text{A.13})$$

The usual generalized least-squares estimate of  $\beta$  is

$$\beta^* = (F^T R^{-1} F)^{-1} F^T R^{-1} Y \quad (\text{A.14})$$

Thus, the predictor is defined by

$$\begin{aligned}\hat{y}(x) &= r^T R^{-1} Y - \left( F^T R^{-1} r - f(x) \right)^T \beta^* \\ &= f(x)^T \beta^* + r^T R^{-1} (Y - F \beta^*) \\ &= f(x)^T \beta^* + r^T \gamma^*\end{aligned}\tag{A.15}$$

and the MSE, by applying [Equation A.12](#) in [Equation A.6](#):

$$\begin{aligned}\varphi(x) &= \sigma^2 \left[ 1 + (F \tilde{\lambda} - r)^T R^{-1} (F \tilde{\lambda} + r) \right] \\ &= \sigma^2 \left[ 1 + \left( F^T R^{-1} r - f \right)^T \left( F^T R^{-1} F \right)^{-1} \left( F^T R^{-1} r - f \right) \right. \\ &\quad \left. - r^T R^{-1} r \right]\end{aligned}\tag{A.16}$$

The gradient and Hessian of the predictor are defined as

$$\begin{aligned}\hat{y}'(x) &= J_f(x)^T \beta^* + J_r(x)^T \gamma^* \\ \hat{y}''(x) &= H_f(x) \beta^* + H_r(x) \gamma^*\end{aligned}\tag{A.17}$$

where

$$(J_f(x))_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \quad (J_r(x))_{ij} = \frac{\partial R(\theta, s_j, x)}{\partial x_j}\tag{A.18}$$

$$(H_f(x))_{ij} = \frac{\partial^2 f_i(x)}{\partial x_j \partial x_j}, \quad (H_r(x))_{ij} = \frac{\partial^2 R(\theta, s_j, x)}{\partial x_j \partial x_j}\tag{A.19}$$

$$\mathcal{R}(\theta, x_i, x_j) = \exp\left(-\theta d_{ij}^p\right), \quad d_{i,j} = x_i - x_j, p \in [0, 2]\tag{A.20}$$

$$\frac{\partial \mathcal{R}(\theta, x_i, x_j)}{\partial x_i} = -\theta p d_{i,j}^{p-1} \mathcal{R}(\theta, x_i, x_j)\tag{A.21}$$

$$\frac{\partial \mathcal{R}(\theta, x_i, x_j)}{\partial x_j} = -\frac{\partial \mathcal{R}(\theta, x_i, x_j)}{\partial x_i}\tag{A.22}$$

$$\frac{\partial^2 \mathcal{R}(\theta, x_i, x_j)}{\partial x_i \partial x_j} = \theta p d_{i,j}^{p-2} \mathcal{R}(\theta, x_i, x_j) \left( p - 1 + \theta_p d_{i,j}^p \right)\tag{A.23}$$

Finally, there is the hyperparameters  $(\theta)$  estimation. This is done by assuming a Gaussian process and maximizing the likelihood  $L$  of  $Y$ :

$$L(x, \theta, \mu) = \frac{1}{(2\pi\sigma^2)^{|R|^{1/2}}} \exp\left(\frac{-(Y - F\beta)^T R^{-1} (Y - F\beta)}{2\sigma^2}\right)\tag{A.24}$$

where the estimator of  $\theta$ ,  $\theta^*$  is

$$\theta^* = \frac{(Y - F\beta^*)^T R^{-1} (Y - F\beta^*)}{m} \quad (\text{A.25})$$