Universidad de Valladolid

ESCUELA DE INGENIERÍAS INDUSTRIALES

DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

TESIS DOCTORAL:

HANDLING UNCERTAINTIES IN PROCESS OPTIMIZATION

Daniel Andrés Navia López

Valladolid, Diciembre 2012

Universidad de Valladolid

ESCUELA DE INGENIERÍAS INDUSTRIALES

DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

TESIS DOCTORAL:

HANDLING UNCERTAINTIES IN PROCESS OPTIMIZATION

Presentada por Daniel Andrés Navia López para optar al grado de doctor
por la Universidad de Valladolid

Dirigida por:

D. César de Prada Moraga

D. Gloria Gutiérrez Rodríguez

UNIVERSIDAD DE VALLADOLID

ESCUELA DE INGENIERÍAS INDUSTRIALES

Secretaría

La presente tesis doctoral queda registrada en el folio

Nº_____

del correspondiente Libro de Registro con el Nº_____

Valladolid, a _____ de _____ de 2012

Fdo. El encargado del Registro

César de Prada Moraga

Catedrático de Universidad del Departamento de Ingeniería de Sistemas y Automática
Universidad de Valladolid

Gloria Gutiérrez Rodríguez

Profesor Contratado Doctor del Departamento de Ingeniería de Sistemas y Automática
Universidad de Valladolid

CERTIFICAN QUE:

DANIEL ANDRÉS NAVIA LÓPEZ ha realizado bajo su dirección el trabajo ***"Handling Uncertainties in Process Optimization"***, en el Departamento de Ingeniería de Sistemas y Automática de la Escuela de Ingenierías Industriales de la Universidad de Valladolid. Considerando que dicho trabajo reúne los requisitos para ser presentado como Tesis Doctoral expresan su conformidad con dicha presentación.

Valladolid a _____ de_____ de 2012.

César de Prada Moraga Gloria Gutiérrez Rodríguez

Reunido el tribunal que ha juzgado la tesis doctoral "**Handling Uncertainties in Process Optimization**" presentada por Daniel Andrés Navia López y en cumplimiento con lo establecido por el Real Decreto 861/2010 (BOE 03.07.2010) ha acordado conceder por _____la calificación de_____.

Valladolid, a de de 2012

PRESIDENTE SECRETARIO

1er Vocal 2do Vocal 3er Vocal

x

Agradecimientos

A la financiación recibida del Séptimo Programa Marco de la Unión Europea [FP7/2007-2013] bajo el contrato de becas nº257462 HYCON2 Red de Excelencia, así como a los recursos otorgados por el proyecto "Control y Optimización de Redes de Distribución MICINN DPI 2009-12805". Al programa de becas Erasmus-Mundus (External Cooperation Window for Chile – Lote 17), por haber hecho posible mi estancia en la Universidad de Valladolid.

A mis directores de tesis César de Prada y Gloria Gutiérrez por el tiempo, la disponibilidad y la confianza depositada en mí para realizar esta tesis. También por la oportunidad de participar activamente con la industria en distintos proyectos. En particular agradecer a César por darme la luz en los momentos oscuros de la tesis así como por todo el conocimiento entregado, espero seguir tu ejemplo en la aventura docente que emprendo.

A los profesores del grupo de investigación en "Control y Supervisión de Procesos" por complementar mi formación y ayudarme a crecer como profesional. A mis compañeros(as) del Departamento de Ingeniería de Sistemas y Automática y del CTA por la compañía y los momentos amenos que me han hecho vivir. En especial agradecer a Daniel y Rubén por su amistad, ayuda y la posibilidad de generar un clima de discusión fecundo de nuevas ideas que espero aplicar en un futuro cercano.

A los amigos hechos en España: Alejandro, Alex, Elisa, Eva, Fernando, Isra, Javi, José, Laura, María, Nacho, Oscar, Raúl y Vanessa, por hacer de mi estadía en Valladolid un período inolvidable y lleno de buenos momentos… son una gran familia adoptiva. A los amigos de Chile: Amory, Andrea, Cristián, Cristián, Jesús, Paulo y Rodrigo, por todo su apoyo y fuerzas para seguir adelante en este proyecto.

A mis padres: Carmen y Jaime por enseñarme el valor del trabajo y darme su apoyo y cariño desde la distancia, gran parte de lo que he logrado se los

debo a ustedes. También a ti Nina por todo tu amor. A mis hermanas, cuñados y sobrinos: Loreto, Mónica, Mauricio, Marcelo, Nicolás, Valentina, Alonso y Gabriela, por estar ahí, acompañándome siempre con una palabra de aliento, un saludo cariñoso y una sonrisa desinteresada. Del mismo modo agradecer a mi familia adoptiva: Mónica, Fernando, Marina, Diana y Carlos por tratarme como uno más de ustedes y siempre darme ánimo.

A Andrea por tu amor, compañía, compresión y todos estos años juntos que han sido los mejores de mi vida (junto con los que vendrán).

Table of Contents

ABSTRACT/RESUMEN

Abstract

Optimization is a powerful tool to be applied in the process industry. It allows obtaining comparative improvements of a given facility finding the operational point that maximizes the profit. In spite of the different applications that concern the use of this tool, the management of uncertainty is a critical issue in order to propose an optimal and feasible solution.

In this thesis, we present the study and application of techniques able to handle the uncertainties from the point of view of random behavior of process variables and errors in the models to be used in the optimization.

Concerning with the random behavior of the process variables, we have focused our study in **Stochastic Optimization** as a tool to be applied in a processes. In particular we have tested the *Two-Stage Optimization* and *Chance Constrained Optimization* methods in a process example from a hydrodesulfuration unit. Applying this techniques, the idea is to propose an optimal and feasible policy to be implemented when there is a change in the load to be treated despite the expected uncertainties. The implementation of these strategies is founded in the fact that both the raw materials and the quality of the product to be transformed are not completely known, and only their probability distribution functions are available.

Regarding the two-stage optimization, we have solved a discrete equivalent problem using a scenario realization of the original probability distribution function, using the scenario aggregation methodology to take into account the nonanticipativity constraint, solving the dynamic optimization problem using a sequential approach.

On the other hand, the chance constrained optimization has been solved using the inverse mapping technique to estimate the probability of the constraints, calculating the equivalent limit on the random variable solving

a parameter estimation problem over the entire time prediction horizon with the single shooting methodology.

Because of the large computational times observed in the resolution of both methods, we have presented an open loop implementation that has been a tested using Montecarlo simulations. Because of the discretization applied in the two-stage approach, we have proposed a generalization method based in the interpolation of the second-stage outcomes to apply this policy in open loop with the original probability distribution function.

The results show that using the implementations described before, we can obtain an optimal trajectory for the load change problem in the hydrodesulfuration process that ensures a given degree of feasibility.

About the management of the modeling errors, we have worked with **Real Time Optimization**, in order to propose the value of the stationary points that optimizes the process in an iterative way. In particular we have focused our attention in the *Modifier-Adaptation Methodology* as a technique to get over the modeling mismatch produced because of the partial knowledge of the system. Using previous information from literature and a simulation example, we have detected three challenges to be addressed in this methodology: infeasible operations produced in intermediate points, problems with the experimental gradient detection and the possibility of using modifiers in dynamic optimization. Each of these topics were studied.

We have proposed an intermediate layer between the RTO and the control one that uses a non-model based controller to modify the value suggested by the RTO algorithm when a violation on the process constraints is detected. If the experimental gradient is calculated using the dual control optimization approach, we have implemented a second controller (the dual one) to manage the excitation level of the process.

With respect to the gradient estimation, we have proposed another way to see the modifier adaptation algorithm, with a methodology that intuitively

finds the optimum of the process directly in the space of the gradient modifiers: the *Nested-Modifier Adaptation* method. In our proposal the decision variables to be applied in the process must be calculated as the outcomes of an inner modified optimization with the same structure as the one solved in the original modifier-adaptation scheme. Nevertheless, the update law of the modifiers is implemented in an upper optimization layer that uses directly the performance index from the process as a cost function.

At last we have presented some preliminary ideas about the implementation of the modifiers to solve dynamic optimization problem in a receding horizon implementation, modifying the NLP equivalent of the problem.

As a result of the implementations proposed in the modifier-adaptation approach, we can say that for the steady-state case it has been increased the field application of this method for those problems where the process gradient is difficult to obtain (or is not available) or when the measurements are contaminated with noise, detecting the real optimum of the process with an erroneous model. With this, we are able to converge in a feasible and robust path to the real optimum. On the other hand, in the dynamic case we have shown that correcting the process measurements with the natural response of system, it is possible to deal with modeling mismatch using the modifiers, provided the objective function can be measured at the end of the prediction horizon.

Resumen

La optimización es una poderosa herramienta en la industria de procesos. Mediante esta es posible obtener ventajas comparativas determinado estimando correctamente el valor de las variables que maximizan el beneficio. Independientemente de la aplicación, el tratamiento de las incertidumbres es un tópico transversal si se desea resolver un problema aplicable y factible.

En esta tesis doctoral, se presenta el estudio y la aplicación de técnicas que permiten manejar estas incertidumbres, desde el punto de vista del comportamiento aleatorio de las variables de proceso y de los errores en los modelos utilizados en la optimización.

Con respecto al tratamiento de las variables de proceso inciertas, se presenta la aplicación de la **Optimización Estocástica** como herramienta para hacer frente al comportamiento aleatorio de las variables de proceso. En particular se ha trabajado con *Optimización Estocástica de dos Etapas* y con *Optimización Probabilística*, aplicada en un ejemplo de un proceso de hidrodesulfuración presente en una refinería de petróleo. Mediante la aplicación de estas técnicas, se pretende encontrar una trayectoria óptima de las variables de decisión, que permita hacer frente a una situación de cambio de alimentación del sistema, asegurando un determinado grado de factibilidad. La idea de utilizar optimización estocástica, se basa en que tanto la calidad de la materia prima utilizada, como la del hidrocarburo a ser tratado no son conocidas completamente y sólo se dispone de su distribución de probabilidad.

Con respecto a la optimización de dos etapas, se ha resuelto un problema simplificado basado en una discretización de la distribución de probabilidad original. Referente a la restricción de no anticipatividad se ha utilizado la técnica de agregación de escenarios para asegurar una única trayectoria durante la primera etapa, resolviendo el problema de optimización dinámico en cada escenario mediante un método secuencial.

En el caso de la optimización probabilística, se ha utilizado el método de mapeo inverso para estimar la probabilidad de cumplir una restricción utilizando la distribución de probabilidad de las variables incierta, siendo necesario resolver un problema de estimación de parámetros sobre todo el horizonte de predicción para obtener el límite equivalente en las variables aleatorias. Este problema se ha abordado usando el método de single shooting.

Debido al alto tiempo computacional observado, se ha propuesto la aplicación de los resultados obtenidos en lazo abierto, para lo cual se han sometido a prueba utilizando simulaciones de Montecarlo. Puesto que se ha resuelto un problema discreto en la optimización de dos etapas, se ha implementado un método basado en interpolaciones para generalizar su solución y así poder aplicar las trayectorias a la función de distribución de probabilidad continua.

Los resultados obtenidos muestran que al tratar explícitamente las incertidumbres mediante las técnicas aplicadas, es posible obtener una solución óptima que garantice un determinado grado de factibilidad en el problema de cambio de carga de una unidad hidrodesulfuradora.

Para el manejo de la incertidumbre derivada del conocimiento parcial de los modelos que rigen un proceso, se ha utilizado el método de **Optimización en Tiempo Real**, el cual consiste en actualizar un modelo con medidas del proceso para proponer de manera iterativa el valor de las variables de decisión que lo optimizan. En particular se ha estudiado el método de adaptación de modificadores como alternativa para encontrar el óptimo de un proceso utilizando un modelo erróneo. A partir de una recopilación bibliográfica y con la ayuda de un ejemplo implementado en simulación se han detectado tres áreas para las cuales se proponen mejoras: convergencia al óptimo con puntos intermedios infactibles, problemas derivados con la estimación del gradiente experimental y la posibilidad de aplicar los modificadores en un problema de optimización dinámica.

Para evitar la ocurrencia de puntos factibles en la evolución del algoritmo de adaptación de modificadores, se ha propuesto e implementado una capa intermedia entre la de optimización y el control de procesos, que corrige el valor sugerido por la optimización en el instante en el que se detecta una infactibilidad, llevando el sistema hasta una región factible mediante la acción de un controlador no basado en modelos. Si el gradiente del proceso ha sido estimado mediante el método de optimización dual se ha implementado un segundo controlador (Dual) que asegura energía suficiente en el problema de estimación de gradientes.

Se ha propuesto una forma de buscar el óptimo del proceso sin necesidad de estimar los gradientes del proceso, mediante la reformulación del método de adaptación de modificadores como un problema de optimización anidada: el método de *Adaptación de Modificadores Anidado*. La estructura de optimización del método anidado presenta una capa superior que no está basada en gradientes y que tiene como variables de decisión a los modificadores, lo que evita la estimación de la curvatura del proceso. Por otro lado, la capa interior conserva la misma estructura del problema de adaptación de modificadores.

Finalmente, se han presentado algunas ideas preliminares respecto a la aplicación del método de adaptación de modificadores basado en gradientes y el anidado, para un problema de optimización dinámica con horizonte móvil.

Los resultados obtenidos a partir de la implementación de los métodos propuestos muestran que, para el caso de la optimización estática, se ha ampliado el campo de aplicación de la metodología original, siendo posible implementarla en problemas donde los gradientes experimentales son difíciles de obtener (o no están disponibles), así como para aquellos procesos cuyas medidas están contaminadas con ruido, detectando el óptimo del proceso mediante un modelo erróneo, de una manera más robusta y siguiendo una evolución factible. Referente a la aplicación

dinámica, se ha mostrado que mediante la corrección de los gradientes del proceso con la dinámica natural del sistema, es posible encontrar el óptimo de un proceso en un problema de optimización dinámica con errores de modelado, siempre y cuando la función objetivo pueda ser medida al final del horizonte de predicción.

1 INTRODUCTION

1.1 Introduction

Optimization is one of the most versatile tools employed in process engineering. Several applications, involving the whole production structure of an industrial facility, are based in the solution of an optimization problem. Since the industrial development implies searching for comparative advantages with respect to the market, the idea of producing with the minimum amount of the available resources, such as: energy, raw materials and utilities, seems to be an excellent way to become more competitive. To do this, it is necessary to make the right decisions in the different time scales that a plant presents:

Thinking in terms of months and weeks, it is necessary to *plan* the production of the factory according to the market prices. Also it is necessary to *schedule* the production, in order to ensure that the manufacture volume can be achieved with the available process units in a reasonable time.

If the time-scale is reduced to days and hours, it is necessary to know the best way to drive the process to the optimal production plan from the point of view of the process variables, in order to minimize the use of the available resources.

At last, when the time scale is fixed to minutes, it is necessary to control and supervise the process, following the given references and rejecting possible disturbances. Moreover, a correct treatment of the data process, allows detecting as early as possible the occurrence of faults in the system.

In general, the occurrence of different time scales in a process makes possible the fact that the outcomes of the upper layers are part of the degrees of freedom of the lower ones. This *Optimization Architecture* allows producing in a safe and optimal way.

The general optimization problem can be represented in equation (1.1), where $u \in \mathbb{R}^{n_u}$ are the decision variables, $x \in \mathbb{R}^{n_x}$ are the dependent variables, $f: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \to \mathbb{R}$ is the performance function, $h: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_x}$ is the model of the process that relates the dependent and the independent variables, $g: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_g}$ are the inequality constraints that describe the feasible region of an optimization problem and $\theta \in \mathbb{R}^{n_\theta}$ can be considered as the parameters of the optimization problem.

$$\begin{aligned} &\min_{u} f(x,u,\theta) \\ &s.t.: \\ &h(x,u,\theta) = 0 \\ &g(x,u,\theta) \leq 0 \end{aligned} \tag{1.1}$$

From the definition of the general optimization problem already presented, several classifications can be proposed in order to describe the application of equation (1.1) to the different problems of the process industry (Figure 1.1).

In terms of the dependence of the model with time, the optimization problem can be classified as Steady state or Dynamic. If the model employed in the optimization problem does not consider the dynamic behavior of the system, the equation (1.1) can be classified as a static optimization problem. These kinds of problems can be applied in the upper layers of the Optimization Architecture, with the aim to obtain the optimal operation point or planning the facility. On the other hand, if the model of the process takes into account the evolution in time of the system, equation (1.1) can be described as a dynamic problem, which has its main applications in the lower layers (or regulatory layers) of the Optimization Architecture, calculating the evolution of the decision variables that change the system from the actual condition to the optimal one.

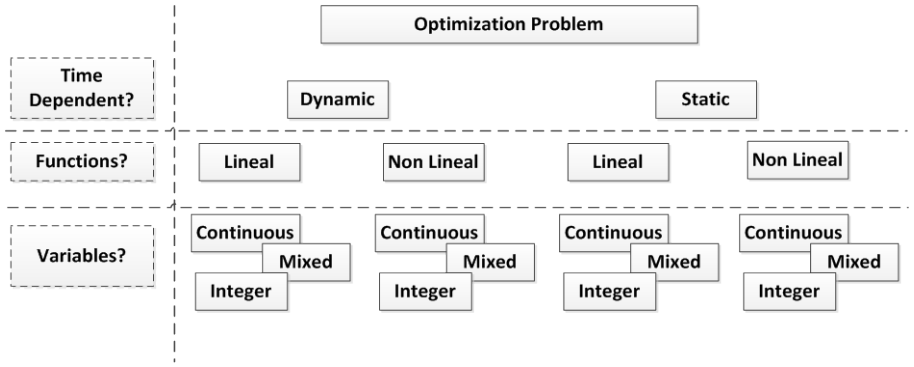

FIGURE 1.1: CLASSIFICATION OF THE OPTIMIZATION PROBLEMS

Another classification of the equation (1.1) can be proposed according to the nature of the variables involved, leading to continuous, integer and mixed integer optimization. By using different types of variables, it is possible to apply the optimization problem in different processes and contexts, e.g. continuous and batch plants, event driven processes, among others.

Furthermore, according to the nature of the functions of equation (1.1), the optimization problem can also be classified as linear and non-linear optimization. The use of linear functions allows computing faster solutions but the problem is in general only an approximation of the real system. Because of this, linear optimizations are used to solve control problems where the linear system is identified locally around the operation point. In contrast to the linear case, nonlinear problems are generally more difficult to solve, and even in some situations there is not guarantee of the existence of solutions. Since complex nonlinear models are employed, the use of these problems is oriented to optimizations with the objective to explore the entire process, finding the economic optimal operational point.

Because of the large and diverse applications of the optimization in the process industry, the study in this field has been very intense. The work of

Biegler and Grossmann (Biegler & Grossmann, 2004) present a retrospective about the most important applications of the optimization in process industry, classifying the applications according to the type of optimization problem to be solved. In that article, the authors also present a summary of the most important methods to solve equation (1.1) and the available solvers. As a second part of the retrospective, Biegler and Grossmann present a future perspective on optimization (Ignacio E. Grossmann & Biegler, 2004), outlining the directions of research that are likely to be subject of further significant research work over the next decade and so on. The authors propose the following four working areas:

Global Optimization: one of the most important limitations of the actual methods to solve nonlinear problems (NLP and MINLP), is the lack of guarantee to find a global optimum when the optimization problem present nonconvexities. Even though several developments have been proposed (especially in Chemical Engineering), there has been an increase in the research activity in this area from the year 1991, when the *Journal of Global Optimization* was published by the first time. This intensification in the research, gives important clues about the importance of this topic in the future.

Logic-Based Methods: the major motivation in this area lies in developing symbolic representation of discrete constraints, with the aim to facilitate their modeling and motivate the more effective solution techniques that can help to reduce computational time of discrete-continuous optimization problems.

Large-Scale Optimization: The success of optimization strategies in process engineering constantly motivates the desire to formulate and solve larger problems over a wider set of domains. Handling larger problems allows including more detailed models in terms of

temporal and spatial representation, and permits the integration of other models in order to improve the physical representation of the reality, leading in more realistic solutions to be applied.

Scientific Computing: Future progresses in the previous three topics necessarily need more power computers in terms of hardware and object oriented modeling software. Therefore, it is critical to emphasize the efforts in developing machines and programming languages.

Even though the authors propose the four topics presented above, there is a transversal subject that is mentioned in the first three of them: *the uncertainty*.

Dealing with uncertainty in process optimization is a key issue if we are interested in proposing realistic and safe solutions. Since the optimization is almost all of the times based in models that are only an abstraction of the reality, taking into account the difference between this representation and the real world is mandatory to suggest valid solutions. Moreover, processes present some variables with random behavior that might affect the feasibility and the optimality of the proposed solution of the optimization; therefore, this behavior might be explicitly treated.

1.2 Motivation

In industrial processes changes in the load and type of the influents is a common situation. If we take as an example the petroleum industry, every three or four days the crude oil changes its origin, meaning that a different raw material must be treated. The management of the feed transition operation increases its difficulty, if we consider that some of the process streams employed to perform the chemical reactions are recycled by-products from side processes and they might have variable and unknown purities too. Therefore, if a feasible and economic operation policy is

proposed to deal with the operational changes it must consider these uncertain variables.

Even when there are important sources of uncertainty in the reactants, some information can be extracted from past operation in order to know at least approximately what are the bounds and the expected distribution of the unknown stream properties. If this information can be exploited in the optimization step, we are able to give better and safer alternative to carry on this operation. Therefore, an adequate use of the stochastic information of the uncertain variables obtained from the DCS system can be very useful to propose an optimal and feasible way to operate under random situations.

On the other hand, obtaining operational points of complex processes is not a trivial task and a systematic methodology based on solving an optimization problem must be applied. However, because of the inherent difference between the model and the real process, continuous corrections estimated from the process data must be applied into the optimization problem. In the process industry, Real-time optimization (RTO) has become a popular tool to look for the optimal conditions despite this kind of uncertainties (Cutler & Perry, 1983). The authors not only conclude that important economic improvements can be expected when this technique is applied, but also remark that using RTO is an effective means of distributing technical know-how from research to the plant.

The optimization that is solved in RTO layer is based in a model of the process, which in general is a gross representation of the reality that might neglect important physical dependences. This may lead to solutions that don not correspond to the real process optimum. For that reason, the study and the development of RTO methods able to overcome this mismatch is an important area to research in order to deal with the uncertainties that affect the process optimization in a better way, overcoming the partial knowledge of the process.

In summary, it can be said that the motivation of this thesis is to handle the uncertainty in process optimization from two points of views:

- Incorporating the random and unknown behavior of some process variables in the optimization by using probabilistic information about their past values.

- Incorporating process measurements into the optimization in order to find the true operational optimum point of the process using RTO, overcoming the uncertainty that can be attained to the errors produced in the modeling stage of the process.

1.3 State of the Art: Handling Uncertainties in Process Optimization

As it was mentioned, taking into account the uncertainties that affect the processes is completely necessary in order to give realistic solutions to different situations where the optimization is required. There are many ways in which these uncertainties can be considered. The most common approach is model updating, but other methods reported in literature manage the uncertainty in process optimization from the view point of:

- Robust optimization
- Stochastic methodologies
- Converting the optimization in a control problem

In this thesis, two points of view of the uncertainty will be studied: the treatment of random behavior of process variables using *Stochastic Programming* and overcoming the modeling mismatch in optimization by means of the *Real Time Optimization*.

The context of both topics will be briefly summarized next.

1.3.1 Stochastic Programming

In the classical approach of optimization, the equations and parameters are considered totally known. However, when the solution is applied to the reality, frequently the value of the objective function is worse than expected one and/or the constraints are violated (Birge & Louveaux, 1997; Rockafellar, 2001; Wendt et al., 2002). These problems can be attributed to the uncertainty that affects the system (Wendt et al., 2002). Usually in process optimization, an important source of uncertainty can be assigned to the random behaviour of some parameters that can be described using random variables ξ that present expected bounds and belong to a probability space with a given probability distribution function (PDF). Taking into account explicitly the stochastic behaviour that the chance variables can present in the optimization problem from probabilistic point of view, is what is called *Stochastic Programming.* There are other alternatives to handle the uncertain behaviour of random variables like fuzzy programming, which are clearly summarized in the review of Sahinidis (Sahinidis, 2004).

Handling the unknown variables can be done with robust optimization as well. The essence of this methodology is finding an optimal and feasible solution for the random optimization problem considering all the possible cases. That is to say, all the values within the bounds of the unknown variables. Stochastic optimization problems on the other hand present a similar approach of the robust optimization, but taking advantage of the fact that the probability distribution of the random parameters can be obtained or estimated from available data. Therefore, it is possible finding an optimal policy that is feasible for almost all of the possible realizations of the chance variables (the degree of feasibility required can be adjusted due to the fact that the PDF of ξ is known), optimizing the expectation of an objective function of the decision and the random variables. Therefore, applying stochastic techniques to the process optimization seems to be an efficient strategy to find a solution that compromises feasibility and

optimality in the operation of a determined process. Because the random variables can have a continuous nature, it can be necessary transforming an infinite dimensional problem into a finite one by means of solving a parameterized version based in scenarios

In the field of stochastic optimization, there are two main ways to understand the optimization problem: Multistage Programming and Chance Constrained Programming (COSP, 2012; "Preface," 2003; Sahinidis, 2004).

1.3.1.1 Multistage Programming

Multistage programming or Programming with recourse is based in the idea that there are several stages of knowledge of the random variables. At the beginning a set of decisions must be taken without full information of some random events, knowing only their probability distribution (which is a common assumption). As time passes, it is assumed that it is possible to measure or estimate the real value of the unknown variables and new decisions must be taken in order to correct the initial choices. It was first presented by Dantzig (Dantzig, 1955) who defines the concept of *decision stages*, and then commented in the work of Beale (Beale, 1955) in the context of solving large scale linear problems. In the paper of Dantzig it is presented a general model to handle explicitly the random variables for linear programming by means of solving *multistage* problems. In each stage, decisions must be taken considering the previous ones and the probability distribution function of the random variable in the future, which is called a *multistage recursion problem*.

Starting with the basic idea of decision stages and the relationship among them, several approaches have been proposed in order to extend the original proposition of Dantzig into a more general and applicable context, including nonlinear problems and numerical algorithms that takes advantage of the particular structure of multistage problems, among others.

In terms of the probability distribution function of the random variables, there have been efforts to propose a numerically solvable optimization problem when the PDF of the random variable is continuous since the problem becomes infinite – dimensional. One alternative to this situation is to postulate a function of the random variable as a solution of the optimization problem, using this approach the problem becomes finite dimensional in the space of the parameters of the postulated function, which is problem – dependent for the nonlinear case. On the other hand, it is possible to use a *Scenario Representation* of the random space as a finite-dimensional approximation of the original problem, transforming the original stochastic optimization problem in a large scale deterministic equivalent that can be solved by means of many general purpose optimization algorithms (Dupacova, 1995). Because of this versatility and generality of the scenario representation, this methodology is commonly used for practical problems. However, the exponential grow of the deterministic equivalent problem that is produced with the discrete assumption of the PDF, can make the problem prohibitively large. To overcome this issue there have been developed scenario reduction techniques based in: decreasing the number of scenarios adapting them using pass information of the random variables, to aggregate some periods and/or some scenarios (Römisch, 2009), selecting only the "important" ones with some statistical data (Dupačová et al., 2000) or using expert's opinions or heuristic strategies (Dupacova, 1995; Karuppiah et al., 2010).

On the other hand there is a lot of interest in modelling the recursive part of the multistage optimization problem, due to the fact that depending on the type of constraints utilized to describe the nonanticipativity behaviour of the multistage variables and the scenario description of the PDF, the resulting optimization problem will have a particular structure that can be exploited by different solvers (Birge & Louveaux, 1997), including the option to split the large – scale resulting problem in a primal problem and dual set

of problems which code can be parallelized in order to reduce the solving time (Kall & Wallace, 1997; Ruszczynski, 1998).

One of the first attempts to solve industrial optimization problems considering the uncertainty in processes based in multistage optimization, appears in the works of Grossmann and co-workers introducing the concept of flexibility indexes (I. E. Grossmann et al., 1983; Halemane & Grossmann, 1983). After these applications some works based on optimization under uncertainty using stages of decisions were modified to be used in the process industry (Pistikopoulos & Ierapetritou, 1995; Rooney & Biegler, 1999, 2001, 2003).

In process engineering, multistage programming (and its particular version, the two-stage programming), has been used in several applications for problems where the time scale between the first and the next stages is clearly delimited, i.e. long term decisions. For example, in optimal facility location problems where the first stage decision is the location and/or the construction of a determined factory, while the second one is the expected production for the already built factories. In almost all of the times this optimization is stated as a Mixed Integer (Non)Linear Multistage Problem and the applications include: off-shore gas fields (Goel & Grossmann, 2004), petroleum refineries (Khor et al., 2008), petrochemical networks (Al-Qahtani et al., 2008), electric energy producers (Beraldi et al., 2008; Fleten & Kristoffersen, 2007) and even the allocation of water supplies from the highland lakes in Texas, USA (Watkins et al., 2000) among others. Long term planning has been also employed by Tarham and Grossmann to include the possibility of investing in pilot plants with the aim to reduce the uncertainty and take into account the possibility of expansion capacities in a particular plant for a given process network (Tarhan & Grossmann, 2008). In the context of medium term decisions, recursive programming recently has been used for the utilization and planning of batch multiproduct plants (Cui & Engell, 2010). Long and medium term decisions are stated as a multistage problem. However, the important outcome is obtained in the first stage

since it gives information about the initial investment of a production facility or general project. The results of the second stage, conversely, are only used as a guide to select correct sizes, but in practice they are not applied directly.

On the other hand, the short term decisions involving uncertainty, such as control and real-time optimization, has been treated mostly from the point of view of the robust (N)MPC approach that provides robustness against disturbances. A significant effort has been invested in this field both theoretically (Limon et al., 2009; Magni & Scattolini, 2007; Rawlings & Mayne, 2009) and practically, with the main idea to reject disturbances in a robust way. The usual robust (N)MPC is based on min-max approaches (Campo & Morari, 1987) that try to minimize the worst-case realization of the uncertainty. They can be classified in open-loop approaches, if they assume open-loop control in the optimization, or closed-loop approaches, if they assume closed-loop control in the optimization, taking feedback explicitly into account (Lee & Yu, 1997). As another point of view, it seems to be a good alternative to explore the use of multistage programming for this kind of applications in order to take advantage of the knowledge of the PDF of the random variables. However, there are no reports of the use of this strategy to cope with the uncertainty in short periods in the literature. At present, the group of System Dynamics and Process from Dortmund is working with the idea of using it to provide less conservative close-loop control than the robust (N)MPC ensuring feasibility in the outcomes of the control layer (Lucia et al., 2012), as it is commented in the HYCON2 report (Navia et al., 2011).

1.3.1.2 Chance Constrained Optimization

Chance constrained optimization also provides a probabilistic treatment of the uncertainty that affects an optimization problem, but unlike the multistage programming it assumes that a decision must be taken knowing only the probability distribution of the random variables, neglecting the idea of knowing the true realization of the uncertainty in the future.

Therefore, the uncertainty is taken into account from the point of view of ensuring a certain degree of feasibility in the solution of the optimization, no matter the value that the random variable might have within its PDF, focusing the approach in the reliability of the system. It was presented in the first time by Charnes and Cooper (Charnes & Cooper, 1959) in the context of independent probability constraints, and then it was extended to the case of problems where different constraints must be fulfilled jointly by Miller and Wagner (Miller & Wagner, 1965) and Prékopa (Prékopa, 1970)

The main challenge in this area is how to calculate the probability of the chance constraints and most of the efforts are focused in this direction. In the case of some linear problems with independent chance constraints, the propagation of the uncertainty can be tracked and then a deterministic linear equivalent problem of the uncertain one can be obtained applying the inverse of the PDF (if exists).

For a general set of linear chance constrained problems, Prékopa has presented the requirements of the PDF in order to transform the original uncertain constraints into standard ones (Prékopa, 1995), also the author has shown an extensive review of methods to solve, either analytically or numerically, problems where the feasible set is convex (Prékopa, 2003). In the context of nonlinear chance constrained optimization, some methods have been proposed using different solution perspectives, for instance: transforming the original problem into a dynamic programming one (Mine et al., 1976), using sampling techniques (Diwekar & Kalagnanam, 1997; Nemirovski & Shapiro, 2006b; Sakalauskas, 2002) and defining convex deterministic approximations (Nemirovski & Shapiro, 2006a), among others. The convexity assumption seems to be a very important drawback to apply previous methodologies in general nonlinear problems used in process optimization, since the main requirements involve dealing with a set of measurements that must be convex and a probability distribution function of the constrained variables that must be log-concave. Finding a convex set of measurements is quite complex in process system

applications, even for the most simple models based in first principles since the mass balance defined for each component produces non convexities in the feasible set of the model, which implies that if the compositions are probabilistic constrained this set cannot be considered for these kind of methods. On the other hand, the requirements of log-concavity of the PDF implies knowing a priori the function of the probability distribution function of the constrained variables, which in general is unknown, and even in the case when the random variables can be described using a log-concave PDF, the nonlinearities of the system might not preserve this characteristic, leading to PDFs in the constrained variables that do not fulfil the basic assumptions of these methodologies again.

To overcome these limitations in the context of process system engineering, Wendt and co-workers have presented a methodology based in the inverse mapping of the PDF of the random variable, to estimate the probability constraints of the process outputs for the case of nonlinear optimizations (Wendt et al., 2002). In general, this problem consists in finding critical values of the random variables that are in the border of the confidence region such that the probability of fulfilment of the constraints is equal to the degree of feasibility required. The methodology proposed by the authors presents the advantage that the only requirement to be applied to solve a chance constrained problem is the existence of monotony between the random variable and the constrained output and the fact that this is not a sampling-based technique.

The first real life application of the probabilistic constraints was formulated for the electrical energy sector in Hungary to reach a determined degree of reliability in the produced power using a linear model that was transformed in a deterministic one (Prékopa et al., 1980), after this application Prékopa summarized the numerical solution method applied in this kind of electricity applications(Prékopa et al., 1998). Following with linear applications, the probabilistic constrained optimization has been formulated and solved to design water reservoirs that can protect a

downstream area from flood (Dupačová et al., 1991). From the point of view of the industrial applications, there are some examples of solving the inventory problem in order to fulfil the uncertain demands from the market(Beraldi & Ruszczynski, 2001). An extensive review of applications in engineering and other fields like finance and food management is presented in the compilation of Prékopa (Prékopa, 2003). In all cases, the author gives details about the examples, indicating that they correspond to linear problems that can be solved using convexity-based techniques.

About the applications of probabilistic constrained optimization in process system engineering, there are some examples for linear cases in the context of MPC (Henrion & Möller, 2003; Pu Li et al., 2002; P. Li et al., 2002) and for nonlinear ones in examples related with the production planning to optimize complex process systems such as continuous and semi-batch distillation columns (Arellano-Garcia & Wozny, 2009; Barz et al., 2008; Li et al., 2008), being the method of the inverse mapping technique (Wendt et al., 2002) the chosen one to estimate the probabilistic constraints.

As it can be noted from the applications, we can say that this methodology is not commonly used to cope with the uncertainties in economic process optimization, since in these cases it is necessary to use nonlinear models.

1.3.2 Real Time Optimization

In the process industry, plants have a large number of units which are interconnected. Finding the optimal operation point of these facilities is not a trivial task, because of the inherent difficulties of the process itself, as well as the uncertainties and disturbances that continuously modify the operating conditions. Therefore, it is necessary to use a systematic mechanism to find out the optimal operating point. Real-time optimization (RTO) has become a popular tool in the chemical and petrochemical industry in order to look for the optimal conditions despite the uncertainties (Cutler & Perry, 1983), which can be due to the partial knowledge of the process. This implies solving an optimization problem

using inaccurate or approximated models in the decision making process where the operational point of the plant is calculated.

In highly automated plants, optimal operation is typically addressed by a decision hierarchy involving several levels that include planning, real-time optimization and process control. At the RTO level, medium-term decisions are made on a time scale of hours to days, considering economic objectives in an explicit way. This step typically consists in an optimizer that determines the optimal conditions under slowly changing conditions (Marlin & Hrymak, 1997), hence steady-state assumptions are considered in the optimization carried out.

Real-time optimization emerged in the late 1970's with a two-stage algorithm: parameter estimation and economic optimization with the idea of making the model as much similar possible to the real process in the present conditions. In the first step, the uncertainties are taken into account updating the parameters of a nonlinear model of the process, while in the second step a new economic stationary point is found with the updated model (Bamberger & Isermann, 1978). The new operating point is then applied to the process in an iterative scheme until no further improvements in the cost function are observed.

The application of the RTO is concerned with implementing economic decisions in real time based on an updated non-linear steady-state model, with a level of detail that upper (planning) layers do not have. In its implementation values are directly passed or translated by using a filter (that can involve plant operators) to the MPC layer, following the hierarchical Optimization Architecture (Engell, 2007).

At the MPC level, there may be one or more MPCs, depending on the controller envelope(s)(Findeisen et al., 1980). The MPC(s) provides the minute-to minute dynamic control of the plant and provide some amount of optimization capability because of the RTO updates.

The implementation of the RTO/MPC configuration may produce important benefits in the plant. However, it increases significantly the complexity of the control system because of the continuous updates, making the model implemented in the MPC layer (identified for a particular condition of operation), inaccurate and erroneous. To justify the application of the RTO, it is necessary to realize that, without this layer, plant operators are not able to optimize set-points within the feasible range according to the disturbances affecting the process (Engell, 2007). Therefore, we can argue that the RTO is recommended when there are degrees of freedom that can be adjusted and when variability of the optimal operating point caused by changes in operating conditions is frequent (Darby et al., 2011). Besides the justification of the implementation in terms of the benefits, RTO presents and important benefit related to be an effective way of distributing technical know-how from research to the plant (Cutler & Perry, 1983).

There are several challenges related to the RTO applications (Darby et al., 2011). About the interaction with the MPC layer, there are important progresses in the study of the tradeoff between benefits related with the time scale separation, the problems derived from the differences between the models used in each layer and the loss of optimality produced with the steady state assumptions. The general solution proposed is to implement the dynamics into the RTO layer, transforming it into a Dynamic RTO or an economic driven NMPC level (Backx et al., 2000; De Prada & Valentin, 1996; Engell, 2007, 2009; Gonzalez et al., 2001). There are different alternatives to make the resulting nonlinear dynamic optimization solvable in real time. Among them, there are proposals that keeps the two-layer structure, such as: the Infrequent DRTO executions by using sensitivity-updated solutions keeping the two layers structure (Kadam & Marquardt, 2004; Würth et al., 2009, 2011), the reduction of search region with stochastic criteria(Ochoa, Repke, et al., 2010; Ochoa, Wozny, et al., 2010) and the frequent application of a static model to explore the objective function in Real Time Evolutions (Sequeira et al., 2002; Sequeira et al., 2004). Other solutions

include merging the DRTO with the MPC layer, solving in each sampling time an NMPC with economic objectives. Since the computational times must be significantly reduced there are several approaches to distribute the entire optimizations in several local NMPCs using different methods (Findeisen et al., 1980; Scattolini, 2009). In particular, it has been developed an promising methodology based in a hierarchical architecture using a coordination layer that makes converge the distributed NMPCS to the optimum of the original one using: game theories, price driven coordination or Lagrangian Relaxation methods (Maestre et al., 2011; Martí et al., 2012; Voos, 2007).

Moreover, there are important advances in improving the capability of the RTO layer to detect the optimum of the process. Because of the modeling mismatch, produced either from the partial knowledge of the process e.g. when physical dependences are not correctly modeled, or from the simplifications that must to be made in order to make the nonlinear model solvable, the classic two-step algorithm does not converge into the real optimum of the process because of the interaction between the two steps, as Roberts demonstrated in latest years of the 70's. Since the modeling mismatch is a common case, there have been several attempts to overcome this situation adapting the model in order to take into account this difference and converge to the optimum of the process, which can be classified in three main groups according to Chachuat and co-workers (Chachuat et al., 2009): *Model Parameter Adaptation*, *Modifier Adaptation* and *Direct Input Adaptation*, which will be briefly mentioned next.

1.3.2.1 Model Parameter Adaptation

Maintaining the classical two-step structure, these methods takes into account the interaction already mentioned identifying the characteristics that the updated parameters must have to achieve the plant optimum.

Forbes and coworkers present a method to evaluate whether the model available has the ability to find the real process optimum defining the

"Point-model Adequacy" (Forbes et al., 1994) and the "*Model Accuracy*" (Forbes & Marlin, 1994). These strategies consist in finding a suitable choice of model parameters such that the optimal point of the plant coincides with the optimal point of the model for a given value of these parameters. Even though this method gives sufficient conditions to ensure process optimality with the two-step approach, it uses the concept of reduced gradient which implies that the set of active constraints must be the same for the process and for the model, which also must be known a priori. This, plus the fact that the value of the real optimum of the process must be estimated in order to perform a gradient matching, makes this classification scheme not very practical.

Another way to take into account the interaction between the parameter estimation and the economical optimization is investigated by Srinivasan and Bonvin in a dynamic optimization problem (Srinivasan & Bonvin, 2002). The authors work in a run-to-run optimization for batch processes, attempting to change the paradigm of "identification of the model" for "modeling for optimization". To do this, they identify a model that allows computing inputs that are nearly optimal for the reality. For this, the cost and constraints of a dynamic optimization problem are included in the identification objective creating an integrated problem, while the interactions between both optimization problems are taken into account by using appropriate Lagrange multipliers.

1.3.2.2 Modifier Adaptation

As it was said in the introduction of the RTO, it is necessary to manage the interaction between the parameter estimation and the economic optimization steps in order to find the process optimum. With this in mind, Roberts developed the ISOPE[1] algorithm (Roberts, 1979) as the result of merging the two classical steps. In a few words, with the ISOPE algorithm

[1] ISOPE is the acronym of "*Integrated System Optimization and Parameter Estimation*"

the convergence to the optimum of the process can be achieved correcting the objective function with a linear term that takes into account the difference between the gradient of the model and the process, for a square identification problem with only constraints in the decision variables.

Later, Tatjewsky demonstrated that the same convergence results of the ISOPE method can be obtained if the parameter estimation step is replaced by zero order corrections for the process measurements, presenting the *Iterative Optimizing Set-Point Control* (Tatjewski, 2002). This technique was extended to general constrained optimizations by Gao and Engell who presented a first order correction in the constraints that takes into account the difference between the process and the model of their gradients and values(Gao & Engell, 2005). The constraint adaptation was proved by Chachuat and co-workers giving the theoretical basis for this methodology (Chachuat et al., 2008).

The adaptation methodology was finally formalized by Marchietti and co-workers who presented the *Modifier Adaptation Methodology* (Marchetti et al., 2009). The authors gave the explanation for the convergence to the real optimum when the adaptation in the economic optimization is performed correcting the gradients of the objective function and the constraints, from the point of view of the KKT matching between the model and the real process.

1.3.2.3 Direct Input Adaptation

Direct-input adaptation methods were developed with the idea of avoiding repeated numerical optimizations performed in the RTO layer by changing the optimization problem into a feedback control problem. The aim of this feedback control is to calculate the set points for the manipulated variables while trying to maintain certain measures of optimality fixed. The challenge then is to find the optimality functions calculated with the measured variables that must be fixed by changing the decision variables such that the system is driven to the real optimum. In other words, the goal is to

choose appropriate controlled variables in order to achieve a similar performance as would be realized by a (fictitious) on-line optimizing controller (Chachuat et al., 2009).

In the literature, three different methods can be found to implement the idea of Direct Input Adaptation: *Self-Optimizing Control*, *Extremum Seeking Control,* and *NCO Tracking*, which will be explained next. For more detailed information, see chapter 2 of the Deliverable 3.1.1 of the HYCON2 Project (Navia et al., 2011).

> *SELF-OPTIMIZING CONTROL:* In this case the on-line optimization is replaced with the design of a control structure for the process, such that when maintaining some variables, or combination of variables, at fixed set points a close-to-optimal situation is attained. Ideas related to self-optimizing control have been presented repeatedly in the process control history. The first quantitative treatment was presented by Morari and coworkers (Morari et al., 1980). Nevertheless, Skogestad defined the problem in detail and was the first to introduce the formally of the methodology (Skogestad, 2000).
>
> *EXTREMUM SEEKING CONTROL:* Following the concepts of Direct – Input Modifier Approach, the task of Extremum seeking control is to find the operating set points that maximize or minimize an objective function (Guay & Zhang, 2003). Two main approaches of Extremum seeking control have been presented (Dochain et al.): perturbation-based (Krstic & Wang, 2000) and model-based (Guay & Zhang, 2003). In the first method, a constant perturbation signal is introduced to the system in order to calculate an estimate of the gradient of the measured objective function without any previous knowledge about the system (black box). On the other hand, in the second method, an approximation of the objective function is assumed to be known. This function presents parametric

uncertainties, and the model is used to calculate analytically its gradient

NCO TRACKING: One common practice when conducting a transient operation in the industry is following a well-defined recipe (from the laboratory, or from the experience of the operators). However, due to the existence of constraints related with equipment limits and quality levels, which conflict with the recipes, operators will naturally introduce a degree of conservatism to guarantee feasibility moving away from these constraints, thus leading in suboptimal results. To reduce this backing off from constraints which is introduced intentionally to take care of disturbances, an optimization-based methodology can be applied to follow the constraints in an optimal way (Bonvin & Srinivasan, 2003). The optimal trajectories can be estimated from nominal dynamic models which can then be applied as reference trajectories to lower layers. Nonetheless, the uncertainty related with the model available and the disturbances that can affect the system makes the application of standard dynamic optimization methods ineffective. To overcome these problems, the authors suggest a method consisting in evaluating offline the optimality conditions of the process using a dynamic model. Once the solution of the problem is fully characterized, the optimality criteria can be applied in a feedback control in order to maintain optimality despite the uncertainties of the process.

A more extensive revision of the three methodologies can be found in Chapter 2 of Deliverable 3.1.1 of HYCON2 project (Navia et al., 2011).

Each of these adaptation strategies presents their own particular advantages and inconveniences. If we fixed our attention in direct adaptation methodology to classify if the model employed in the optimization is adequate and accurate, it is completely mandatory knowing

a priori the set of active constraint in the real optimum of the process, which can be almost impossible if some of the physical dependences are neglected in the modeling step. Therefore, this focus can be impractical when modeling mismatch is present. On the other hand, the modifier adaptation methodology ensures convergence to the optimum of the process without previous information about its location in the feasible space. However, the KKT matching with the real process needs a correct estimation of the difference between the process and the model, and since this method adds a first order correction in the economic optimization, it is necessary to measure in some way the experimental gradient of the process. Even when there are some alternatives to estimate this magnitude, they need a lot of previous information or they require applying some kind of perturbations to the real process, which can be inapplicable for some plants. Finally, the direct input adaptation presents the advantage that they become model-independent since online optimization is no longer required. However, Self-optimizing control only guarantees a solution close to the optimum of the process where the optimality has been obtained with a model that can present mismatch, Extremum-seeking control present the same problem than the modifier adaptation methodology about the continuous perturbation of the plant to estimate the gradient, in addition with the incapability to manage the process constraints, and finally NCO tracking depends on the invariance of the set of active constraints with the uncertainty.

Even when every adaptation strategy presents their pluses and cons, from the perspective of the applicability, only the modifier Adaptation Methodology is able to: manage constraints, not require previous information of the true optimum of the process (like the set of active constraint) and ensure optimality for a given model. Of course, the need to estimate the gradient of the process, as well as, other problems related with its implementations must be considered in order to make this strategy more realistic to be applied in a real plant.

1.4 Objectives

The general objective of this thesis is to study the applicability of optimization methods in the process industry taking into account explicitly the uncertainty in models and variables. First the problem was considered from the stochastic approach, but the computation times obtained were not very promising. This motivated a change in focus toward RTO problems with the modifier-adaptation methodology. So, this thesis is split in two parts, each of them with general objectives detailed at next.

1.4.1 Stochastic Optimization

The main objective of this part of the thesis is to study the applicability of methods to manage the stochastic behavior of the random variables in an optimization problem, from the viewpoint of the stochastic programming. As a test plant a hydrodesulfuration process is chosen.

To reach this general objective, the following particular objectives are proposed:

- Identify the main sources of uncertainties in a transition problem corresponding to a hydrodesulfuration unit.

- Study the particularities of the two-stage algorithm in terms of possible decompositions to be applied in dynamic optimization.

- Study the simplification in the implementation of the two-stage optimization when the probability distribution function has a continuous nature.

- Propose a method to generalize the discrete simplified solutions from the two-stage optimization, to be applied with the original uncertainty

- Study the implementation of the chance constrained optimization in the context of dynamic optimization with the sequential approach.

- Apply the general algorithms of stochastic programming in the particular problem of the transition in a hydrodesulfuration unit.

- Solve the problem of optimization under stochastic uncertainty and test the solution in terms of feasibility, optimality and applicability.

1.4.2 Real Time Optimization

The main objective of this part of the thesis is to propose improvements to be implemented in the Modifier-Adaptation Methodology, to increase its field of application as a valid tool to overcome the uncertainties produced in the RTO layer as a consequence of the modeling mismatch.

To reach this general objective, the following particular objectives are proposed:

- Study the generalities of the modifier Adaptation Methodology and compare its outcomes with the classic RTO method.

- Identify the main challenges that the Modifier Adaptation Methodology presents.

- Propose alternatives to solve the main problems detected, in terms of:
    - Infeasibilities that can be generated in the evolution of the process.
    - Gradient estimation step and an alternative to avoid it.
    - Dynamic versus static behavior in order to remove this requirement and save time.

- Test the alternatives proposed using process examples.

- Discuss and conclude about the solutions proposed and their applicability in a real facility.

1.5 Summary

The summary of this thesis is the following:

Chapter 2 deals with the first part of this thesis: the management of stochastic variables. It presents the following subsections:

- Introduction to the stochastic programming methodology. In this section the basic concepts about stochastic programming are presented, including the explanation of the algorithms that will be used.

- Description of the test problem. This section introduces the problems presented in a hydrodesulfuration plant when changes in the feed take place.

- Formulation of the two-stage and chance constrained programming for the problem of the hydrodesulfuration plant tested.

- Results obtained with the methods studied and tested with simulated random variables. Also the discussion of the results is presented

- Conclusions about the results and their applicability.

Chapter 3 presents the second part of this thesis: the study of the Modifier Adaptation Methodology in the context of the Real Time Optimization and

the improvements proposed. The following subsections are detailed in this chapter.

- Introduction to the Real Time Optimization.

- The Modifier Adaptation Methodology. In this section the main generalities of the methodology studied are explained.

- Implementation of the Modifier Adaptation Methodology. In this section we identify the main challenges of the methodology with the help of an example.

- Handling Infeasibilities in Modifier Adaptation Methodology. In this section the introduction of infeasibility controllers is presented as an alternative to the classic method to avoid process infeasibilities in the evolution of the system. The methodology proposed is tested in a process example.

- Reformulation of the Modifier Adaptation Methodology as a Nested Optimization problem. Throughout this section it is presented a completely new way to understand the modifier adaptation methodology, with the idea to avoid the gradient estimation step required in the original formulation. The methodology is tested in process examples.

- A first Approach for Modifier Adaptation Methodology in Dynamic Real Time Optimization. In this section, an attempt to extend the Modifier Adaptation Methodology for Dynamic Real Time Optimization is presented, with the aim of using the information obtained in the transient of a process to reach the optimum of the process. The methodology proposed is tested in a process example.

- Conclusions: advances proposed in the field of Real Time optimization using Modifier Adaptation Methodology.

Chapter 4 summarizes the main conclusions of this thesis in terms of the previously presented results. Also it presents the open issues for future work and the contributions.

1.6 Objetivos

El objetivo general de esta tesis es estudiar la aplicabilidad de métodos de optimización en la industria de procesos que consideren explícitamente las incertidumbres que pueden encontrarse en modelos y variables. Al principio el problema fue abordado utilizando un enfoque probabilístico. Sin embargo, los tiempos de cálculo observados no fueron adecuados, por lo que se cambió el enfoque hacia la Optimización en Tiempo Real con el método de adaptación de modificadores. De esta forma, esta tesis se ha dividido en dos partes, cada una de ellas con objetivos generales diferentes.

1.6.1 Optimización Estocástica

El objetivo principal de esta parte de la tesis es el estudio de métodos que permitan manejar el comportamiento aleatorio de las variables de proceso desde el punto de vista de la programación estocástica, aplicados en una unidad de hidrodesulfuración.

Para cumplir esta meta, se proponen los siguientes objetivos particulares:

- Identificar las fuentes de incertidumbre principales observadas en un problema de cambio de carga producido en una unidad de hidrodesulfuración

- Estudiar las particularidades del método de optimización de dos etapas referidas a la descomposición que puede implementarse en un problema dinámico.

- Estudiar la aplicación de simplificaciones en la resolución de un problema de dos etapas, referidas a la discretización de una función de distribución de probabilidad continua.

- Proponer un método para generalizar los resultados obtenidos a partir de la discretización de la distribución de probabilidad para ser aplicados con la incertidumbre original.

- Estudiar la implementación del método de optimización probabilística en el contexto de un problema de optimización dinámica resuelto con un enfoque secuencial.

- Aplicar los métodos antes estudiados a un problema de cambio de carga en una unidad hidrodesulfuradora.

- Resolver el problema de optimización con incertidumbre estocástica, probando la solución obtenida en términos de optimalidad y factibilidad.

1.6.2 Optimización en Tiempo Real

El objetivo principal de esta parte de la tesis es proponer y aplicar mejoras al método de Adaptación de Modificadores, y de esta forma ampliar su campo de aplicabilidad como herramienta para manejar las incertidumbres en la optimización de procesos, producidas en la capa de RTO como consecuencia de errores de modelado.

Para lograr esta meta, se proponen los siguientes objetivos particulares:

- Estudiar las generalidades del método de adaptación de modificadores y comparar sus resultados con el método clásico de RTO basado en dos etapas.

- Identificar los mayores desafíos del método de adaptación de modificadores

- Proponer alternativas para resolver estos problemas, en términos de:
    - Violación en las restricciones de proceso, durante la evolución del sistema.
    - Evitar la estimación del gradiente del proceso para detector el óptimo del sistema.
    - Estudiar la posibilidad de implementar el método de adaptación de modificadores en el contexto de optimización dinámica.

- Probar las propuestas en ejemplos de simulación.

- Discutir y concluir sobre las soluciones propuestas y su campo de aplicabilidad.

1.7 Esquema

El esquema de esta tesis es el siguiente:

El Capítulo 2 presenta la primera parte de la tesis: el manejo de las variables inciertas, con las siguientes secciones:

- Introducción a la programación estocástica, donde se explican los conceptos básicos de la optimización estocástica incluyendo los algoritmos que serán implementados en el problema de hidrodesulfuración.

- Descripción del problema. En esta sección se comentan los problemas relacionados en el cambio de carga en una unidad hidrodesulfuradora.

- Formulación del problema de optimización de dos etapas y probabilística para el sistema compuesto por la unidad hidrodesulfuradora.

- Resultados y discusión de las trayectorias obtenidas con ambos métodos.

- Conclusiones sobre los resultados y su aplicabilidad

El capítulo 3 presenta la segunda parte de esta tesis: el estudio y las mejoras propuestas para el método de adaptación de modificadores en el contexto de la Optimización en Tiempo Real. Se presentan las siguientes secciones:

- Introducción a la Optimización en Tiempo Real

- Método de Adaptación de Modificadores. En esta sección se explican las generalidades del método.

- Implementación del método de adaptación de Modificadores. En esta sección se identifican los principales desafíos a abordar en el método de adaptación de modificadores con ayuda de un ejemplo de simulación.

- Manejo de Infactibilidades en el método de Adaptación de Modificadores. En esa sección se presenta una alternativa para evitar la ocurrencia de infactibilidades en el proceso durante la evolución del sistema que está siendo actualizado con el método de adaptación de modificadores, mediante la implementación de controladores de infactibilidades.

- Reformulación del Método de Adaptación de Modificadores como un problema de Optimización Anidado. En esta sección se presenta

una forma nueva de entender la Adaptación de Modificadores como un problema de optimización anidado, que permite detectar el óptimo de un proceso sin la necesidad de estimar los gradientes reales de la planta.

- Método de Adaptación de Modificadores en Optimización Dinámica. En esta sección se presenta una propuesta para extender el método a optimización dinámica en tiempo real, desde el punto de vista clásico basado en gradientes así como del anidado propuesto en esta tesis.

- Conclusiones de las mejoras propuestas.

El capítulo 4 resume las conclusiones generales de esta tesis en relación a los resultados presentados previamente. También se comentan los puntos abiertos para seguir la investigación y las contribuciones realizadas.

1.8 References of the Chapter

Al-Qahtani, K., Elkamel, A., & Ponnambalam, K. (2008). Robust Optimization for Petrochemical Network Design under Uncertainty. Industrial & Engineering Chemistry Research, 47, 3912-3919.

Arellano-Garcia, H., & Wozny, G. (2009). Chance constrained optimization of process systems under uncertainty: I. Strict monotonicity. Computers & Chemical Engineering, 33, 1568-1583.

Backx, T., Bosgra, O., & Marquardt, W. (2000). Integration of model predictive control and optimization of processes. Aachen: LPT, RWTH.

Bamberger, W., & Isermann, R. (1978). Adaptive on-line steady-state optimization of slow dynamic processes. Automatica, 14, 223-230.

Barz, T., Arellano-Garcia, H., & Wozny, G. (2008). Implementation of Optimal Decisions in the Presence of Uncertainty. In 17th World Congress The International Federation of Automatic Control (pp. 11931 - 11936). Seoul, Korea.

Beale, E. M. L. (1955). ON MINIMIZING A CONVEX FUNCTION SUBJECT TO LINEAR INEQUALITIES. Journal of the Royal Statistical Society Series B-Statistical Methodology, 17, 173-184.

Beraldi, P., Conforti, D., & Violi, A. (2008). A two-stage stochastic programming model for electric energy producers. Computers & Operations Research, 35, 3360-3370.

Beraldi, P., & Ruszczynski, A. R. (2001). The probabilistic set covering problem. Operations Research, 50, 956-967.

Biegler, L. T., & Grossmann, I. E. (2004). Retrospective on optimization. Computers & Chemical Engineering, 28, 1169-1192.

Birge, J. R., & Louveaux, F. (1997). Introduction to stochastic programming. New York: Springer.

Bonvin, D., & Srinivasan, B. (2003). Optimal operation of batch processes via the tracking of active constraints. ISA Transactions, 42, 123-134.

Campo, P. J., & Morari, M. (1987). Robust model predictive control. In American Control Conference (Vol. 2, pp. 1021-1026).

COSP. (2012). Introductions to Stochastic Programming. In (Vol. 2012): The Committee on Stochastic Programming (COSP).

Cui, J., & Engell, S. (2010). Medium-term planning of a multiproduct batch plant under evolving multi-period multi-uncertainty by means of a moving horizon strategy. Computers & Chemical Engineering, 34, 598-619.

Cutler, C. R., & Perry, R. T. (1983). Real time optimization with multivariable control is required to maximize profits. Computers & Chemical Engineering, 7, 663-667.

Chachuat, B., Marchetti, A., & Bonvin, D. (2008). Process optimization via constraints adaptation. Journal of Process Control, 18, 244-257.

Chachuat, B., Srinivasan, B., & Bonvin, D. (2009). Adaptation strategies for real-time optimization. Computers & Chemical Engineering, 33, 1557-1567.

Charnes, A., & Cooper, W. W. (1959). CHANCE-CONSTRAINED PROGRAMMING. Management Science, 6, 73-79.

Dantzig, G. B. (1955). LINEAR PROGRAMMING UNDER UNCERTAINTY. Management Science, 1, 197-206.

Darby, M. L., Nikolaou, M., Jones, J., & Nicholson, D. (2011). RTO: An overview and assessment of current practice. Journal of Process Control, 21, 874-884.

De Prada, C., & Valentin, A. (1996). Set point optimization in multivariable constrained predictive control. In 13 th World Congress of the International Federation of Automatic Control (Vol. D, pp. 351-356). San Francisco, USA.

Diwekar, U. M., & Kalagnanam, J. R. (1997). Efficient sampling technique for optimization under uncertainty. AIChE Journal, 43, 440-447.

Dochain, D., Perrier, M., & Guay, M. Extremum seeking control and its application to process and reaction systems: A survey. Mathematics and Computers in Simulation, In Press, Corrected Proof.

Dupacova, J. (1995). MULTISTAGE STOCHASTIC PROGRAMS - THE STATE-OF-THE-ART AND SELECTED BIBLIOGRAPHY. Kybernetika, 31, 151-174.

Dupačová, J., Consigli, G., & Wallace, S. W. (2000). Scenarios for Multistage Stochastic Programs. Annals of Operations Research, 100, 25-53.

Dupačová, J., Gaivoronski, A., Kos, Z., & Szántai, T. (1991). Stochastic programming in water management: A case study and a comparison of solution techniques. European Journal of Operational Research, 52, 28-44.

Engell, S. (2007). Feedback control for optimal process operation. Journal of Process Control, 17, 203-219.

Engell, S. (2009). Online Optimizing Control: The Link Between Plant Economics and Process Control. In C. A. O. d. N. Rita Maria de Brito Alves & Evaristo Chalbaud Biscaia, Jr. (Eds.), Computer Aided Chemical Engineering (Vol. Volume 27, pp. 79-86): Elsevier.

Findeisen, W., Bailey, F. N., Malinowski, K., & Tatjewski, P. (1980). Control and coordination in hierarchical systems.

Fleten, S.-E., & Kristoffersen, T. K. (2007). Stochastic programming for optimizing bidding strategies of a Nordic hydropower producer. European Journal of Operational Research, 181, 916-928.

Forbes, J. F., & Marlin, T. E. (1994). Model Accuracy for Economic Optimizing Controllers: The Bias Update Case. Industrial & Engineering Chemistry Research, 33, 1919-1929.

Forbes, J. F., Marlin, T. E., & MacGregor, J. F. (1994). Model adequacy requirements for optimizing plant operations. Computers & Chemical Engineering, 18, 497-510.

Gao, W., & Engell, S. (2005). Iterative set-point optimization of batch chromatography. Computers & Chemical Engineering, 29, 1401-1409.

Goel, V., & Grossmann, I. E. (2004). A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Computers & Chemical Engineering, 28, 1409-1429.

Gonzalez, A. I., Zamarreño, J. M., & De Prada, C. (2001). Non-linear model predictive control in a batch fermentator with state estimation. In European Control Conference. Porto, Portugal.

Grossmann, I. E., & Biegler, L. T. (2004). Part II. Future perspective on optimization. Computers & Chemical Engineering, 28, 1193-1218.

Grossmann, I. E., Halemane, K. P., & Swaney, R. E. (1983). Optimization strategies for flexible chemical processes. Computers & Chemical Engineering, 7, 439-462.

Guay, M., & Zhang, T. (2003). Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. Automatica, 39, 1283-1293.

Halemane, K. P., & Grossmann, I. E. (1983). Optimal process design under uncertainty. AIChE Journal, 29, 425-433.

Henrion, R., & Möller, A. (2003). Optimization of a continuous distillation process under random inflow rate. Computers & Mathematics with Applications, 45, 247-262.

Kadam, J. V., & Marquardt, W. (2004). SENSITIVITY-BASED SOLUTION UPDATES IN CLOSED-LOOP DYNAMIC OPTIMIZATION. In DYCOPS 7. Cambridge, USA.

Kall, P., & Wallace, S. W. (1997). Stochastic programming. Chichester; New York: Wiley.

Karuppiah, R., Martín, M., & Grossmann, I. E. (2010). A simple heuristic for reducing the number of scenarios in two-stage stochastic programming. Computers & Chemical Engineering, 34, 1246-1255.

Khor, C. S., Elkamel, A., Ponnambalam, K., & Douglas, P. L. (2008). Two-stage stochastic programming with fixed recourse via scenario planning with economic and operational risk management for petroleum refinery planning under uncertainty. Chemical Engineering and Processing: Process Intensification, 47, 1744-1764.

Krstic, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. Automatica, 36, 595-601.

Lee, J. H., & Yu, Z. (1997). Worst-case formulations of model predictive control for systems with bounded parameters. Automatica, 33, 763-781.

Li, P., Arellano-Garcia, H., & Wozny, G. (2008). Chance constrained programming approach to process optimization under uncertainty. Computers & Chemical Engineering, 32, 25-45.

Li, P., Wendt, M., Arellano-Garcia, H., & Wozny, G. (2002). Optimal operation of distillation processes under uncertain inflows accumulated in a feed tank. AIChE Journal, 48, 1198-1211.

Li, P., Wendt, M., & Wozny, G. (2002). A probabilistically constrained model predictive controller. Automatica, 38, 1171-1176.

Limon, D., Alamo, T., Raimondo, D., de la Peña, D., Bravo, J., Ferramosca, A., & Camacho, E. (2009). Input-to-State Stability: A Unifying Framework for Robust Model Predictive Control. In L. Magni, D. Raimondo & F. Allgöwer (Eds.), Nonlinear Model Predictive Control (Vol. 384, pp. 1-26): Springer Berlin / Heidelberg.

Lucia, S., Finkler, T. F., Basak, D., & Engell, S. (2012). A New Robust NMPC Scheme and Its Application to a Semi-Batch Reactor Example. In V. Kariwala, L. Samavedham & R. D. Braatz (Eds.), 8th IFAC International Symposium on Advanced Control of Chemical Processes (Vol. 8). Furama Riverfront, Singapore: IFAC.

Maestre, J. M., Muñoz de la Peña, D., Camacho, E. F., & Alamo, T. (2011). Distributed model predictive control based on agent negotiation. Journal of Process Control, 21, 685-697.

Magni, L., & Scattolini, R. (2007). Robustness and Robust Design of MPC for Nonlinear Discrete-Time Systems. In R. Findeisen, F. Allgöwer & L. Biegler (Eds.), Assessment and Future Directions of Nonlinear Model Predictive Control (Vol. 358, pp. 239-254): Springer Berlin / Heidelberg.

Marchetti, A., Chachuat, B., & Bonvin, D. (2009). Modifier-Adaptation Methodology for Real-Time Optimization. Industrial & Engineering Chemistry Research, 48, 6022-6033.

Marlin, T. E., & Hrymak, A. N. (1997). Real-time Operations Optimization of Continuous Processes. AICHE SYMPOSIUM SERIES, 156-164.

Martí, R., Navia, D., Sarabia, D., & De Prada, C. (2012). Shared Resources Management by Price Coordination. In B. Ian David Lockhart & F.

Michael (Eds.), Computer Aided Chemical Engineering (Vol. Volume 30, pp. 902-906): Elsevier.

Miller, B. L., & Wagner, H. M. (1965). CHANCE CONSTRAINED PROGRAMMING WITH JOINT CONSTRAINTS. Operations Research, 13, 930-&.

Mine, H., Ohno, K., & Fukushima, M. (1976). Decomposition of nonlinear chance-constrained programming problems by dynamic programming. Journal of Mathematical Analysis and Applications, 56, 211-222.

Morari, M., Arkun, Y., & Stephanopoulos, G. (1980). 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, 26, 220-232.

Navia, D., Sarabia, D., Marti, R., Gómez, E., De Prada, C., Lucia, S., & Sonntag, C. (2011). Preliminary report on the state of the art and research needs in hierarchical coordination in energy systems and processes. In: Technical report of Highly-complex and networked control systems (HYCON2) European Project.

Nemirovski, A., & Shapiro, A. (2006a). Convex Approximations of Chance Constrained Programs. SIAM Journal on Optimization, 17, 969-996.

Nemirovski, A., & Shapiro, A. (2006b). Scenario Approximations of Chance Constraints. In G. Calafiore & F. Dabbene (Eds.), Probabilistic and Randomized Methods for Design under Uncertainty.

Ochoa, S., Repke, J.-U., & Wozny, G. (2010). Integrating real-time optimization and control for optimal operation: Application to the bio-ethanol process. Biochemical Engineering Journal, 53, 18-25.

Ochoa, S., Wozny, G., & Repke, J.-U. (2010). Plantwide Optimizing Control of a continuous bioethanol production process. Journal of Process Control, 20, 983-998.

Pistikopoulos, E. N., & Ierapetritou, M. G. (1995). NOVEL-APPROACH FOR OPTIMAL PROCESS DESIGN UNDER UNCERTAINTY. Computers & Chemical Engineering, 19, 1089-1110.

. Preface. In. (2003)A. Ruszczynski & A. Shapiro (Eds.), Handbooks in Operations Research and Management Science (Vol. Volume 10, pp. v-vi): Elsevier.

Prékopa, A. (1970). On Probabilistic Constrained Programming. In Princeton Symposium on Mathematical Programming (pp. 113 - 138). Princeton, NJ: Princeton University Press.

Prékopa, A. (1995). Stochastic programming. Dordrecht ; Boston: Kluwer Academic Publishers.

Prékopa, A. (2003). Probabilistic Programming. In A. Ruszczynski & A. Shapiro (Eds.), Handbooks in Operations Research and Management Science (Vol. Volume 10, pp. 267-351): Elsevier.

Prékopa, A., Ganczer, S., Deák, I., & Patyi, K. (1980). The STABIL Stochastic Programming Model and its Experimental Application to the Electrical Energy Sector of the Hungarian Economy. In Stochastic Programming, Proceedings of the International Conference on Stochastic Programming (pp. 369-385). Oxford, England: Academic Press.

Prékopa, A., Vizvári, B., & Badics, T. (1998). Programming under probabilistic constraint with discrete random variable. In New trends in mathematical programming (pp. 235-255): Kluwer Acad. Publ.

Rawlings, J., & Mayne, D. Q. (2009). Model predictive control : theory and design. Madison, Wis: Nob Hill Pub.

Roberts, P. D. (1979). AN ALGORITHM FOR STEADY-STATE SYSTEM OPTIMIZATION AND PARAMETER-ESTIMATION. International Journal of Systems Science, 10, 719-734.

Rockafellar, R. T. (2001). Optimization Under Uncertainty. In. Washington: University of Washington.

Römisch, W. (2009). Scenario reduction techniques in stochastic programming. In Proceedings of the 5th international conference on Stochastic algorithms: foundations and applications (pp. 1-14). Sapporo, Japan: Springer-Verlag.

Rooney, W. C., & Biegler, L. T. (1999). Incorporating joint confidence regions into design under uncertainty. Computers & Chemical Engineering, 23, 1563-1575.

Rooney, W. C., & Biegler, L. T. (2001). Design for model parameter uncertainty using nonlinear confidence regions. AIChE Journal, 47, 1794-1804.

Rooney, W. C., & Biegler, L. T. (2003). Optimal process design with model parameter uncertainty and process variability. AIChE Journal, 49, 438-449.

Ruszczynski, A. R. (1998). Some advances in decomposition methods for stochastic linear programming. In 8th International Conference on Stochastic Programming (Vol. 85, pp. 153-172). Vancouver, Canada: Baltzer Sci Publ Bv.

Sahinidis, N. V. (2004). Optimization under uncertainty: state-of-the-art and opportunities. Computers & Chemical Engineering, 28, 971-983.

Sakalauskas, L. L. (2002). Nonlinear stochastic programming by Monte-Carlo estimators. European Journal of Operational Research, 137, 558-573.

Scattolini, R. (2009). Architectures for distributed and hierarchical Model Predictive Control - A review. Journal of Process Control, 19, 723-731.

Sequeira, S. E., Graells, M., & Puigjaner, L. (2002). Real-Time Evolution for On-line Optimization of Continuous Processes. Industrial & Engineering Chemistry Research, 41, 1815-1825.

Sequeira, S. E., Herrera, M., Graells, M., & Puigjaner, L. (2004). On-line process optimization: parameter tuning for the real time evolution (RTE) approach. Computers & Chemical Engineering, 28, 661-672.

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

Srinivasan, B., & Bonvin, D. (2002). Interplay between identification and optimization in run-to-run optimization schemes. In Proceedings of the 2002 American Control Conference, Vols 1-6 (pp. 2174-2179).

Tarhan, B., & Grossmann, I. E. (2008). A multistage stochastic programming approach with strategies for uncertainty reduction in the synthesis of process networks with uncertain yields. Computers & Chemical Engineering, 32, 766-788.

Tatjewski, P. (2002). Iterative Optimizing Set-Point Control - The Basic Principle Redesigned. In 15th Triennial World Congress, Barcelona, Spain: IFAC.

Voos, H. (2007). Resource Allocation in Continuous Production using Market-Based Multi-Agent Systems. In Industrial Informatics, 2007 5th IEEE International Conference on (Vol. 2, pp. 1085-1090).

Watkins, D. W., McKinney, D. C., Lasdon, L. S., Nielsen, S. S., & Martin, Q. W. (2000). A scenario-based stochastic programming model for water supplies from the highland lakes. International Transactions in Operational Research, 7, 211-230.

Wendt, M., Li, P., & Wozny, G. (2002). Nonlinear Chance-Constrained Process Optimization under Uncertainty. Industrial & Engineering Chemistry Research, 41, 3621-3629.

Würth, L., Hannemann, R., & Marquardt, W. (2009). Neighboring-extremal updates for nonlinear model-predictive control and dynamic real-time optimization. Journal of Process Control, 19, 1277-1288.

Würth, L., Hannemann, R., & Marquardt, W. (2011). A two-layer architecture for economically optimal process control and operation. Journal of Process Control, 21, 311-321.

2 STOCHASTIC PROGRAMMING

ABSTRACT

This Chapter describes two methods of stochastic economic programming and their application in a hydrogen consuming plant: Two – stage programming and Chance constrained optimization. We assume that the system under study presents sources of uncertainty that can be modeled with a binormal probability distribution function (PDF). Both dynamic optimization methods were expressed in the continuous time domain. To calculate the probabilistic constraints inverse mapping method was formulated as a nested estimation problem. On the other hand, to solve the two stage optimization, a discretization of the PDF in scenarios was applied with a scenario aggregation formulation to take into account the nonanticipativity constraints. Finally it was proposed a framework based in interpolation to generalize this solution. Both optimization methods were tested in terms of feasibility and optimality using Montecarlo simulation for the application case considered. The main problem appears to be the large computation times associated.

KEYWORDS

Two-stage stochastic optimization, Chance constraints optimization, Montecarlo simulation, Economic dynamic optimization, Hydrodesulfuration process

RESUMEN

Este capítulo describe dos métodos de programación estocástica y su aplicación a un problema de optimización económica en una hidrodesulfuradora: programación de dos etapas y optimización probabilística. Se ha supuesto que el sistema presenta fuentes de incertidumbre modeladas mediante una distribución de probabilidad binormal (PDF). Ambos métodos de optimización dinámica se han expresado en términos del dominio continuo del tiempo. Para calcular las restricciones probabilísticas, se ha utilizado el método de mapeo inverso reformulado como un problema de estimación de parámetros anidado. Por otro lado, se ha utilizado el método de agregación de escenarios en la programación de dos etapas para considerar las restricciones de no anticipatividad. Finalmente, se propone un método para la aplicación de los resultados obtenidos basado en la interpolación de los mismos, con la idea de generalizar la solución discreta obtenida en cada escenario. Los resultados, han sido probados en la planta de hidrodesulfuración mediante simulaciones de Montecarlo para dar una idea de la factibilidad en una aplicación real

PALABRAS CLAVES

Programación de dos etapas, Optimización probabilística, Simulaciones de Montecarlo, Optimización dinámica, Hidrodesulfuración

2.1 Introduction

Uncertainty is always present in the operation of processes. Therefore, when optimal decisions have to be made, differences between the model and the reality must be considered in order to propose optimal and feasible policies. In some situations the uncertainty that affects the system can be attained to the random behavior of some process variables, which must considered explicitly in the optimization problem to propose a feasible policy to be applied in a real process. In this chapter, we have used two approaches of stochastic programming to overcome this problem in a test problem related to the optimization of the hydrogen consumption of a desulfuration unit from a petrol refinery, when we are facing changes in the operating conditions. Stochastic programming uses a probabilistic viewpoint to propose a feasible solution in an optimization problem when some of the variables present a random behavior. As it was commented in the state of the art, there are two main ways to solve this problem: multistage (and its particular case with two-stages) and chance constrained programming. In this thesis both methods have been used.

In contrast to other approaches that appear in the literature, the stochastic dynamic programming problem has been solved in the continuous domain, using a sequential approach. To do this, a control vector parameterization was used combining optimization methods and dynamic simulation. Two changes have been proposed in the stochastic methods according to the continuous formulation. In the implementation step of the two-stage one, an interpolation method is presented to overcome the loss of generalization that takes place when scenarios are used to describe the continuous PDF and an open loop policy must be applied. In the same way, in the chance constraint method a new approach for calculating the limits of the probability integrals as the solution of a parameter estimation problem has been proposed. The optimization results obtained with these methods were tested and analyzed using Montecarlo simulations.

The structure of this chapter is as follows: Section two presents a brief introduction to the stochastic optimization methods. Section three shows the description of a hydrodesulfuration plant and the application of the stochastic optimization methods to its optimal operation. In section four, the outcomes of the stochastic optimization with a test using Montecarlo simulations are presented and discussed, in order to evaluate the generalization method proposed. The chapter ends with some conclusions and comments about future work.

2.2 Generalities of Stochastic Programming

In general, a problem of dynamic optimization under uncertainty can be summarized as:

$$\begin{aligned} &\min_{u} f(x,u,\xi,t_f) \\ &s.t.: \\ &h(\dot{x},x,u,\xi,t)=0, \qquad x(t_0)=x_0 \\ &g(x,u,\xi,t)\leq 0 \\ &x\in X,\ \ u\in U,\ \ \xi\in \Xi, \qquad t\in[t_0,t_f] \end{aligned} \tag{2.1}$$

Where $x \in \mathbb{R}^{n_x}$ is the vector of states, $u \in \mathbb{R}^{n_u}$ are the decision variables, t is time and $\xi \in \mathbb{R}^{n_\xi}$ represents a random variable that belongs to a probability space with a certain probability distribution function (PDF) Ξ. The model of the process is given by the set of equations $h: \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \to \mathbb{R}^{n_x}$, the cost function to be minimized is represented by $f: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \to \mathbb{R}$, while $g: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \to \mathbb{R}^{n_x}$ denotes the constraints of the optimization problem.

Many practical problems can be formulated as equation (2.1) states due to the presence of unknown elements. The nature of the random behavior of the uncertain variables can be very different, ranging from fairly constant but unknown values (e.g. compositions) to values that change continuously in a casual way (e.g. wind). The decision variables can also be very different, and here it is assumed that they will take values over a time horizon

according to an input vector parameterization. With respect to the solution methods, two approaches have been chosen considering the application target.

2.2.1 Two – Stage Programming

Two stage programming is a particularization of the multistage programming. In two–stage formulation the key idea is: we need to make a decision now taking into account that after a certain period of time, more information will be available as measurements that will contribute to decrease the incertitude from that time on. So, when decisions have to be made over a time horizon, there are stages of decisions that differ in the degree of knowledge of the uncertain variable: in the first one (stage 0), a choice must be made knowing the initial conditions of the system and without any certainty about the random variables, except that they belong to a certain PDF. Then, in the second stage (stage 1), the decision variables can be chosen taking into account that the value of the random variable is available (measured or estimated) (Dantzig, 1955), and can be equated to the value that it had in the previous stage. Therefore, the decision variable in second stage will depend on the earlier values and on the random variables. If the subscripts denotes the decision stages (0 and 1), the general problem from equation (2.1) can be reformulated from the two-stage point of view as (Rockafellar, 2001):

$$
\begin{aligned}
&\min_{u(\cdot)} \mathbb{E}_{\xi}[f_0(u_0, x_0(\xi), \xi) + f_1(u_1(\xi), x_1(\xi), \xi)\,] \\
&s.t.: \\
&g_0(u_0, x_0(\xi), \xi) \le 0 \\
&g_1(u_1(\xi), x_1(\xi), \xi) \le 0 \\
&h_0(\dot{x}_0(\xi), x_0(\xi), u_0, \xi, t) = 0, \qquad x_0(\xi, t_0) = x_{0,i}, \qquad t \in [t_0, t_1] \\
&h_1(\dot{x}_1(\xi), x_1(\xi), u_1(\xi), \xi, t) = 0, \qquad x_{1,i}(\xi, t_1) = x_0(\xi, t_1), \\
&\qquad\qquad t \in (t_1, t_f] \\
&x_1 \in X_k, \;\; u_k \in U_k, \qquad k = \{0,1\} \\
&\xi \in \Xi
\end{aligned}
\tag{2.2}
$$

In equation (2.2), $u_0 \in \mathbb{R}^{n_u}$ and $u_1(\xi) \in \mathbb{R}^{n_u}$ represent the decision variables applied in stages 0 and 1 respectively. Notice that these variables can take several values in time according to a certain parameterization in each stage, but the notation has been shortened for simplicity. The initial states for both stages are represented as $x_{0,i} \in \mathbb{R}^{n_x}$ and $x_{1,i}(\xi) \in \mathbb{R}^{n_x}$ for stage 0 and 1 respectively, being the second one obtained as a result of the decision variables applied. It can be noted how the value of the uncertain parameter affects the evolution of the state variables in both stages. The functions that must be minimized are denoted as $\{f_0, f_1\}$: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \to \mathbb{R}$ for each decision stage. Due to the fact that these functions depend on the value of the random variables, their expected value ($\mathbb{E}_\xi$) must be used to group all these situations in a single objective function to be minimized. Analogous to the objective function, the model of the process and the inequality constraints are represented by $\{h_0, h_1\}$: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \to \mathbb{R}^{n_x}$ and $\{g_0, g_1\}$: $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \times \mathbb{R} \to \mathbb{R}^{n_g}$ for both stages respectively.

In order to solve this problem for a continuous PDF it is necessary to solve a nested numerical integration (Birge & Louveaux, 1997). To avoid this step, it is possible to discretize the original PDF (**D-PDF**) allowing only a finite number of values for the random variable, called scenarios, and then solve problem from equation (2.2) using a weighted sum of the cost function (Birge & Louveaux, 1997; Dupacova, 1995; Dupačová et al., 2000; Ruszczynski & Shapiro, 2003; Sahinidis, 2004). The problem with the scenario approach is the loss of generalization in the optimization because the solution is valid only for the discrete values considered in the D-PDF. However, it allows obtaining solutions that otherwise will not be available numerically.

In this work, the scenario formulation was used with the two-stage approach described previously. From the point of view of the uncertain variable ξ, the situation can be represented in the schematic of Figure 2.1(a) where in stage 0 we must consider that it can have any value of the

scenarios chosen, while in stage 1 the value of the random variable is equal to the one from the first stage. In contrast, Figure 2.1(b) shows different realizations that could take place over time in different stages if the information that can be acquired in the future is not considered, leading to an exponential increase in the number of scenarios.

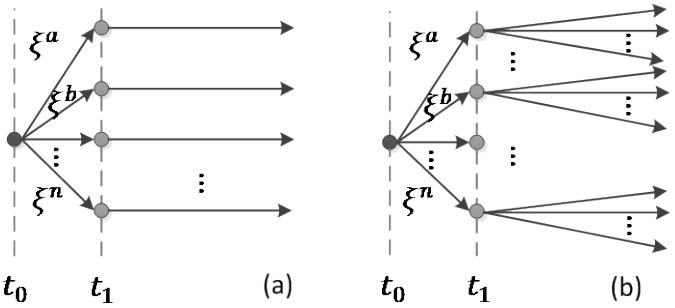

FIGURE 2.1: INFORMATION ABOUT THE STOCHASTIC VARIABLE IN TWO-STAGE APPROACH

The selection of one or other implementation strategy is related to how the uncertain variables affect the process. In the first case it is considered that the uncertainty only disturbs the system from the actual time. On the other hand, in the second case there are additional sources of uncertainty that will appear in the future. To understand this idea, consider an investment problem to produce two kinds of crop plants with different yields depending on the weather. In the first stage we must take a decision related with the proportion to seed that is affected by the weather, while in the second stage, we must decide the selling quantities of each product, which depends on the market prices. As we can see, in this example there are to sources of uncertainty that appear in different times, therefore we must represent the two-stage implementation of this problem as Figure 2.1(b). Now imagine that before seeding the field, we have a customer that buys the entire harvest at a fixed price. In this case, we have not the second stage uncertainty, which implies that the stochastic optimization problem must be implemented as Figure 2.1(a).

For the decision variables, Figure 2.2 shows that in the stage 0, it is possible to choose only a single policy, considering all possible values of ξ. On the other hand, in stage 1 the choice can be made according to the realization of the random variable considered in stage 0, because is already known. Of course, the computation of the optimum decision variables over the entire time horizon is made at current time t_0 considering the value of the cost function over both stages.

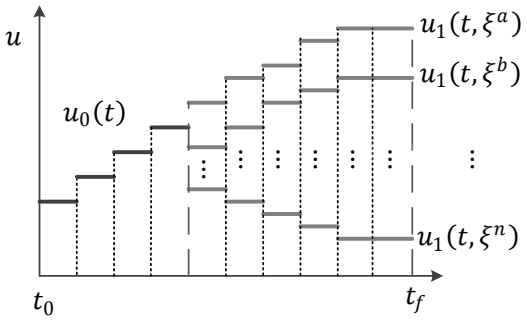

FIGURE 2.2: DECISION VARIABLES IN THE TWO STAGES

In the context of real time process optimization, the application of this method would be made according to a receding horizon policy, computing u_0 and u_1 at current time but applying only u_0 and computing again the solution at the next sampling period. Nevertheless, there are situations in which the open loop solution must be applied, either due to the nature of the system or because an on-line solution is not available. In these cases, the second stage solution $u_1(\xi^i)$ is only available for the discrete values ξ^i which may not correspond to its estimated value in the second stage. To overcome this problem, several policies can be chosen as the nearest neighbor or oversizing the discrete solution. This is shown to be more important than expected, as one can see in the Montecarlo test performed in section 2.4.3, here a method based on interpolations is proposed.

To see how it works, notice that as the solution of the second stage depends on the value of the random variable. This can be represented as an optimal function that depends on the uncertain variables, i.e., $u_1^\star = f_{opt}(\xi)$. If the random variables have a continuous nature, it is expectable that f_{opt} would be continuous too (for a continuous process), but when the problem is solved discretizing the probability function only some values of this optimal function can be known. This will lead to a loss of generality in the application of the optimal solution, because the uncertain variables are continuous and the solution is available only for a finite number of scenarios. Therefore, if there is a way to connect these solution points guaranteeing a definite degree of feasibility and optimality in the operation, for all the possible values included in the original PDF, the problem can be solved with a lower complexity in the optimization by using an acceptable approximation of the continuous solution, like Figure 2.3 shows.

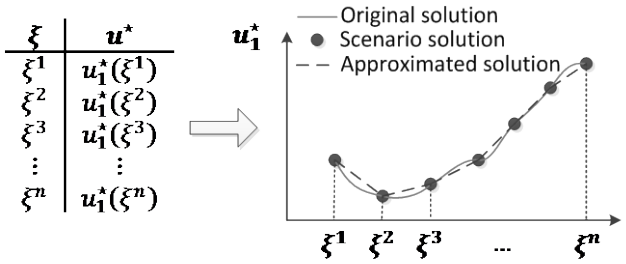

FIGURE 2.3: APPROXIMATED SOLUTION OBTAINED FROM SCENARIO APPROACH

To implement the previous idea, first it is necessary to solve the stochastic optimization problem with the D-PDF. This will lead to a particular solution for each scenario considered in the D-PDF. As we are considering an open loop solution, in order to select the best interpolation method, some of them can be tested in a Montecarlo simulation with the continuous probability distribution function. Once that all the interpolation methods are tested the one that guarantees feasibility and optimality will be chosen for the final open-loop implementation. Figure 2.4 summarizes the

interpolation methodology proposed to generalize the second stage solution obtained for each scenario.

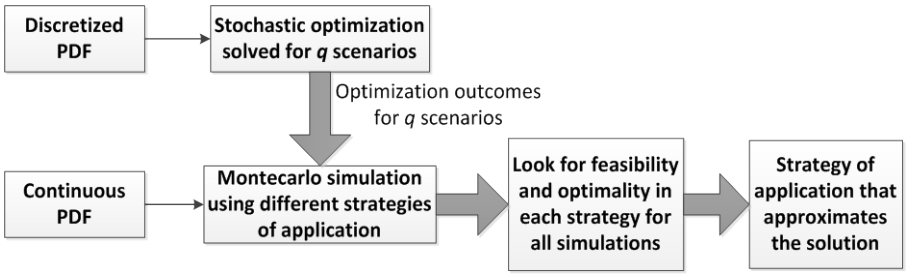

FIGURE 2.4: DIAGRAM OF THE METHOD PROPOSED TO APPLY THE SECOND STAGE SOLUTION

2.2.2 Chance Constraint formulation

The works of Charmes and Cooper (Charnes & Cooper, 1959), Miller and Wagner (Miller & Wagner, 1965) and Prékopa (András Prékopa, 1970), present the method of optimization based in probabilistic constraints. This formulation can be understood as: finding some decision variables that minimize a given objective function, ensuring a probability of feasibility greater than a certain level of confidence.

In general, a problem of optimization under uncertainty has the form of equation (2.1). Referring to the constraints g, in the process industry it is very common to impose limits to the output variables (y) like the temperature, pressure, mole fractions, etc. guaranteeing that these variables must stay within a certain range, i.e.:

$$y_i^L \leq y_i(u,\xi) \leq y_i^U, \qquad i = 1 \dots I \tag{2.3}$$

The uncertainty that affects the system is propagated itself throughout the model. Therefore, it is expectable that the output variables will have a random behavior too. Thus, the output constraints in equation (2.3) can be redefined in order to ensure a probability of feasibility greater than a confidence level α, as equation (2.4) states.

$$\Pr(y_i^L \leq y_i(u,\xi) \leq y_i^U) \geq \alpha_i\,, \qquad i = 1 \dots I \tag{2.4}$$

Given the values of the decision variables, the model h must be satisfied for any value that the random variable can take (if the problem is feasible). Therefore the state variables can be seen as a projection of ξ, which is equivalent to say that for a given value of the decision variables, the states are function of the random variables. We can also note that for a given value of ξ, a sequential resolution of the optimization problem in equation (2.1) can be stated. Sequential resolution means solving the optimization problem using an intermediate step of simulation, that is to say: for a certain value of the decision variables given by an optimization algorithm, the state variables are calculated simulating the model of the process. This procedure can be repeated for different values of the random variables obtaining a mapping of the state variables and the inequality constraints with respect to the uncertain variables as a function of the decision variables. With this approach we are able to transfer the equality constraints from the optimization to simulation step, using the model of the process to calculate the state variables for a given value of the decision and the random variable (Harvey Arellano-Garcia & Wozny, 2009; Li et al., 2008). The sequential approach transforms the chance optimization problem to equation (2.5).

$$\begin{aligned} &\min_u \mathbb{E}_\xi\left[f\left(h\left(x,u,\xi,t_f\right),u,\xi,t_f\right)\right] \\ &s.t.: \\ &\Pr(y_i^L \leq y_i(u,\xi) \leq y_i^U, i = 1 \dots I) \geq \alpha \\ &u \in U, \xi \in \Xi \end{aligned} \tag{2.5}$$

Now, if we want to solve the optimization problem from equation (2.5) it is necessary to calculate the probability of the output variables y. This is not a trivial task since the PDF of these variables is in general unknown (commonly the PDF is only available for ξ). To overcome this problem, several methods of solution have been proposed that assumes convexity or linearity (Mayne & Polak, 1976; Nemirovski & Shapiro, 2006; András

Prékopa, 1970; A. Prékopa, 1995; András Prékopa, 2003). Other methods based on sampling and genetic algorithms had been suggested too (Loughlin & Ranjithan, 2001; Poojari & Varghese, 2008; Yang & Wen, 2005). On the other hand, the chance constrained optimization can be approached using an inverse mapping of the random variables to estimate the PDF of the output variable (Harvey Arellano-Garcia et al., 1998; H. Arellano-Garcia et al., 2003; Harvey Arellano-Garcia & Wozny, 2009; Barz et al., 2008; Li et al., 2008; Li et al., 2002; Li et al., 2003; Wozny & Arellano-Garcia, 2007). This method has no assumptions about the nonlinearities in the system and is not based in sampling for each step of optimization, but a regularity condition is required that guarantee monotony in the response of the model output with respect to the uncertain variable.

The idea of the inverse mapping methodology is to translate the calculus of the probability of the chance constrained output satisfying the constraints, to the probability of the stochastic variables (which are known) being within a certain range. To do this, it is necessary to obtain an equivalent of the bounds registered in the chance constraints called ξ^*, where its probability of occurrence is the same. Then the probability can be calculated as:

$$\Pr(y \leq y^*) = \Pr(\xi \lesseqgtr \xi^*) \tag{2.6}$$

Where y^* can be y^U and/or y^L. A method for computing ξ^* will be presented in Section 2.3.5.

Next, both formulations of optimization under uncertainty will be applied to a case study taken from a process of the refining industry.

2.3 Problem Formulation

2.3.1 Process Description

In petroleum refineries the hydrodesulfuration process is used to remove sulfur from the hydrocarbons to fulfill environmental policies. Another

important reason for removing sulfur from the intermediate product naphtha streams within a petroleum refinery is that sulfur, even in extremely low concentrations, poisons the noble metal catalysts (platinum and rhenium) in the catalytic reforming units that are subsequently used to upgrade the octane rating of the naphtha streams.

To perform the desulfuration reaction, hydrogen is put in contact with the hydrocarbon in fixed bed reactors with a specific catalyst (Commercial Ni-Mo based catalyst) (Bellos & Papayannakos, 2003). The optimal management of the hydrogen provided is very important in order to operate efficiently and safely. That is to say: if the quantity of hydrogen supplied is less than the minimum required, then the catalysts used in the desulfurization reactors can suffer important damage. On the other hand, if the supply is in excess significant economic losses will be experienced (Sarabia et al., 2009).

To understand the process, let us consider the simplified structure of the core part of a hydrodesulfuration plant represented in Figure 2.5. It can be seen that the hydrogen come from three sources: H4, H3 and LP. H4 and H3 are collectors transporting hydrogen manufactured in especially dedicated production units where methane and steam are reformed. Each production unit can produce the hydrogen in different quantities and purity levels. On the other hand, the LP source is a recirculated stream with low hydrogen concentration that can proceed either from side units where the hydrogen is a by-product from other reactions, or from units where the excess hydrogen fed in the reactor is partially purified using membranes (Gómez et al., 2008). The mixture goes through a compressor and then is mixed with the hydrocarbon stream (F_C) in order to react in a packed bed reactor (R-1, R-2). The products of the reaction are separated using a flash tank (T-1). One part of the excess of hydrogen is recirculated to the reactors while other (F_{10}) is purged in order to maintain a minimal purity because of the catalyst.

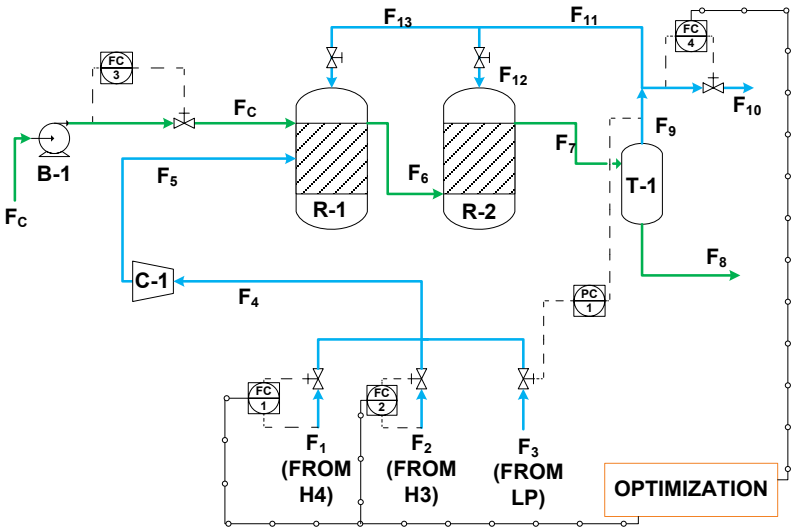

FIGURE 2.5: DIAGRAM OF THE HYDRODESULFURATION UNIT WITH DECISION VARIABLES

The key operation is performed in the reactor to eliminate the undesired sulfur down to a given level at the plant output. The operators adjust the total hydrogen supply to the reactor, its temperature, etc. to attain this target, so that a given mode of operation implies certain hydrogen consumption in the reactor according to the load conditions. This is a sub-system where settling times can be in the order of hours. To carry out the operation, the operators (and the basic control system) modify the different hydrogen streams (fresh and recycled) in order to maintain the required supply to the reactor(s) fulfilling the constraints imposed by compressors, purities and catalysts. Since there are different ways of providing the same amount of hydrogen that the reactor is consuming, the operational target is to supply the required flow to the reactors using the best combination of these sources (from an economic point of view), satisfying the set of operational constraints.

Several problems are related to the hydrogen management and optimization that are worth to mention; among them, the lack of reliable

information about many streams and compositions, the uncertainty of the demands and the large scale of the system that creates additional difficulties. Regarding the first one, it is clear that reliable information from the network is required if one wishes to perform optimal decisions. This covers several aspects, on one hand for accounting and to compare different ways of functioning, and on the other hand as a basic element of any model based optimization. Part of this uncertainty comes from the measurement system, where there are many unmeasured variables and some instruments may need better calibration. In fact, the main problem is related to partial measurements, in particular, gas flows are usually measured with volumetric flow meters that require compensation in order to convert the readings to mass or normalized flows used in models based on mass balances (Kelly & Mann, 2005). This compensation involves pressure, temperature and molecular weight of the streams. Nevertheless, the last one is quite often unavailable because of the price and the reliability of the instruments measuring hydrogen purity.

The other significant source of uncertainty comes from the changes in the flow, or the composition, of the hydrocarbon streams being treated in the hydrodesulfuration plants, which are linked to the precedence of the oil crude or to production policies. As mentioned before, the precedence of the crude oil change quite often, which modify the hydrogen demands in the reactor. A certain planning of the most important changes in the flow of the different streams of hydrocarbons is performed by the production planning department of the refinery, but still uncertainty exists because its compositions that determine the specific hydrogen consumption are not very well known. A typical pattern in the operation of a HDS plant is a transient lasting some hours followed by a stable demand. This situation becomes critical when a change in the hydrocarbon load takes place. This happens approximately every two days when new products are processed and the plant takes some hours to stabilize in the new operating point, which is desired to be optimum.

Therefore, it can be concluded from the description of the operation of the HDS that the optimization must be stated when a change in the load is produced, in order to estimate the correct policy to perform the transition and the optimal operating point at the end of the time horizon. As it was mentioned, in this situation the main sources of uncertainty that affects the process are: the composition of the hydrocarbon load that can change according to the type of oil that is processed and the composition of the hydrogen supply, and in particular the LP stream since it depends on the operation of other units (Sarabia et al., 2011).

2.3.2 Model of the Process

The internal operation of HDS involve many aspects an is the responsibility of the operators to maintain the hydrogen level above the minimal requirements to produce the desired reaction in spite of hydrocarbon load changes in a safe way. The internal mechanisms of the hydrodesulfuration process are quite complex (Gómez et al., 2008) but a simplified model will be used to represent its dynamics. The evolution of the hydrogen consumption inside the Reactors ($F_X^{H_2}$) is approximated by a first order dynamics as equation (2.7) shows. This consumption depends on the flow of hydrocarbon to be desulfurized (F_{HC}) and the stochastic variable (ρ), which represents the specific hydrogen consumption rate as a characteristic of the sulfur content in the hydrocarbon to be treated.

$$\tau \frac{dF_X}{dt} + F_X = F_{HC}\rho, \qquad F_X^{H_2}(t_0) = F_{X0}^{H_2} \tag{2.7}$$

For the mixing point of the three hydrogen sources and the flow through the compressor, the total and hydrogen mole balances can be defined.

$$\begin{aligned} &F_1 + F_2 + F_3 = F_5 \\ &F_1X_1 + F_2X_2 + F_3X_3 = F_5X_5 \end{aligned} \tag{2.8}$$

Where X_1 and X_2 are known hydrogen compositions and F represent flows according to Figure 2.5.

Inside the reactor, the pressure is maintained constant by changing the flow from LP, therefore the dynamics of the total mole can be neglected unlike the hydrogen concentration that can change in time:

$$\begin{aligned} 0 &= F_5 - F_{10} - F_X^{H2} \\ \frac{V}{ZRT}\left[P\frac{dX_{H2}}{dt}\right] &= F_5 X_5 - F_X - F_{10}X_{H2}, \qquad X_{H2}(t_0) = X_{H2,0} \end{aligned} \tag{2.9}$$

In addition, some operational constraints must be defined:

The purity of the mix stream that enters to the desulfuration unit and the one inside the reactor must be greater than a lower bound, because of the compressor an the catalyst requirements respectively.

$$\begin{aligned} X_5^L &\leq X_5 \\ X_{H2}^L &\leq X_{H2} \end{aligned} \tag{2.10}$$

The flows of the hydrogen streams must be within a certain range, which is given by the production limits.

$$\begin{aligned} F_1^L &\leq F_1 \leq F_1^U \\ F_2^L &\leq F_2 \leq F_2^U \\ F_3^L &\leq F_3 \leq F_3^U \end{aligned} \tag{2.11}$$

The nomenclature of the model used from equation (2.7) to (2.11) is presented in Table 2.1.

TABLE 2.1 NOMENCLATURE EMPLOYED ON THE HYDROGEN CONSUMPTION UNIT

Variable	Meaning	Units
F_i	Molar flow of stream i	$kmol/h$
F_{HC}	Flow of hydrocarbon to be treated	m^3/h
F_X^{H2}	Consumption of H_2 inside the reactor	$kmol/h$
X_i	Molar fraction of H_2 in stream i	Adimensional
X_{H2}	Molar fraction of H_2 inside the reactor	Adimensional
τ	Time constant of the H_2 consumption	h
ρ	Ratio hydrogen/hydrocarbon	$kmol/m^3$
V	Reactor volume	m^3
P	Pressure in the reactor	atm
T	Temperature in the reactor	K
R	Universal constant of gases	$atm \cdot m^3/K\ kmol$
Z	Compressibility factor	Adimensional
F_i^L, F_i^U	Bounds of the flow rate of stream i	$kmol/h$
X_i^L	Bound of the molar fraction of H_2 in stream i	Adimensional

2.3.3 Optimization Problem

As it was stated in the process description, the optimization problem is formulated in the framework of economic dynamic optimization and it consists in finding the best combination of hydrogen sources that produces the desulfuration of a given hydrocarbon when a change in the load is produced, fulfilling the constraints of the process. The objective function is given by equation (2.12), where C_{H4} and C_{H3} are the costs of pure hydrogen from each fresh hydrogen source. This cost must be minimized over the entire transient that involves the load change. The dynamic model of the plant has three degrees of freedom, which means that if three of the variables have a fixed value and the random variables are known, the complete system is determined. Therefore, it can be said that the optimization problem based in this model has three decision variables, which are chosen as the same ones that the operators can modify in the plant: F_1, F_2 and F_{10}. Hence, the dynamic optimization problem can be summarized as: compute the flows F_1, F_2 and F_{10} that minimize the equation (2.12), subject to the model of the process and the operational

constraints represented from equations (2.7) to (2.11), in spite of the uncertainty of ρ and X_3, for a given change in the hydrocarbon load.

$$Cost = \int_{t_0}^{t_f} C_{H4} X_{H4} F_1 + C_{H3} X_{H3} F_2 \, dt \tag{2.12}$$

For the nominal values of Table 2.2, a deterministic dynamic optimization has been solved using the single shooting method with control vector parameterization.

TABLE 2.2 VALUE OF THE PARAMETERS, BOUNDS AND INITIAL CONDITIONS FOR DETERMINISTIC OPTIMIZATION

Parameter	Value	Parameter	Value
C_{H4}	88100	t_f	10
C_{H3}	77000	T	623.15
ρ	12.6	F_3^L	0
F_{HC}	102	F_3^U	5000
τ	0.3	X_5^L	0.9
V	100	X_{H2}^L	0.7
P	68	F_1^L	0
$F_{X^0}^{H2}$	682.5	F_1^U	1400
X_{H2}^0	0.9	F_2^L	0
X_1	0.991	F_2^U	790
X_2	0.931	F_{10}^L	0
X_{LP}	0.85	F_{10}^U	1500

The optimal trajectories of the decision and the state variables are presented in Figure 2.6. These trajectories were obtained applying the single shooting approach, for a time horizon of 10 hours. The decision variables were discretized in 16 periods with constant values. The optimization took 1.5 minutes.

Results are as one can expect: because the purities are above their lower limits at the initial time, the optimal trajectories of the decision variables try to bring the hydrogen concentrations to their bounds. To do this, first the

hydrogen purge is completely closed to accumulate impurities inside the reactor and to consume less fresh hydrogen, optimizing the cost function. Once that the purity inside the reactor reaches its minimal value the optimal policy increases the flow of the purge with the aim to maintain the purity constraint active. About the hydrogen fed it can be noted that only the stream coming from H4 is used, because the ratio purity/cost is lower in this unit than in H3.

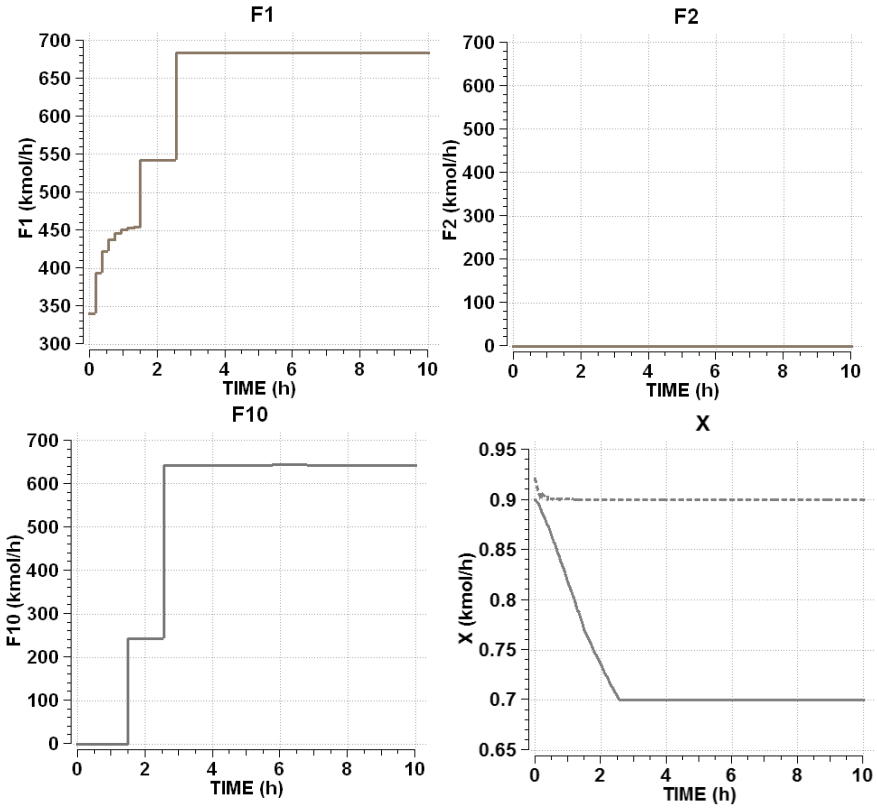

FIGURE 2.6: OPTIMAL RESULTS OF DETERMINISTIC PROBLEM

Nevertheless, in practice there are two main sources of uncertainty that add difficulty to the operation and can produce infeasibilities: (1) the demands of hydrogen in the reactors depend on the quality of fuel loaded,

which generally is an unknown mixture, and (2) the purity of the recirculation stream that comes from other processes can suffer changes over time. Using historical data from a refinery, these variables can be modeled using a binormal probability distribution function (Figure 2.7 and equation (2.13)). In this graph ξ_1 and ξ_2 are the random variables that correspond to the specific hydrogen consumption ρ and to the purity X_3 in the LP stream respectively.

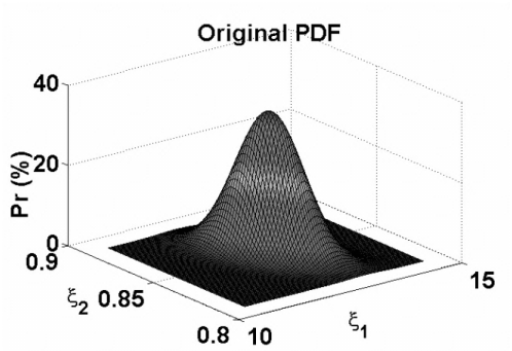

FIGURE 2.7: PROBABILITY DISTRIBUTION OF THE RANDOM VARIABLES

$$\Xi(\xi_1,\xi_2) \coloneqq \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r}}\exp\left(-\frac{1}{2(1-r)}\left[\frac{\mu_1-\xi_1}{\sigma_1} - \frac{(\mu_1-\xi_1)*(\mu_2-\xi_2)}{\sigma_1\sigma_2} + \frac{\mu_2-\xi_2}{\sigma_2}\right]\right) \quad (2.13)$$

Being μ_i and σ_i the mean and standard deviation of the random variable ξ_i, while r is the correlation index.

A Montecarlo simulation can applied to the results from the deterministic optimization presented in Figure 2.6, in order to see how the random behavior of the uncertain parameters ξ_1 and ξ_2 affect the constrained purities. The results of these simulations are presented in Figure 2.8, showing their evolution in time and probability distribution.

From the simulation, it can be noted that the distribution of purities is centered in their lower bound, but approximately the 50% of the times the system is operating in the infeasible part of the optimization problem, decreasing the life cycle of the catalysts and the compressors, which is not an optimal way to operate.

Consequently, an adequate management of these uncertainties is required in order to ensure an optimal and feasible operating policy taking into account the random nature of these variables, which leads to the use of stochastic optimization in this system.

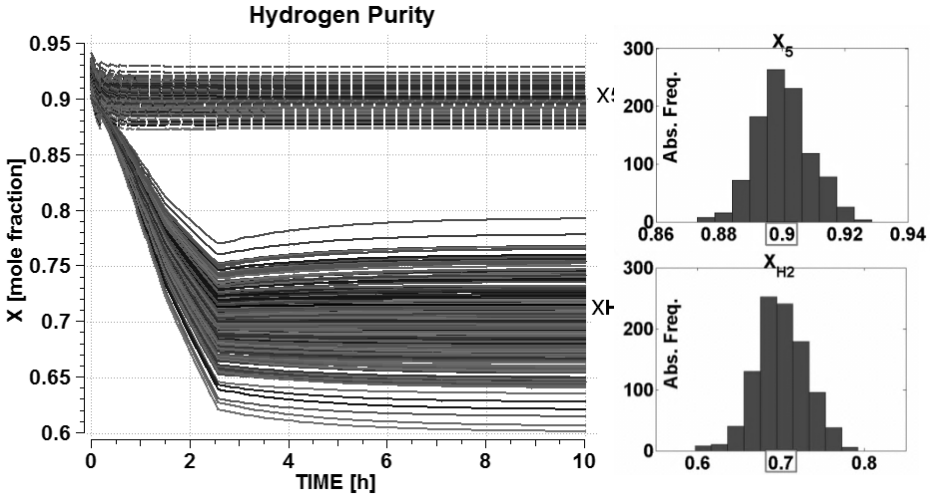

FIGURE 2.8: HYDROGEN PURITY OBTAINED IN MONTECARLO SIMULATION

2.3.4 Two – Stage Formulation in the HDS

In order to set up the stochastic problem using this formulation, it is necessary to define the stages of decision and the information that can be available in each of them. In the first stage (stage 0) the hydrogen flows and the purge $F_{1,0}$, $F_{2,0}$ and $F_{10,0}$, will be decided knowing only the PDF of the random variables. In the second stage, the value of these random variables is assumed to be available (measured in the laboratory or estimated), and

the decisions about the hydrogen production for the second stage: $F_{1,1}$, $F_{2,1}$ and $F_{10,1}$, will be calculated as a function of them. In terms of time, the first stage goes from t_0 to t_1, which is a reasonable period for estimating (or measuring) the real value of the random variables.

Now the original optimization problem presented before must be reformulated from the point of view of the two-stage approach. To do this, a discretization of the PDF in 25 values (5 for each variable) was performed in order to solve the problem by using the scenario approach. The discretized PDF is shown in Figure 2.9.

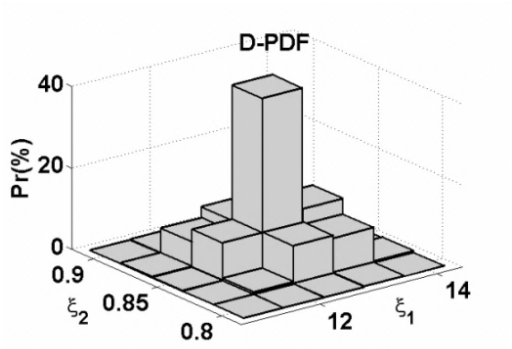

FIGURE 2.9: SCENARIO DISCRETIZATION OF THE PDF

Equation (2.14) shows the stochastic reformulation of the problem presented from the point of view of the two stage representation, where k denotes the stage of decision and *j* the scenarios, N_{Sc} represents the total number of scenarios, and the probability of occurrence for each of them is defined as $\Pr(j)$.

It can be noted from equation (2.14a) that the decision variables from the first and the second stages can have as many values as scenarios are. However, during the first stage the random variable has not been estimated yet, hence a single decision must be computed, which implies that each decision variable $F_{1,0}^{j}$, $F_{2,0}^{j}$ and $F_{10,0}^{j}$ must have the same value for all the

scenarios. This condition is called "nonanticipativity" and it is represented in the last constraint of the optimization problem from equation (2.14b), where if the problem is feasible, the first stage decision in every scenario must be equal to the most representative one (since it is multiplied by its probability of occurrence). This large optimization problem can be separated in as many problems as scenarios are by using the scenario aggregation method (Kall & Wallace, 1997). To do this it is necessary to eliminate the constraints that relate all the scenarios in the first stage. This can be done using the Augmented Lagrangian method.

$$
\begin{aligned}
&\min_{F_1,F_2,F_{10}} \sum_{j\in N_{Sc}} \Pr(j)\left\{\int_{t_0}^{t_1} C_{H4}X_{H4}F_{1,0}^{j} + C_{H3}X_{H3}F_{2,0}^{j}dt \right.\\
&\qquad\qquad \left. + \int_{t_1}^{t_f} C_{H4}X_{H4}F_{1,1}^{j} + C_{H3}X_{H3}F_{2,1}^{j}dt\right\}\\
&s.t.:\\
&F_{1,k}^{j}X_1 + F_{2,k}^{j}X_2^{j} + F_3^{j}\xi_2^{j} = F_5^{j}X_5^{j}\\
&F_{1,k}^{j} + F_{2,k}^{j} + F_3^{j} = F_5^{j}\\
&0 = F_5^{j} - F_{10,k}^{j} - F_X^{H2\,j}\\
&\frac{VP}{ZRT}\left[\frac{dX_{H2}^{j}}{dt}\right] = F_5^{j}X_5^{j} - F_X{}^{j} - F_{10,k}^{j}X_{H2}^{j}, \qquad X_{H2}^{j}(t_0) = X_{H2_0}\\
&\tau\frac{dF_X^{j}}{dt} + F_X^{j} = F_{HC}\xi_1^{j}, \qquad F_X^{j}(t_0) = F_{X_0}\\
&t \in [t_0, t_f]\\
&F_3^{L} \le F_3^{j} \le F_3^{U}\\
&X_5^{L} \le X_5^{j}, \qquad X_{H2}^{L} \le X_{H2}^{\mathrm{j}}\\
&F_1^{L} \le F_{1,k}^{j} \le F_1^{U}\\
&F_2^{L} \le F_{2,k}^{j} \le F_2^{U}\\
&\xi_1^{j}, \xi_2^{j} \in \Xi\\
&k = \begin{Bmatrix} 0, & t \in [t_0, t_1) \\ 1, & t \in [t_1, t_f] \end{Bmatrix}, \qquad j = 1 \dots N_{SC}
\end{aligned}
\tag{2.14a}
$$

$$
F_{m,0}^{j} = \frac{\sum_{i=1}^{N_{Sc}} \Pr(i)\, F_{m,0}^{i}}{\sum_{i=1}^{N_{Sc}} \Pr(i)}, \qquad m = \{1,2,10\} \tag{2.14b}
$$

The augmented Lagrangian method consists in removing complicated equality constraint (h) from an optimization problem, augmenting the objective function with a term that takes into account the Lagrange form of the particular constraint to be removed plus a penalty term. Each of these modifications of the objective function must be multiplied with a proper variable corresponding to the Lagrange multiplier (v) and a penalty multiplier (r) respectively. Once that the original optimization is redefined with the augmentation in the objective function, the idea is solving it iteratively updating the multipliers until eventually the results of the modified problem are equal to the original one. This condition is verified when the updated Lagrange multipliers converge to a stationary. In practice, the update formula from equation (2.15) can be used in two-stage programming (Kall & Wallace, 1997), while the penalty term can be increased slowly or kept fixed at a given value (Luenberger & Ye, 2008).

$$v^{n+1} = v^n + rh \tag{2.15}$$

In the two-stage representation of the HDS, we are interested in removing the nonanticipativity constraint since is the only one that links all the scenarios in the first stage. To do this, first of all we need to define the aggregated variable from equation (2.16) that can be understood as the most representative decision variable in the first stage for all the scenarios.

$$\bar{F}_{m,0} \coloneqq \frac{\sum_{i=1}^{N_{Sc}} \Pr(i)\, F_{m,0}^{i}}{\sum_{i=1}^{N_{Sc}} \Pr(i)}, \quad m = \{1,2,10\} \tag{2.16}$$

With the aggregate variable we can define the modified cost function augmented with the nonanticipativity constraint that is reformulated using the definition of $\bar{F}_{m,0}$.

$$\min_{F_1,F_2,F_{10}} \left\{ \begin{array}{c} \int_{t_0}^{t_1} (C_{H4}X_{H4}F_{1,0}^{j} + C_{H3}X_{H3}F_{2,0}^{j})dt + \\ \int_{t_1}^{t_f} \left(C_{H4}X_{H4}F_{1,1}^{j} + C_{H3}X_{H3}F_{2,1}^{j}\right)dt + \\ \int_{t_0}^{t_1} \sum_{m=\{1,2,10\}} \mu_m\left(F_{m,0}^{j} - \bar{F}_{m,0}\right) + r_m\left(F_{m,0}^{j} - \bar{F}_{m,0}\right)^2 dt \end{array} \right\} \quad (2.17)$$

Even though we have removed the hard nonanticipativity constraint from equation (2.14), there still remains a bonding between all the scenarios in the first stage but now in the augmented objective function. Therefore, we cannot split the large problem into smaller ones for each scenario yet. Nevertheless, we can note that if an estimator of the aggregated variable ($\bar{F}_{m,0}^{E}$) is known before the optimization is solved, then it is possible to solve equation (2.14) for each scenario *j* because there is not linking constraints among them as we can see from equation (2.18).

The estimator $\bar{F}_{m,0}^{E}$ can be computed before the optimization using the iterative procedure of the augmented Lagrangean method, replacing the solution of each scenario obtained in the previous iteration in equation (2.16).

Therefore, the complete two – stage optimization for the HDS can be solved in the following iterative way:

Step 0: Set the iteration counter in zero, $n = 0$. Choose $\mu_m^n, r_m^n, \bar{F}_{m,0}^{E\ n}, m = \{1,2,10\}$. The first guess can be obtained from the solution of the deterministic optimization for each scenario.

Step 1: $n = n + 1$

Step 2: Solve the dynamic optimization problem from equation (2.17) $\forall j = 1 \ldots N_{Sc}$ using a sequential approach and a control vector parameterization.

Step 3: Update the estimator $\bar{F}_{m,0}^{E}{}^{n}, m = \{1,2,10\}$, using equation (2.16)

Step 4: if $\left\|\bar{F}_{m,0}^{E}{}^{n} - F_{m,0}^{j}\right\|_{p} < \epsilon,\ \forall j = 1 \dots N_{Sc}$ stop, else go to step 5

Step 5: Update Lagrange multiplier μ_m^n and penalty factor r_m^n according to the Augmented Lagrangian method from equation (2.15) and go to Step 1.

$$
\begin{aligned}
&\min_{F_1,F_2,F_{10}} \left\{ \begin{array}{c} \int_{t_0}^{t_1} (C_{H4}X_{H4}F_{1,0}^{j} + C_{H3}X_{H3}F_{2,0}^{j})dt + \\ \int_{t_1}^{t_f} \left(C_{H4}X_{H4}F_{1,1}^{j} + C_{H3}X_{H3}F_{2,1}^{j}\right)dt + \\ \int_{t_0}^{t_1} \sum_{m=\{1,2,10\}} \mu_m\left(F_{m,0}^{j} - \bar{F}_{m,0}^{E}\right) + r_m\left(F_{m,0}^{j} - \bar{F}_{m,0}^{E}\right)^2 dt \end{array} \right\} \\
&s.t.: \\
&F_{1,k}^{j}X_1 + F_{2,k}^{j}X_2^{j} + F_3^{j}\xi_2^{j} = F_5^{j}X_5^{j} \\
&F_{1,k}^{j} + F_{2,k}^{j} + F_3^{j} = F_5^{j} \\
&0 = F_5^{j} - F_{10,k}^{j} - F_X^{H2\,j} \\
&\frac{VP}{ZRT}\left[\frac{dX_{H2}^{j}}{dt}\right] = F_5^{j}X_5^{j} - F_X{}^{j} - F_{10,k}^{j}X_{H2}^{j}, \qquad X_{H2}^{j}(t_0) = X_{H2_0} \\
&\tau\frac{dF_X^{j}}{dt} + F_X^{j} = F_{HC}\xi_1^{j}, \qquad F_X^{j}(t_0) = F_{X_0} \\
&t \in [t_0, t_f] \\
&F_3^{L} \le F_3^{j} \le F_3^{U} \\
&X_5^{L} \le X_5^{j} \\
&F_1^{L} \le F_{1,k}^{j} \le F_1^{U} \\
&F_2^{L} \le F_{2,k}^{j} \le F_2^{U} \\
&k = \begin{Bmatrix} 0, & t \in [t_0, t_1) \\ 1, & t \in [t_1, t_f] \end{Bmatrix}
\end{aligned}
\tag{2.18}
$$

As a result of solving the algorithm presented, there will be available a single optimal trajectory for the decision variables in the first stage according to the control parameterization chosen. For the second stage, there will be as many trajectories of the decision variables as scenarios were considered, in a similar way as Figure 2.2 shows. If a receding horizon

policy is used, only the value of the decision variables computed at current time are going to be applied to the process.

2.3.5 Chance Constrained Formulation

In contrast with the two-stage formulation chance constrained optimization does not assume that there are diverse stages of knowledge of the random variable. The method allows obtaining decision variables that guarantee a certain degree of feasibility for all the period considered in the optimization, without the knowledge of the uncertain values. Taking this into account, the original problem of optimal hydrogen feeding can be reformulated as equation (2.19), where the probabilistic constrained output is represented in equation (2.19b).

The model can be included in a simulation step to solve the dynamic optimization using a sequential approach. This implies that the decision variables are updated in each iteration using some optimization technique (SQP, for example), and these values are used to run a dynamic simulation in order to calculate the states, the constraints and the objective function of the system, which are passed back to the optimization routine to update the decision variables again. Therefore, the chance constraints from equation (2.19b) must be computed in each optimization step, by solving the simulation of the system for a given value of the decision variables.

The main difficulty of the chance constrained optimization is linked to the evaluation of the probability of the constraints, as it was mentioned previously. Arellano – García and co-workers proposed the inverse mapping method (Harvey Arellano-Garcia & Wozny, 2009) for calculating the probability of the constraints through the probability distribution of the random variables, which are assumed known. The only assumption that the system must fulfill is the monotony condition of the chance constraints with respect to, at least, one random variable.

$$\min_{F_1,F_2,F_{10}} \int_{t_0}^{t_f} (C_{H4}X_{H4}F_1 + C_{H3}X_{H3}F_2)dt$$
$$s.t.:$$
$$F_1X_1 + F_2X_2 + F_3\xi_2 = F_5X_5$$
$$F_1 + F_2 + F_3 = F_5$$
$$0 = F_5 - F_{10} - F_X^{H2}$$
$$\frac{VP}{ZRT}\left[\frac{dX_{H2}}{dt}\right] = F_5X_5 - F_X - F_{10}X_{H2}, \qquad X_{H2}(t_0) = X_{H2_0} \tag{2.19a}$$
$$\tau\frac{dF_X}{dt} + F_X = F_{HC}\xi_1, \qquad F_X(t_0) = F_{X_0}$$
$$t \in [t_0, t_f]$$
$$F_1^L \le F_1 \le F_1^U$$
$$F_2^L \le F_2 \le F_2^U$$
$$F_3^L \le F_3 \le F_3^U$$
$$\xi_1, \xi_2 \in \Xi$$

$$\Pr\{X_5 \ge X_5^L\} = 1 - \Pr\{X_5^{LO} \ge X_5\} \ge \alpha_1$$
$$\Pr\{X_{H2} \ge X_{H2}^L\} = 1 - \Pr\{X_{H2}^{LO} \ge X_{H2}\} \ge \alpha_2 \tag{2.19b}$$

As it was mentioned, the inverse mapping method is based in estimating the value of the probability constrained output variables by using the PDF of the random variables, estimating an equivalent of the boundary of the chance constraints y^* in the domain of ξ, called ξ^*. The monotony condition that is required for this methodology implies that, keeping all the system constant, for each value of ξ, lets say $\xi = \xi^a$, there is an equivalent value of $y = y^a$ that can be obtained solving the model h and that increases or decreases its value monotonically with ξ. This indicates that the probability of occurrence of y^a must be the same than the probability of occurrence of ξ^a (See Figure 2.10).

Therefore, if we are interested in calculating the probability of $y = y^*$ then it is necessary to solve the inverse mapping of h in order to estimate the equivalent value ξ^* that is replaced in its PDF Ξ. Taking in to account now the chance inequality constraints, the probability of $y \ge y^*$ might be equivalent to $\Pr(\xi \lesseqgtr \xi^*)$, where the direction of the inequality in the

random variables will depend on the sign of the monotony. That is to say, if an increase in the value of y increases the equivalent value of ξ the monotony is positive and the sign of the chance constraint is preserved, meaning that $\Pr(y \geq y^*) = \Pr(\xi \geq \xi^*)$. On the other hand, if an increase in y produces a decrease in its equivalent ξ, then the monotony is negative and the sign of the chance constraint changes its direction, which is equivalent to $\Pr(y \geq y^*) = \Pr(\xi \leq \xi^*)$.

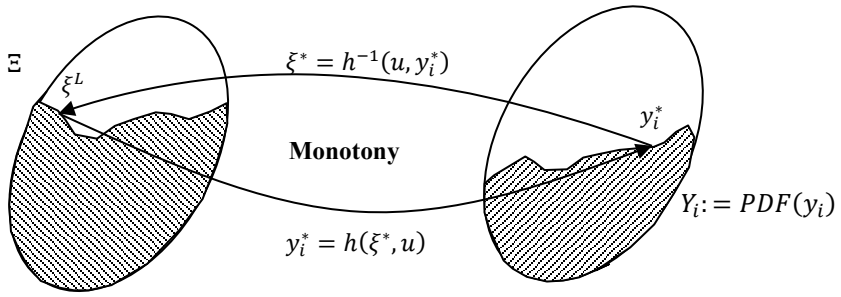

FIGURE 2.10: DIAGRAM OF THE INVERSE MAPPING METHODOLOGY

For the hydrodesulfuration plant the hydrogen purities X_5 and X_{H2} are the chance constrained outputs, as it is presented in equation (2.19b). For these variables the monotony can be checked in the following way: if all variables remain constant and the purity of the stream LP (ξ_2) increase its value, the hydrogen purities X_5 and X_{H2} will also increase its value. On the other hand, if ξ_2 decreases X_5 and X_{H2} will decrease too. Hence, the monotony of the chance constrained variables with respect to the random variable ξ_2 is positive, which implies that the inequality of the chance constraints preserves their direction. The problem then is to compute for which values of ξ_2 the chance constrained variables reach their limits. Consequently, the chance constraints can be translated to the fulfillment of a similar inequality but on the variable ξ_2 (equation (2.20)).

$$\begin{aligned}
&\Pr\{X_5 \geq X_5^L\} = 1 - \Pr\{\xi_2^{*X5} \geq \xi_2\} \geq \alpha_1 \\
&\Pr\{X_{H2} \geq X_{H2}^L\} = 1 - \Pr\{\xi_2^{*XH2} \geq \xi_2\} \geq \alpha_2 \\
&\Pr\{\xi_2^{*X5} \geq \xi_2\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\xi_2^{*X5}} \Xi(\xi_1, \xi_2) d\xi_2 \, d\xi_1 \\
&\Pr\{\xi_2^{*XH2} \geq \xi_2\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\xi_2^{*XH2}} \Xi(\xi_1, \xi_2) d\xi_2 \, d\xi_1
\end{aligned} \tag{2.20}$$

Here Ξ is the joint PDF of the uncertain variables.

From the integrals of equation (2.20) it can be noted that the probability of fulfillment of the constraints is calculated as a function of $\xi_2^{*X_5}$ and $\xi_2^{*X_{H2}}$. So, when computing the integrals, for every value of ξ_1 from $-\infty$ to ∞, the corresponding values of $\xi_2^{*X5}$and ξ_2^{*XH2} must be computed in order to evaluate the inner integral. These limits can be obtained for given values of the decision variables, replacing the value of the constrained variables in the model equations by its bounds and solving the equations h for the variable ξ_2. To do this Arellano García and Wozny (Harvey Arellano-Garcia & Wozny, 2009) propose to discretize the variables by using orthogonal collocation and then replacing the bounds of the constraints at the final time, solving a set of nonlinear equations for ξ^*. Regarding the good results that they present, their procedure only guarantees that the equality $y = y^*$ can be ensured at the end of the time horizon, meaning that the original problem is forced to be feasible only at $t = t_f$. In this work we have used a different method that considers the entire time horizon. Taking the advantage of the sequential approach employed to solve the dynamic optimization, the limits ξ^* can be computed minimizing the difference between the constrained variable and its bound from t_0 to t_f, solving the parameter estimation problem presented in equation (2.21), where the values of F_1, F_2 and F_{10} are known because are the ones given by the optimizer in each iteration of the optimization problem from equation (2.19).

$$\begin{aligned}&\min_{\xi_2^{*X5}} \|X_5 - X_5^L\| \\ &h(F_1, F_2, F_3, F_5, F_{10}, F_X, X_{H2}, X_5^L, \xi_1, \xi_2^{*X5}) = 0 \\ &\min_{\xi_2^{*XH2}} \|X_{H2} - X_{H2}^L\| \\ &h(F_1, F_2, F_3, F_5, F_{10}, F_X, X_{H2}^L, X_5, \xi_1, \xi_2^{*XH2}) = 0\end{aligned} \tag{2.21}$$

Being h the model of the system from equations (2.7), to (2.9) with the values of X_{H2} and X_5 replaced by their lower bounds. Once that the limits $\xi_2^{*X5}$and ξ_2^{*XH2} are calculated, the inner integral of the PDF in (2.20) can be evaluated and the procedure is repeated until the range of $\xi_1$is covered, obtaining in this way the value of the probability for a given values of the decision variables. For the computation of the integrals, a trapezoidal rule was used. Since the computation of the probability implies that ξ_1 can have infinite values, the whole range of this value used in equation (2.21) was its 99.5% confidence interval, that is to say: $\xi \in [\mu_1 - 3\sigma_1, \mu_1 + 3\sigma_1]$. A schematic of the sequential optimization method is represented in Figure 2.11

The following algorithm summarizes the proposed chance constrained method using the sequential approach for solving the stochastic dynamic optimization problem:

Step 0: Set the iteration counter in zero $n = 0$. Choose the initial guesses for decision variables F_1^n, F_2^n, F_{10}^n. Calculate the objective function and the constraints from equation (2.19).

Step 1: $n = n + 1$

Step 2: Estimate the new decision variables F_1^n, F_2^n, F_{10}^n , using an optimization procedure (for example SPQ).

Step 3: Calculate the cost function of equation (2.19) by dynamic simulation with the current decision variables.

Step 4: For each ξ_1 within its range, obtain the value of $\xi_2^{*XH2}, \xi_2^{*X5}$ solving (2.21).

Step 5: Calculate the probability of the chance constrained variables from equation (2.20)

Step 6: If the KKT conditions are fulfilled under a certain degree of tolerance, stop. Otherwise, go to step 2.

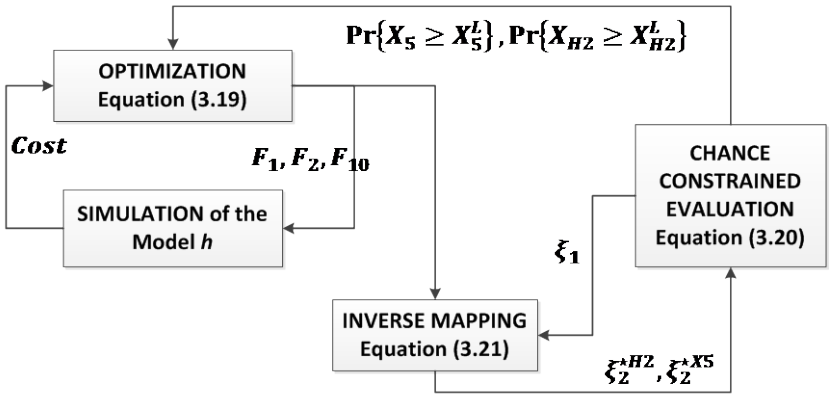

FIGURE 2.11: DIAGRAM OF THE SOLUTION METHOD OF CHANCE CONSTRAINED OPTIMIZATION

2.4 Results and Discussion

In order to test the two ways of coping with stochastic optimizations they were applied in the hydrodesulfuration problem of the example, assuming at t_0 a step change on the flow and quality of the hydrocarbon stream that feeds the reactor, simulating a change in the type of hydrocarbon to be treated. The load change can be described as: starting from a hydrocarbon flow of F_{HC}^0 with a sulfur quantity equivalent to ρ_{HC}^0, change the conditions in the feed of the system to F_{HC}^f and ξ_1 respectively. As before, the other stochastic variable ξ_2 will be the purity of LP feed: X_3. The values of the parameters used in the optimizations for both methods are the same ones

used in the deterministic optimization (Table 2.2). In terms of the available information during the optimization, only the value of F_{HC}^{f} is known, since in general this is a decision imposed by production requirements. On the other hand, ξ_1 and $\xi_2$are unknown at t_0 but it is possible to estimate their value after a certain period of time t_1 by means of laboratory analysis that can be implemented offline. Means (μ), variances (σ) and correlation coefficient (r) of the stochastic variables, as well as other additional parameters like the minimum degree of feasibility imposed in the optimization (α), are summarized in Table 2.3.

TABLE 2.3 ADDITIONAL VALUES FOR STOCHASTIC OPTIMIZATION

Parameter	Value	Parameter	Value
$\mu_1\left(\frac{kmol}{m^3}\right)$	12.6	α_1, α_2	0.9
μ_2	0.85	r	- 0.5
$\sigma_1\left(\frac{kmol}{m^3}\right)$	0.4	$t_1(h)$	1.5
σ_2	0.013		

The results obtained with each methodology, were tested using a Montecarlo simulation with 1e3 realizations of the random variables, simulated from the original PDF of Figure 2.7. This test has been proposed with the idea of confirming the level of feasibility of the outcomes obtained with the approximated methods implemented in the stochastic optimization. The solutions proposed by both methods and the results of the Montecarlo tests are presented below.

The single shooting methodology for both stochastic optimization methods tested, were implemented in EcosimPro 4.6.(Internacional, 2009). This is a simulation environment that uses DASSL as numerical integrator to solve the DAE system (Brenan et al., 1989). DASSL is a code for solving index zero and one system of differential/algebraic equations of the form:

$$h(\dot{x}, x, u, t) = 0, \qquad x(t_0) = t_0 \tag{2.22}$$

The basic idea for solving DAE systems using numerical ODE methods is to replace the derivative in equation (2.22) by a difference approximation, and then to solve the resulting (non)linear system for the solution at the current time t_{n+1} using Newton's method. For example, replacing the derivative in equation (2.22) by the first order backward difference, we obtain the following implicit Euler formula:

$$h\left(\frac{x_{n+1} - x_n}{\Delta t_{n+1}}, x_{n+1}, u, t_{n+1}\right) = 0, \tag{2.23}$$

Being Δt_{n+1} the step size of the integration. The algorithm that is implemented in DASSL is an extension of this idea, but instead of using always a first order approximation of the derivatives, it approximates the variation over time by a k^{th} order backwards differentiation formula, where k can vary from one to five. On every time step it chooses the order k and the step size, based in the behavior of the solution.

Regarding the optimization layer of the single shooting approach applied in EcosimPro, it was employed the Sequential Quadratic Algorithm (SQP) implemented in the NAG routines (Numerical Algorithms, 1998). The SQP method consists in solving the KKT conditions of the problem by using iteratively a quadratic approximation of the original nonlinear optimization problem of equation (2.24). Starting from an initial guess, the idea of the methodology is to build a second order approximation of the Lagrangean function and a linear representation of the constraints to solve an approximated quadratic problem. The outcome of the optimization procedure is used to find the direction where the next iterant will be, which is obtained after solving a line search problem. This procedure is then repeated in a sequential manner until no improvements in the objective function, or no changes in the decision variables are observed (Biegler Lorenz, 2010).

$$\begin{aligned} &\min_{u} f(u) \\ &s.t.: \\ &h(u) = 0 \\ &g(u) \leq 0 \end{aligned} \tag{2.24}$$

To solve equation the SQP problem it is necessary to estimate first and second order derivatives. The implementation of the NAG routines in EcosimPro estimates the gradient with a perturbation-based method applying the finite differences calculation, while the Hessian is obtained from a Quasi-Newton approximation.

2.4.1 Results from Two – Stage Optimization

As it was mentioned, to cope with the numerical difficulties that suppose the solution of the two-stage implementation when the uncertainty is described with a continuous PDF, the scenarios technique has been used to transform the original problem into a finite-dimensional one. Figure 2.9 represents the scenario representation of the PDF that describes ξ_1 and ξ_2. It was built discretizing the confidence interval of the 99.95% of each random variable in five values, giving a total number of 25 scenarios, which is translated in the same number of individual optimizations to be solved in every iteration of the Lagrangian aggregation methodology. In order to reduce this number, all the scenarios with less than 1% of probability of occurrence were neglected. This action decreased the number of random variable combinations until 10, explaining the 99.04% of probability.

For each scenario, the dynamic optimization problem was solved using the single shooting method. The discretization of the decision variables was performed by using control vector parameterization with eight possible values for each stage. Therefore, there are 48 decision variables per scenario, giving a total number of decision variables in the entire stochastic optimization problem equal to 480. The optimization took about 9h. Results are given in Figure 2.12.

Notice that this high computation time prevent the method from being applied on-line, but the results are still applicable in open loop if changes in operation mode are planned in advance and enough historical data are available to estimate the probability distribution of the uncertain variables, which is the case of an HDS operation. Computing the solution off-line provides a pattern to bring the process in the best possible way to a new optimal steady state. Then, when the feed change takes place, the stage-0 solution is used to guarantee feasibility even if there is no information about the value of the random variable and then, provided the uncertain variables can be estimated from measurements, stage-1 policies gives optimality. This way of acting is the basic idea of the multistage optimization, since the assumption of several stages of knowledge implies that at the beginning, when we do not know about the value of the uncertain variables, we have to take a decision that must present a compromise between feasibility and optimality. And later, when we are able to estimate or have some feedback about the value of random variable, then it is necessary to implement corrective actions that drive the process to the optimality.

Figure 2.12 shows the trajectories of the decision variables for the control parameterization chosen and the constrained purities over time. It can be noted that for the first stage (until $t_1 = 1.5h$), the decision variables have only one value because of the nonanticipativity requirement that now can be understood as a constraint imposed over the complete trajectory described for each decision variable during the entire first stage, hence there are 280 nonanticipativity constraints to augment the objective function. Once that the unknown variables are estimated, after t_1, the system decides which are the optimal trajectories according to the value of the random parameters, that is why there are ten different control trajectories for times greater than 1.5h, which depends on the value of the combination of ξ_1 and ξ_2 for every scenario considered, regarding to the corrective actions that the two-stage optimization presents.

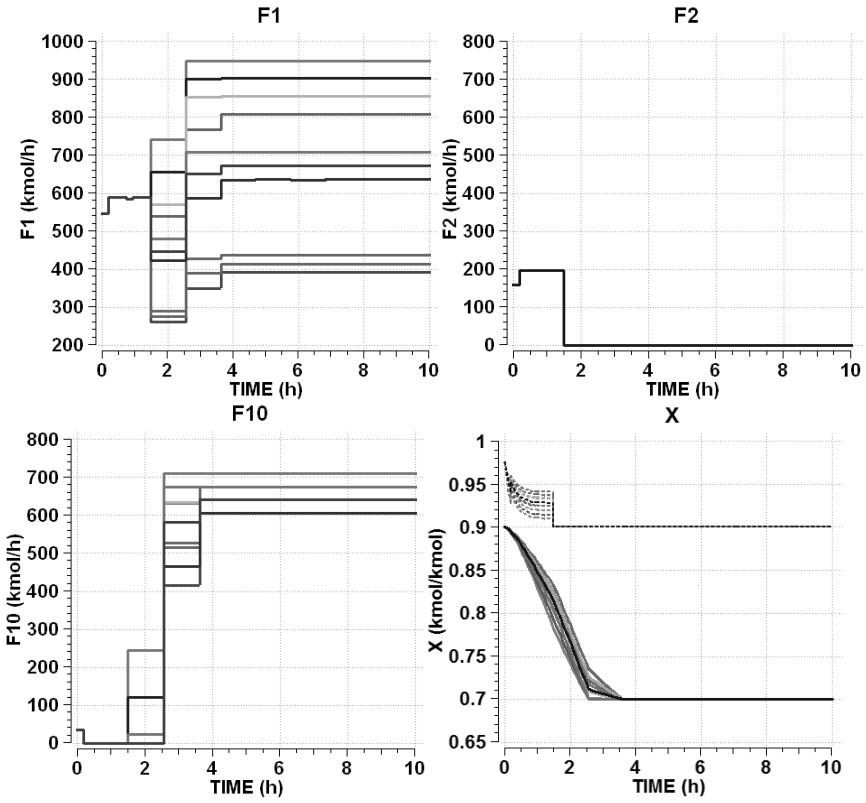

FIGURE 2.12: RESULTS OF THE 2-STAGE OPTIMIZATION FOR EACH SCENARIO CONSIDERED

The effect that these trajectories have in the constrained variables is also represented in Figure 2.12. It can be noted that during the first stage, the purities shows as many trajectories as scenarios are, since the process dependent variables are function of the value of the random and the decision variables. In each of the trajectories presented, the purities decrease their values trying to reach the constraint, giving the idea of the optimality that the two-stage methodology tries to obtain, however, the uncertainty in the value of the random variable forces the system to stay inside the feasible region for all the scenarios considered, taking into account the feasibility for all the possible values that the random variable may have. Again, if now we focus our attention to the period when the

uncertain parameters are estimated (after 1.5h), the different trajectories calculated for the decision variables sends the constraints to their lower limit, which again can be understood from the point of view of the corrections that the method must implement in order to look for the optimality once that the random variables are known. These corrective actions can be observed also comparing these results with the ones from deterministic optimization (Figure 2.6). In order to ensure an optimal use of the hydrogen, classic optimization suggests not to use hydrogen that comes from H3 (F_2), because is more expensive in terms of the purity delivered. Unlike deterministic optimization, the results from two – stage optimization suggests opening F_2 in order to maintain the feasibility, until the value of the random variable is estimated. At this moment, this value is set to zero searching for optimality.

In the literature of multistage optimization, some indexes are defined that attempts to look for justifications of the stochastic optimization, in terms of comparing the efforts that are necessary to make to solve a very complicated problem, with the possibility of invest the time improving the estimation step of the uncertain variable. That is to say, evaluate the use of stochastic optimization in terms of whether it is necessary its application or it is better to improve the efforts required to estimate the uncertain variables in real time. One of the indexes used for this purpose is the *Estimated Value of Perfect Information* ($EVPI$) (Birge & Louveaux, 1997). This index presents the implicit assumption that it is possible, in some way, measuring the uncertain variable during the entire optimization period. The index compares the value in objective function that would be obtained if for each scenario the uncertain variable is known perfectly, with the one obtained in two-stage programming. Since the perfect situation supposes that the uncertain variable is available, this index makes sense in process optimization when it is possible to measure directly or indirectly the random variables by means of reliable online sensors. It is defined as the difference between the cost function under perfect information ($COST_{PI}$),

calculated as the weighted sum of the objective function for each scenario if the random variable is known in the first stage, and the objective function obtained in the two stage optimization ($COST_{TS}$). If this index is defined relative ($EPVI_R$) to the $COST_{PI}$ as equation (2.25) shows, it represents the percentage of the costs that must be spent due to not having complete knowledge of the system, i.e. how much can be saved if hydrogen and hydrocarbon analyzers are used in order to measure these variables online, instead of using the results from this optimization. For this problem the $EVPI_R$ is 6.2%.

$$EVPI_R = \left(\frac{COST_{TS} - COST_{PI}}{COST_{PI}}\right) 100$$
$$COST_{PI} = \frac{\sum_{i=1}^{N_{Sc}} \Pr(i)\, COST(i)}{\sum_{i=1}^{N_{Sc}} \Pr(i)} \quad (2.25)$$

2.4.2 Results from Chance Constrained Optimization

The iterative procedure from Figure 2.11 has been used to solve the chance constrained formulation for the HDS example. In order to calculate the limits of integration in the calculus of probability it has been implemented a nested dynamic optimization using the single shooting approach. The total optimization took 8.5h to reach the optimal point. Notice that the same discussion as in two-stage optimization applies regarding the use of the solution in practice in real time. Figure 2.13 summarizes the value of the optimal decision variables for a probability of feasibility for X_{H2} and X_5 equal to 90%.

The evolution of the optimal trajectories obtained with the chance constrained approach, are quite similar to the results obtained with the deterministic optimization, since there is a single trajectory for each decision variable, but with the difference that probabilistic criteria has been taken into account in order to ensure a fixed degree of feasibility.

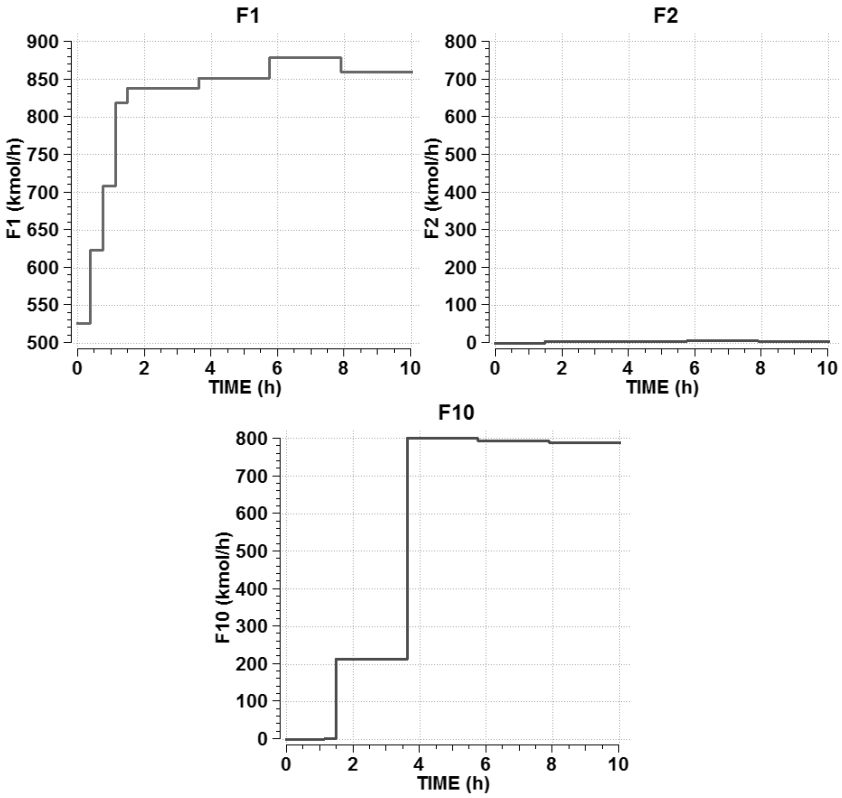

FIGURE 2.13: VALUE OF THE DECISION VARIABLES FOR CHANCE CONSTRAINED OPTIMIZATION

Regarding the actions that the decision variables produce in the system, it can be noted that there are no corrective actions in terms of feasibility and optimality. This is reflected in the value of F_2 which is always close to zero, but is still a positive number. Because F_2 presents a less convenient ratio between purity and production cost (as it was pointed out in the deterministic optimization outcomes), the gap between zero and the optimal trajectory of this variable can be interpreted as the compromise among optimality and feasibility, that is to say, the additional price that must be paid, in order to ensure a feasible operation with a given value of probability. About the other decision variables, comparing with Figure 2.6 or 2.12, it can be said that F_{10} has increased its value, because it is

necessary to take in to account more scenarios where the system must be feasible, meaning that the most of the probability distribution of the constrained variable X_{H2} must be inside the feasible region which can be obtained only if there is an increase in the flow purged from the reactor. The growth on the purged flow compared with the deterministic case, implies that more hydrogen must enter in the reactor in order to keep the pressure constant and since the hydrogen concentration in the low purity stream has a random behavior, the optimization increases the consumption of high purity hydrogen (F_1) because of the optimality of working with this type of source, as it was pointed out previously, but keeping the shape of the trajectory, implying also that the method tries to bring the limit of the 90% of the distribution function of the constrained purities close to their lower bound. As a summary we can argue that according to the formulation of this method of optimization, the increments in these variables can be explained in order to ensure feasibility 90% of the times. This statement has been tested with a Montecarlo simulation, in order to confirm the analysis done about the trajectories of the state constrained variables. The results of the trajectories as well as their cumulative histograms at are represented in Figure 2.14.

The cumulative histogram of the constrained variables in the stochastic simulations shows how the methodology moves the center of mass of their probability distribution function into a feasible region in order to ensure the accomplishment of the probabilistic constraints, placing the limit of the feasibility imposed in the tail representing the 10% of the probability of occurrence, meaning that solving the parameter estimation problem to estimate the value of the limit of the probability integral ξ^*, we have achieved the goal to propose a policy that is feasible in 90% of the cases, during the whole transient of the experiment considered in the optimization.

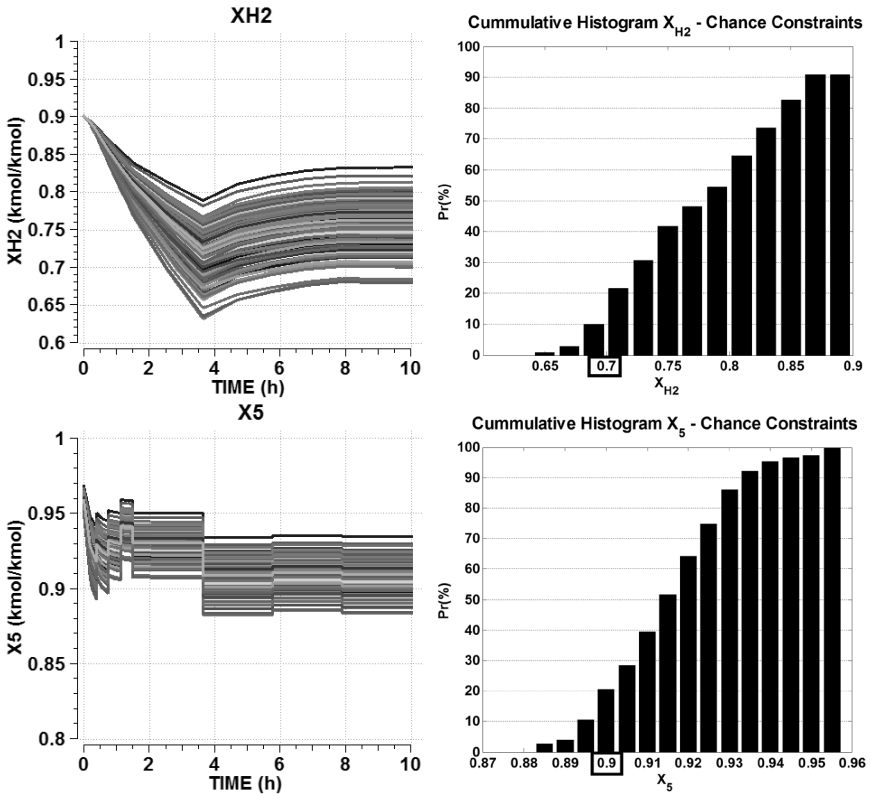

FIGURE 2.14: EVOLUTION OF THE PURITIES USING CHANCE CONSTRAINED FORMULATION

The center of mass of a probability distribution can be viewed as the case that presents a probability of occurrence equal to 50%, which in X_{H2} is equal to 0.77 and in X_5 is 0.913. The gap between the lower bound of the constrained purities and the center of the mass of their probability distribution functions gives information about the optimality of the chance results, and in particular about the tradeoff that must be considered to cope with infeasibilities. It tell us how much we need to move the lower limit of the constrained variable in order to ensure a feasible operation in the 90% of the cases when applying an optimal policy obtained from an optimization, i.e. it is the security back-off that we need to take into account in order to overcome the unknown part of the process in

optimization, as Figure 2.15 schematizes. As it is expected, we can note that an increase in the percentage of feasibility imposed in the probabilistic optimization will be translated in a loss of optimality with respect to the deterministic one, however, this is the price that we must pay in order to operate in a safe way. The confidence level is a designer criteria and it depends on the type of constraints. In particular in this application, this is a compromise among the degree of degradation of the catalyzer, the capacity of the compressors and the economic losses of the overall process.

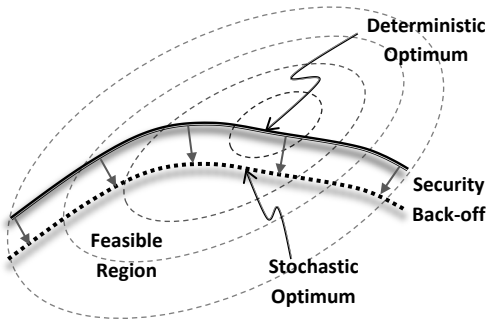

FIGURE 2.15: SCHEMATIC REPRESENTATION OF THE EFFECT OF THE BACK-OFF IN THE OPTIMALITY

From the evolution in time of the purities of Figure 2.14, it can also be noted the change in the center of the mass of the distributions of the trajectories. Moreover, the shape of these trajectories give some insights about the optimality that the method tries to reach, noting that, as in the deterministic case, the decision variables tries to decrease their value in order to consume less fresh hydrogen, moving the process into a more economic point, which again can be understood as the mean of the probability distribution of the constrained variables.

2.4.3 Open Loop Application

Since both optimizations approaches, two-stage and chance constraints, have been solved in times that makes the online application impossible, it is

necessary to study the open loop implementation of the previous solutions. To do this, we have tested the previous outcomes using a Montecarlo simulation with 1e3 realizations of the random variables, simulated from the original continuous PDF.

Regarding the optimal trajectories from chance constrained optimization we already show their performance with stochastic simulations. As a result of the simulations it can be said that the direct open loop implementation of the trajectories increases the degree of feasibility to 90%. Bearing in mind the security that must be considered, the decisions that the method takes to solve a problem are equivalent to increasing the lower bounds of the inequality constraints.

On the other hand, the results obtained from the two-stage optimization cannot be applied directly since it was solved using the scenarios approach, because the trajectories for F_1, F_2 and F_{10} in the second stage are only available for the discretized scenarios considered. Owing to the original PDF has a continuous nature, it is necessary to find a way to apply these results for the values of the uncertain parameters not considered previously. In this thesis we have evaluated the use of three strategies:

Nearest Neighborhood: With this method, the idea is to use the closest value of the random variable compared with the scenarios that had been considered in the optimization step. Therefore, the random variables from the continuous PDF will be compared with the discretized ones and it will be applied the decision variable calculated for the nearest scenario used in optimization, in order to take into account the optimality.

Oversize the System: The decision applied in simulation will be the one calculated for the nearest scenario that oversize the system, i.e.: for the random variables simulated, the decision applied will be the one calculated for the nearest scenario with greater ratio hydrogen – hydrocarbon and lower hydrogen purity in the recycled stream, in order to ensure feasibility.

Interpolation: The decision variable applied in the simulation will be interpolated between the results obtained for the scenarios. Linear and Akima interpolation methods were used, the last one being a continuously differentiable sub-spline interpolation with piecewise third order polynomials (Akima, 1970).

Figure 2.16 shows a summary of the methods for open loop implementation of the two-stage optimization, schematized over the surface of the optimal trajectories of the decision variables that would be obtained if the two- stage optimization was solved with the original PDF. In the figure, ξ_i^D is the discrete value of the random variables for the scenarios, ξ_i^S is the value of the random variable simulated with the original continuous PDF. The grid represents the optimal function, the white dots indicate the trajectory of the decision variables obtained for each scenario used in optimization and the star marks the location of the true optimal trajectory of the decision variable for ξ_i^S. The black dot denotes the trajectory of the decision variables that finally was applied in the simulation.

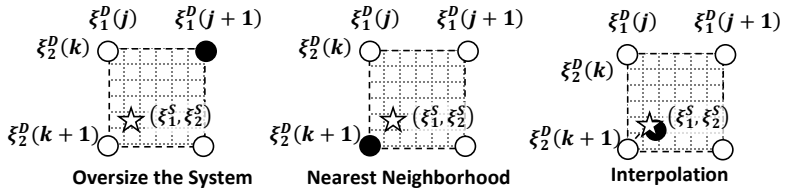

FIGURE 2.16: REPRESENTATION OF THE OPEN-LOOP IMPLEMENTATION METHODS OF THE TWO-STAGE OPTIMIZATION

Because of the decision variables implemented in the process simulation can be different to the ones calculated in the optimization, the $EVPI_R$ may change with respect to the one previously presented. Therefore, the value from equation (2.25) does not represent a fair evaluation of the open loop implementation for the two-stage solution over the alternative of online

measurements. To consider this change, an analogous index is defined: the relative estimated value of perfect information in simulation ($EVPI_R^S$). This index is calculated as the average of the difference between the cost function under perfect information and the real cost obtained applying the strategies of implementation, as equation (2.26) shows. Here, N_{McS} is the number of Montecarlo simulations performed.

$$EVPI_R^S = \frac{\sum_{i=1}^{N_{McS}} \left(\frac{COST_{TS}(i) - COST_{PI}(i)}{COST_{PI}(i)} \right)}{N_{McS}} 100 \tag{2.26}$$

Table 2.4, summarizes the results of the Montecarlo simulations in terms of mean and standard deviation of the hydrogen purities for the cases studied. In each row, the results obtained applying open loop solutions from the methods tested are presented. In columns: $EVPI_R^S$, mean, variance and $\%\, Inf$ for the chance constrained variables. Notice that the index named $\%\, Inf$ corresponds to the percentage of the times when the purities were lower than their bounds. Regarding optimality the last two columns of Table 2.4 presents the mean and the standard deviation of the costs. It can be noted that the results of two-stage optimization are separated in: "Nearest scenario", "Overestimate", "Linear Interpolation" and "Akima Interpolation", taking into account the different ways to apply the discrete outcomes in the second stage of this technique, as mentioned above.

About the infeasibility percentage for the two constrained variables (columns 4 and 7), it can be noted that applying the trajectories obtained with deterministic optimization, about 50% of the times the operation is out of the bounds, in concordance with Figure 2.8. This is because the distributions of the hydrogen purities are centered in the lower bound of the constraints, due to the fact that the optimization had been performed using the mean value of the random variables.

TABLE 2.4 SUMMARY OF THE MONTECARLO SIMULATIONS

Optimization Method	$EVPI_R^S$ (%)	X5			XH2			COST (k€)	
		μ	σ 10^3	% Inf.	μ	σ 10^2	% Inf.	μ	σ 10^2
Deterministic	N/A	0.900	8.7	44.1	0.700	3.2	49.4	0.548	N/A
Chance Constrained	N/A	0.913	8.4	12.6	0.770	2.7	9.9	0.658	N/A
Two – Stage: Nearest Scenario	6.2	0.900	4.7	44.2	0.700	1.5	48.4	0.567	10.3
Two – Stage: Overestimate	22.8	0.908	4.4	0	0.728	1.4	0.8	0.655	9.2
Two – Stage: Linear Interpolation	7.1	0.900	2.1	0	0.700	0.7	1	0.571	9.5
Two – Stage: Akima Interpolation	7.8	0.900	2.1	0	0.701	0.7	1	0.575	9.5
Worst Scenario	-	0.93	7.6	0	0.8	2.5	0.8	0.730	N/A

The infeasibilities in chance constrained optimization are closer to 10%, which was already presented in the Figure 2.14. This result is expected considering that a confidence level of 90% has been imposed in the resolution of the optimization problem. The differences between these percentages and the values imposed can be due to the numerical approximation of the probability calculus. In addition, it can be noted that the mean of the hydrogen purities has been moved above their lower bound, and the standard deviations are similar with respect to deterministic optimization. The previous facts indicate that the optimizer search for some decision variables that allocate the infeasibilities in the tail of their distribution, confirming earlier discussions about the optimal trajectories. Regarding optimality, it can be noted from the last columns that the cost of

using a policy that ensures feasibility without knowing the value of the uncertain variable implies an amount of 20% with respect to the trajectories obtained with deterministic optimization, which can be understood as the increase in the prices because of growing the lower bound of the purities.

For the two – stage formulation, Table 2.4 shows differences in the degree of feasibility and the optimality depending on the strategy used to implement the results.

The percentage of infeasibilities and the mean of the purities in the nearest scenario strategy are similar to the ones registered in deterministic optimization. However, the standard deviation has been reduced in a half. This situation can be explained understanding that using this strategy is equivalent to split the original PDF in ten pieces (the number of scenarios) and then solving a deterministic optimization. Each of them will be solved with the mean of the random variables inside every piece, hence it's expectable that in about a 50% of the times the uncertain variables will underestimate the system and the other half of the times will be overestimated (Harvey Arellano-Garcia & Wozny, 2009). In addition, the decrease in the standard deviation is explained due to the range of the chance variable is bounded in each scenario with respect to the complete PDF. About optimality, it can be noted that the average of the cost function is the same than from deterministic optimization, which leads to the $EVPI_R^S$ equal to the $EVPI_R$. This is quite expected because of the analogy of deterministic optimization over each scenario using the mean value of stochastic variables.

Using the method that overestimates the system, feasibility is ensured for almost all the possible values of the uncertain variables. As the system is always overrated, the mean of purity constraints must be inactive. Both situations can be observed in Table 2.4. This strategy also can be explained by splitting the original PDF in pieces and then solving several deterministic

optimizations. However, in this case rather to use the mean of the uncertain variables, the worst case inside each piece is applied. Therefore, is analogous to the worst case optimization, but with the advantage that the unknown variable is bounded by each scenario. This overestimation of the system leads to an operation which is, in average, more expensive than the expected by the stochastic optimization. Therefore, the $EVPI_R^S$ is more than 5 times bigger than the observed in optimization.

Regarding the results obtained interpolating the discretized optimal solution, Table 2.4 shows that the feasibility percentage, mean and standard deviation are the same for the two methods tested: linear and Akima interpolation. It can be noted that the feasibility percentages, being very low, are not as good as the results from worst case optimization. However, in terms of dispersion of the constrained variables, these values are lower than all the other strategies testes, which mean that a most homogeneous operation policy can be expected. Because of this, the mean of the purities are equal to their lower bounds. About the optimality of applying the interpolated solutions, it can be said that the cost function is the lowest for the strategies able to ensure an appropriate level of feasibility. About the $EVPI_R^S$, it can be said that is bigger than the $EVPI_R$ in 15%. Comparing this index and the percentages of feasibility with respect to the other strategies of application it can be argued that this is the optimal way to apply the results of the two – stage optimization for this problem.

2.5 Conclusions

From the results presented, it can be concluded that it is possible to solve a two-stage optimization for an optimal process operation using the scenario aggregation method and the single shooting technique. Nevertheless, due to the large computation times, in order to be an attractive way to handle uncertainties in processes, it is necessary to review the dynamic optimization procedure and the optimization resources. This is because in the current implementation, most of the time employed by the SQP solver

used in this work is spent in the estimation of the gradients using finite differences. Another way to reduce the solution time is parallel computation, so that each scenario problem can be solved separately in a processor to solve the entire problem.

The results obtained using the scenario approach can be generalized for continuous uncertain variables using the interpolation method proposed here. This idea allows solving the two – stage optimization, simplifying the calculus in order to be applied in a continuous process. The way in which implementation of the second stage decision is made proved to be more important than expected, according to the test performed.

Regarding the chance constrained optimization, it can be concluded that by solving a parameter estimation problem, the inverse mapping method can be applied in a continuous way in order to calculate the probability for constrained variables, and to estimate the security back-off in the bounds of the inequality constraints. Anyhow, again is necessary to review the methods of dynamic optimization and the resources used, to implement these ideas from the point of view of process optimization. As a new point of view, the chance constrained optimization in open loop, might be useful for a new close loop procedure with deterministic optimization, provided is available an adequate mapping between the feasibility percentage required for the process and the corresponding security back-off to be implemented in the inequality constraints.

About the comparison of both methods tested, the ratio between the percentage of feasibility and the objective function is better for the two – stage optimization with the interpolation procedure. To conclude what formulation works better for this case, it is necessary to complement this analysis. Because the two – stage optimization is valid only if it is possible to estimate (using some method) the random variables before the second stage, unlike chance constrained formulation that ensures a level of feasibility with no need of this estimation. In the problem of hydrogen

consumption, the random variables could be estimated before the second stage occurs by combining state estimation methods and some laboratory analysis. Therefore, the two – stage formulation would work better in absence of important uncertainties related to the measurement procedure.

The work developed here can be viewed as a set point updater for the hydrogen flow controllers. If improvements in the computation time were available, it would be interesting to close the loop reformulating this application as a stochastic model predictive controller. But, even in its current formulation, it can contribute to easy the implementation path of advanced optimal decision making tools in industry, and the integration with other approaches to the optimization of complete hydrogen distribution network for a refinery (Sarabia et al., 2009).

2.6 References of the Chapter

Akima, H. (1970). A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures. J. ACM, 17, 589-602.

Arellano-Garcia, H., Henrion, R., Li, P., Möller, A., Römisch, W., Wendt, M., & Wozny, G. (1998). A model for the online optimization of integrated distillation columns under stochastic constraints. In S. D. D. FORSCHUNGSGEMEINSCHAFT (Ed.): ECHTZEITOPTIMIERUNG GROSSER SYSTEME.

Arellano-Garcia, H., Martini, W., Wendt, M., Li, P., & Wozny, G. (2003). Robust optimization of a reactive semibatch distillation process under uncertainty. In A. Kraslawski & I. Turunen (Eds.), 13th European Symposium on Computer Aided Process Engineering (ESCAPE-13) (Vol. 14, pp. 551-556). Lappeenranta, Finland.

Arellano-Garcia, H., & Wozny, G. (2009). Chance constrained optimization of process systems under uncertainty: I. Strict monotonicity. Computers & Chemical Engineering, 33, 1568-1583.

Barz, T., Arellano-Garcia, H., & Wozny, G. (2008). Implementation of Optimal Decisions in the Presence of Uncertainty. In 17th World Congress The International Federation of Automatic Control (pp. 11931 - 11936). Seoul, Korea.

Bellos, G. D., & Papayannakos, N. G. (2003). The use of a three phase microreactor to investigate HDS kinetics. Catalysis Today, 79-80, 349-355.

Biegler Lorenz, T. (2010). Nonlinear programming : concepts, algorithms, and applications to chemical processes: Society for Industrial and Applied Mathematics.

Birge, J. R., & Louveaux, F. (1997). Introduction to stochastic programming. New York: Springer.

Brenan, K. E., Campbell, S. L., & Petzold, L. R. (1989). Numerical solution of initial-value problems in differential-algebraic equations. New York: North-Holland.

Charnes, A., & Cooper, W. W. (1959). CHANCE-CONSTRAINED PROGRAMMING. Management Science, 6, 73-79.

Dantzig, G. B. (1955). LINEAR PROGRAMMING UNDER UNCERTAINTY. Management Science, 1, 197-206.

Dupacova, J. (1995). MULTISTAGE STOCHASTIC PROGRAMS - THE STATE-OF-THE-ART AND SELECTED BIBLIOGRAPHY. Kybernetika, 31, 151-174.

Dupačová, J., Consigli, G., & Wallace, S. W. (2000). Scenarios for Multistage Stochastic Programs. Annals of Operations Research, 100, 25-53.

Gómez, E., De Prada, C., Sarabia, D., Méndez, C. A., Cristea, S., Sola, J. M., & Unzueta, E. (2008). Grey-box modelling of an industrial hydrodesulphurization process. In B. Braunschweig & X. Joulia (Eds.), 18th European Symposium on Computer Aided Process Engineering – ESCAPE 18. Lyon, France: Elsevier.

Internacional, E. A. (2009). EcosimoPro Modelling and Simulation Software - EL Modelling Language Version 4.6. In: Empresarios Agrupados Internacional.

Kall, P., & Wallace, S. W. (1997). Stochastic programming. Chichester; New York: Wiley.

Kelly, J. D., & Mann, J. L. (2005). Improve yield accounting by including density measurement explicitly. Hydrocarbon Processing, 93 - 95.

Li, P., Arellano-Garcia, H., & Wozny, G. (2008). Chance constrained programming approach to process optimization under uncertainty. Computers & Chemical Engineering, 32, 25-45.

Li, P., Wendt, M., Arellano-Garcia, H., & Wozny, G. (2002). Optimal operation of distillation processes under uncertain inflows accumulated in a feed tank. AIChE Journal, 48, 1198-1211.

Li, P., Wendt, M., Arellano-Garcia, H., & Wozny, G. (2003). PROCESS OPTIMIZATION AND CONTROL UNDER CHANCE CONSTRAINTS. In ADCHEM. Hong kong.

Loughlin, D. H., & Ranjithan, S. R. (2001). Chance-Constrained Optimization Using Genetic Algorithms: An Application in Air Quality Management. In D. Phelps & G. Sehlke (Eds.), (40569 ed., Vol. 111, pp. 50-50). Orlando, Florida, USA: ASCE.

Luenberger, D. G., & Ye, Y. (2008). Linear and nonlinear programming. New York, NY: Springer.

Mayne, D. Q., & Polak, E. (1976). FEASIBLE DIRECTIONS ALGORITHMS FOR OPTIMIZATION PROBLEMS WITH EQUALITY AND INEQUALITY CONSTRAINTS. Mathematical Programming, 11, 67-80.

Miller, B. L., & Wagner, H. M. (1965). CHANCE CONSTRAINED PROGRAMMING WITH JOINT CONSTRAINTS. Operations Research, 13, 930-&.

Nemirovski, A., & Shapiro, A. (2006). Convex Approximations of Chance Constrained Programs. SIAM Journal on Optimization, 17, 969-996.

Numerical Algorithms, G. (1998). NAG C library manual : mark 5. Oxford, U.K.: Numerical Algorithms Group Ltd.

Poojari, C. A., & Varghese, B. (2008). Genetic Algorithm based technique for solving Chance Constrained Problems. European Journal of Operational Research, 185, 1128-1154.

Prékopa, A. (1970). On Probabilistic Constrained Programming. In Princeton Symposium on Mathematical Programming (pp. 113 - 138). Princeton, NJ: Princeton University Press.

Prékopa, A. (1995). Stochastic programming. Dordrecht ; Boston: Kluwer Academic Publishers.

Prékopa, A. (2003). Probabilistic Programming. In A. Ruszczynski & A. Shapiro (Eds.), Handbooks in Operations Research and Management Science (Vol. Volume 10, pp. 267-351): Elsevier.

Rockafellar, R. T. (2001). Optimization Under Uncertainty. In. Washington: University of Washington.

Ruszczynski, A., & Shapiro, A. (2003). Stochastic Programming Models. In A. Ruszczynski & A. Shapiro (Eds.), Handbooks in Operations Research and Management Science (Vol. Volume 10, pp. 1-64): Elsevier.

Sahinidis, N. V. (2004). Optimization under uncertainty: state-of-the-art and opportunities. Computers & Chemical Engineering, 28, 971-983.

Sarabia, D., Cristea, S., Gómez, E., Gutierrez, G., Méndez, C. A., Sola, J. M., & de Prada, C. (2009). Data reconciliation and optimal management of hydrogen networks of a real refinery. In International Symposium on Advanced Control of Chemical Processes. Istambul, Turkey.

Sarabia, D., de Prada, C., Gómez, E., Gutierrez, G., Cristea, S., Sola, J. M., & Gonzalez, R. (2011). Data reconciliation and optimal management of hydrogen networks in a petrol refinery. Control Engineering Practice.

Wozny, G., & Arellano-Garcia, H. (2007). Special issue on process optimization and control in chemical engineering and processing. Chemical Engineering and Processing: Process Intensification, 46, 1041-1042.

Yang, N., & Wen, F. (2005). A chance constrained programming approach to transmission system expansion planning. Electric Power Systems Research, 75, 171-177.

3 REAL TIME OPTIMIZATION

ABSTRACT

This chapter is focused in the modifier-adaptation methodology implemented in the RTO layer of the optimization architecture, to overcome the uncertainties related with the mismatch produced in the modeling stage. In particular we present three proposals to deal with some challenges related with its application. (1) An intermediate layer between the RTO and the model control stage to avoid the infeasibilities that can be produced in the evolution of the method. (2) A reformulation of the modifier-adaptation method as a nested optimization problem to evade the step related with the estimation of the process gradients, increasing the application of the method for systems where the process derivative is not available. (3) A proposal to include the modifiers in the context of dynamic optimization, in order to manage the modeling mismatch for a receding horizon problem. Every proposal has been tested and compared with previous development in simulation examples, with the aim to evaluate their performance. As a conclusion we can say that applying the strategies presented, we are able to converge to the real optimum of the process for a RTO implementation, in a feasible way and skipping the process gradient estimation, making the whole algorithm more robust in terms of real life applications. Regarding the inclusion of the modifiers in dynamic optimization, only preliminary results and a very limited implementation have been presented, however some ideas can be useful for more general applications in the future.

KEYWORDS

Real time optimization, Modeling mismatch, Uncertainty, Modifier-Adaptation methodology, Infeasibility controller, Nested modifier-adaptation, Dynamic optimization.

RESUMEN

Este capitulo se enfoca en el método adaptación de modificadores utilizado en la capa de optimización en tiempo real para evitar la incertidumbre en la optimización producida por errores de modelado. En particular se han propuesto tres modificaciones para mejorar su implementación. (1) Inclusión de una capa intermedia entre la optimización y el control que modifica el valor de las variables de decisión sugeridas por el optimizador si es que se detecta una violación en las restricciones del proceso, haciendo factible la convergencia del algoritmo. (2) Reformulación del método de adaptación de modificadores como un problema de optimización anidado, con el objetivo de evitar la estimación de los gradientes del proceso y así ampliar su marco de aplicación. (3) Inclusión de los modificadores en un problema de optimización dinámica con horizonte móvil, con la idea de manejar la incertidumbre en la capa de optimización si es que esta tiene características dinámicas. Cada una de las propuestas ha sido aplicada a ejemplos de simulación para estudiar su desempeño y compararlo con el método actual. A modo de conclusión, se puede indicar que para la optimización estática se ha detectado el óptimo del proceso de manera factible y sin la necesidad de estimar las derivadas del proceso, haciendo el método más robusto a las condiciones reales. Respecto a la optimización dinámica, se ha propuesto la inclusión de los modificadores para un reducido conjunto de sistemas, sin embargo estas ideas pueden ser útiles para aplicaciones más generales en el futuro.

PALABRAS CLAVES

Optimización en tiempo real, Error de modelado, Incertidumbre, Adaptación de modificadores, Control de infactibilidades, Adaptación de modificadores anidada, Optimización dinámica.

3.1 Introduction

Real time optimization is a tool that is used in industrial facilities to improve the benefits of a process. In many real applications, an optimization layer is present above the control level with the aim of updating the set-points of the control layer using an economic criterion. A modern process plant may have thousands of measurements and control loops, so, in practice, a hierarchy for control and decision making is used as shown in Figure 3.1, where this structure is based on the successive refinement of time scales from top to bottom (Darby et al., 2011; Young, 2006). In addition, there are also varying spatial scales, from the plant planning layer at the top to unit regulatory control via single PID loops at the bottom.

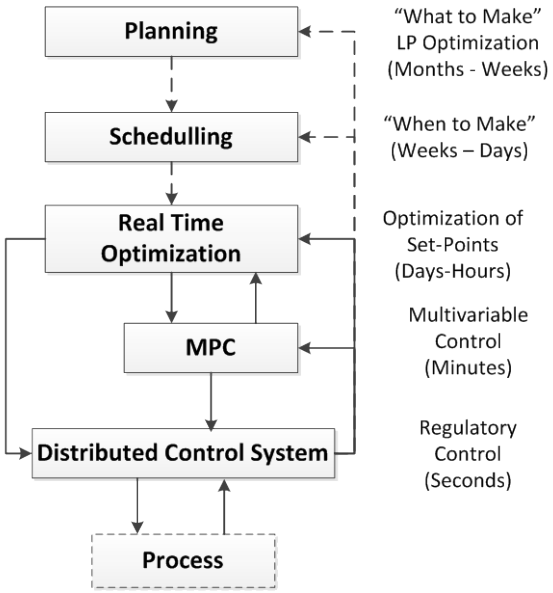

FIGURE 3.1: OPTIMIZATION HIERARCHY

Each of the layers present in the optimization hierarchy has different objectives, and therefore different optimization problems to be solved.

PLANNING: is concerned with "what and how" based on economics and forecasts, and answers such questions as what feedstock to purchase, which products to make, and how much of each product to make. In almost all refineries and larger chemical plants, a linear program (LP) or successive LP is used for planning and is based on an overall plant profit objective function.

SCHEDULING: is concerned with "when". Scheduling addresses the timing of actions and events necessary to execute the chosen plan, with the key consideration of being feasible. Scheduling deals with such issues as the timing of the deliveries of feeds, product lifting and operating mode changes, and avoiding storage problems (overflow or shortage). A range of tools are used across the industry for scheduling, from whiteboards and spreadsheets to tools involving simulation models, heuristic rules, and optimization (Valleur & Grue, 2004).

REAL TIME OPTIMIZATION: is executed in real time at different intervals, and there is automatic, continual feedback from the process (indicated by the solid lines entering from the right in Figure 3.1). At the planning and scheduling levels, feedback and model updating is not automatic and is performed intermittently (indicated by dotted lines in Figure 3.1). Outputs from planning and scheduling are shown as dotted lines to signify that human intervention is required to determine and set certain objectives of the next level, which are normally not simply a one-to-one translation of objectives to targets and limits. Day-to-day logistical issues from scheduling can have a significant impact on plant operations, overriding decisions that are made at the local optimization level and requiring target and limit changes at the constraint control level. The goal is to calculate the process optimal operating point or optimal operating trajectory to be applied for

feedback controllers of directly subordinate layers. Usually, the task of this layer is implemented by a real-time optimization package (RTO) which is concerned with implementing business decisions in real time based on a calibrated non-linear steady-state model, with model detail that the planning model does not have. RTO is implemented for economically justified cases and is typically formulated based on a profit function of the plant. However, RTO is not only optimization but a complex structure consisting on several subsystems for measurement validation, steady-state detection, process model updating, and model-based optimization, all of which are important for a successful performance. Gattu and coworkers and Bieker and coworkers describe the industrial requirements and challenges of RTO in refineries and offshore oil and gas production systems (Bieker et al., 2006; Gattu et al., 2003). Sometimes, instead of RTO, this level reduces to a local economic optimization, limited to a process unit using the same models as in the MPC layer.

CONSTRAINT (MULTIVARIABLE MPC) CONTROL: The MPC(s) provide the minute-to-minute dynamic control of the plant and give some amount of optimization capability. Although industries have invested heavily in MPC controllers, an adequate use and judicious implementation is necessary (Young, 2006) due to sub-optimal and/or poor outcomes reported by many industry experts (Wang, 2011).

DISTRIBUTED CONTROL SYSTEM (DCS): which is designed to collect and record sensor measurements from the process and implement the regulatory control (typically PID), usually it is executed on a second time scale (sub-second to multi-second). The DCS layer is typically the main operator interface for monitoring and controlling the plant including generating alarms for abnormal situations. Also

implemented at this level are advanced regulatory controls such as cascade controllers, ratio controllers, constraint controllers, and sequencing controls.

The hierarchical nature of planning, scheduling, and control follows the natural flow of decisions in an organization and exists in every production environment. Clearly, they have to work together in a coordinated and integrated fashion. The integration should be based not only on technical solutions but on changes in human behavior to get acceptance and use of more sophisticated tools and changes in the proper organizational structure (Shobrys & White, 2002) . The hierarchical structure, also allows separating in an appropriate way the different decisions that must be taken in the process, since the layers are executed in different time scales and with different kinds of models.

RTO presents important challenges to be addressed, such as, proposing algorithms that manage the existing model uncertainty and handling the interaction between the RTO and MPC layers.

The developments of algorithms able to find the optimum of the process under modeling mismatch assumptions, is a topic that has been studied since the RTO methodology first appears with the two-stage algorithm proposed by Bamberger and Isermann (Bamberger & Isermann, 1978): parameter estimation and economic optimization, that is solved in an iterative way. In their implementation the authors proposed adapting the nonlinear model of the plant, by means of solving a parameter estimation step before the economic optimization using the available measurements from the process (equation (3.1)), obtaining the value of the uncertain parameters α that makes the model as close as possible to the real plant. With the updated model, the second step is solving the economic optimization to obtain the next value of the manipulated variables to be applied in the upper layers of the process (equation (3.2)). Figure 3.2 summarizes the two-level algorithm implementation.

$$\begin{aligned} &\min_{\alpha} F_{id} \coloneqq (y-\bar{y}_k)^T R (y-\bar{y}_k) \\ &s.t.: \\ &h(y, u_{k-1}, \alpha) = 0 \\ &g(y, u_{k-1}, \alpha) \leq 0 \end{aligned} \tag{3.1}$$

$$\begin{aligned} &\min_{u \in U} F_{eco} \coloneqq f(y, u, \alpha_k) \\ &s.t.: \\ &h(y, u, \alpha_k) = 0 \\ &g(y, u, \alpha_k) \leq 0 \end{aligned} \tag{3.2}$$

Being: $y \in \mathbb{R}^{n_y}$ the measured variables, $u \in \mathbb{R}^{n_u}$ the decision variables, $\alpha \in \mathbb{R}^{n_\alpha}$ the uncertain parameters, $f: \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\alpha} \to \mathbb{R}$ the objective function, $g: \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\alpha} \to \mathbb{R}^{n_g}$ the inequality constraints of the optimization and $h: \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\alpha} \to \mathbb{R}^{n_y}$ the model of the process. The bar " ¯ " represent the variables that are directly measured or estimated from the process, while the subscript k denotes the actual RTO iteration.

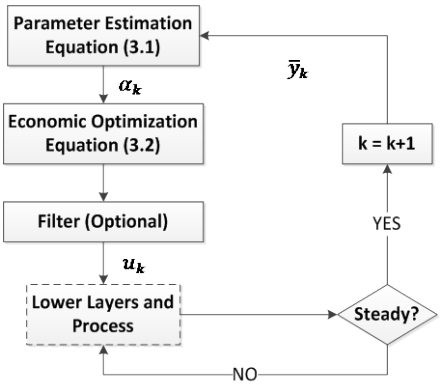

FIGURE 3.2: IMPLEMENTATION OF THE TWO-STAGE ALGORITHM

In Figure 3.2, it can be noted that the outcomes of the optimization problem, can be filtered in order to make the evolution of the RTO application smoother. Even though this step is optional, in practice it is applied most of the times by the operators (Darby et al., 2011). In general

terms the filtering procedure can be understood as a maximum size of the step, reducing locally the feasible region of the optimization around the actual operating point or using the past values from RTO iterations in an affine combination with the actual solution of the equation (3.2).

Most of the times, the model will not be a faithful representation of the reality, either because of the partial knowledge of the process or due to the simplifications that makes the optimization solvable in a reasonable time. Moreover, the two problems interact such that the solution of the optimization problem is dependent upon the values of the model parameters, that is to say, $u_k(\alpha_k)$. While the parameter estimates will change according to the controller settings given by the economical optimization in previous iteration, i.e., $\alpha_k(u_{k-1})$. These two factors: the interaction of the two layers and the modeling mismatch, produce that the two-step algorithm by itself will not necessarily converge to the optimum of the process (Roberts, 1979). Therefore, this interaction must be taken into account in the adaptation step in order to achieve the real optimum of the process under modeling mismatch. The convergence to this point has been studied by Biegler and coworkers who present the conditions to reach the optimum of a complex problem by using a simplified one: (1) The gradient of both models must be the same in every point within the entire feasible region, and (2) the KKT conditions of the complex model must correspond to the KKT conditions of the simplified one (Biegler et al., 1985). Later, Forbes and Marlin apply this idea to real process optimization, giving analogous conclusions about necessary conditions to reach the process optimum in a model-based optimization: at the real optimum, the KKT conditions of the process must be the same than the ones of model (Forbes & Marlin, 1994; Forbes et al., 1994).

To cope with the uncertainty already mentioned and to drive the process to its real optimum point, there have been several developments in RTO, including steady-state and dynamic systems. Chachuat and coworkers (Chachuat et al., 2009) classified these developments in three main groups

according to how the adaptation of the model and/or the decision variables are performed when measurements are available: *model-parameter adaptation*, *modifier-adaptation,* and *direct input adaptation*. The authors present an exhaustive analysis of the three methodologies concluding that the modifier-adaptation method presents important advantages over the other approaches since it combines the main characteristics of both: model-based optimization and direct adaptation of inputs using measurements directly from the process. Taking this into account, in this thesis we have worked using this strategy.

3.2 Modifier-Adaptation Methodology

As it was said previously, it is necessary to manage the interaction between the parameter estimation and the economic optimization step in order to find the process optimum. With this in mind, the ISOPE[2] algorithm was presented by Roberts (Roberts, 1979). Briefly, the ISOPE decouples both problems by introducing two additional constraints (equation (3.3)).

$$\begin{aligned} \alpha_k &= \sigma_k \\ u_k &= v_k \end{aligned} \tag{3.3}$$

With the coupling relation, the two steps of the RTO algorithm can be rewritten as:

$$\begin{aligned} &\min_{\alpha} F_{id} \coloneqq (y - \bar{y}_{k-1})^T R (y - \bar{y}_{k-1}) \\ &s.t.: \\ &h(y, v_{k-1}, \alpha) = 0, \qquad g(x, y, v_{k-1}, \alpha) \leq 0 \end{aligned} \tag{3.4}$$

$$\begin{aligned} &\min_{u \in U} F_{eco} \coloneqq f(y, u_k \sigma_k) \\ &s.t.: \\ &h(y, u, \sigma_k) = 0, \qquad g(y, u, \sigma_k) \leq 0 \end{aligned} \tag{3.5}$$

[2] ISOPE is de acronym of "*Integrated System Optimization and Parameter Estimation*"

Assuming that the optimum is inside the feasible region and denoting the economic optimization problem as $\min_u F_{eco}(u, \sigma)$, the stationary point of the integrated problem can be summarized as:

$$\begin{aligned} \nabla_u \mathcal{L} &= \nabla_u F_{eco_k} - \lambda_k = 0, \qquad & \nabla_\alpha \mathcal{L} &= \nabla^2_{\alpha\alpha} F_{id_k} - \mu_k = 0 \\ \nabla_v \mathcal{L} &= \nabla^2_{\alpha v} F_{id_k} + \lambda_k = 0, \qquad & \nabla_\sigma \mathcal{L} &= \nabla_\sigma F_{eco_k} + \mu_k = 0 \end{aligned} \tag{3.6}$$

Being λ and μ the Lagrange multipliers of equation (3.3) and $\mathcal{L}$ represents the Lagrangean function of the integrated problem, defined as:

$$\mathcal{L} = F_{eco_k} + \lambda^T (v_k - u) + \mu^T (\alpha_k - \sigma_k) + \eta^T \nabla_\sigma F_{id_k} \tag{3.7}$$

Since the optimum is inside the feasible region, the stationary point of the equation (3.6) correspond to the conditions of the optimum of the unconstrained problem from equation (3.8), which is a modified version of equation (3.5) when the inequality constraints g are not active.

$$\begin{aligned} &\min_{u \in U} F_{eco} \coloneqq f(y, u, \alpha) + \lambda_k^T u \\ &s.t.: \\ &h(y, u, \alpha) \end{aligned} \tag{3.8}$$

The parameter λ_k can be obtained solving equation (3.6). If it is assumed that the parameter estimation problem is square, i.e., $n_y = n_\alpha$, the expression to calculate the Lagrange multiplier λ_k is:

$$\lambda_k^T = [\nabla_v \bar{y}_{k-1} - \nabla_v y_{k-1}][\nabla_\alpha y_{k-1}]^{-1}[\nabla_\sigma f_{k-1}], \tag{3.9}$$

In equation (3.9) and during the entire chapter, the $k-1$ subscript indicates that the quantities have been calculated with the values of the previous RTO iteration, also $\nabla_n m$ and $\nabla^2_{np} m$ represent the partial derivatives $\frac{\partial m}{\partial n}$ and $\frac{\partial^2 m}{\partial n \partial p}$ respectively.

Assuming that equation (3.3) holds, and applying the chain rule to equation (3.9), λ_k becomes in a modifier that takes into account the difference of the gradient of the cost function with respect to the decision variable u, between the model and the process evaluated at the previous RTO iteration. Therefore, λ_k can be viewed as a corrector of the gradient, because it changes the direction of the cost function calculated with the model using data available from the measurements: the gradient of the real outputs. Convergence analysis of the algorithm to the true optimum of the process is presented by Brdys and Roberts (M. Brdys et al., 1987; M. Brdys & Roberts, 1987). Starting with the ISOPE algorithm, Roberts and coworkers present some variations, in order to include a wide range of industrial situations (Roberts, 1995), and take into account indirectly the process dependent inequality constraints (M. Brdys et al., 1986; M. Brdys & Roberts, 1986; Lin, Chen, et al., 1988; Lin, Hendawy, et al., 1988; Michalska et al., 1985).

The success of the ISOPE method (and its derivatives) relies not only on the idea that the interaction between both optimization problems is considered, but on the fact that the optimality conditions of the model and the process are fulfilled in the same point (under the supposition of a perfect matching between the predicted and measured outputs). Hence, the true optimal operation can be found if the algorithm converges (M. Brdys & Roberts, 1987). Later, Tatjewski showed that the convergence to the true optimum in the ISOPE method does not depend on the parameter estimation problem, but only on the equality between the output of the process and the model in each iteration. For this reason, the same results can be obtained using a bias corrector in the constraints of the model that is updated in each iteration using the process measurements (Tatjewski, 2002). Because of the parameter estimation is no longer needed, the author proposed another name for the algorithm: *Modifier gradient optimization set-point control* (equation (3.10)).

$$\begin{aligned} &\min_{u\in U} F \coloneqq f(y+b_k,u,\alpha)+\lambda_k^T u \\ &h(y+b_k,u,\alpha)=0 \end{aligned} \tag{3.10}$$

$$with:$$
$$\lambda_k^T \coloneqq \nabla_u \bar{f}_{k-1} - \nabla_u f_{k-1}, \qquad b_k \coloneqq \bar{y}_{k-1} - y_{k-1}$$

It can be noted that the subscript *"eco"* has been neglected, since the parameter estimation step has been replaced by the bias corrector b_k that ensures the equality between the measurements from the process and the predictions of the model, under convergence assumptions. The modified optimization from equation (3.10), takes into account first and zero order corrections of the model, which are calculated with data from the process. Following these ideas, Gao and Engell proposed an extension of the method of Tatjewski in order to handle process-dependent constraints (Gao & Engell, 2005). The authors defined the constraint modifier γ_k and b_k, to adapt the process dependent inequality constraints with a zero and first order correction (equation (3.11)).

$$\begin{aligned} &\min_{u\in U_m} F \coloneqq f(y+b_k,u,\alpha)+\lambda_k^T u \\ &s.t.: \\ &h(y+b_k,u,\alpha)=0 \\ &G_m \coloneqq g(u,\alpha)+\gamma_k^T(u-u_{k-1})+\epsilon_k \le 0 \\ &U_m = [u^L,u^U] \\ &u^L \coloneqq u_{k-1}-\Delta u, \qquad u^U : u_{k-1}+\Delta u \end{aligned} \tag{3.11}$$

$$with:$$
$$\lambda_k^T \coloneqq \nabla_u \bar{f}_{k-1} - \nabla f_{k-1}, \qquad b_k \coloneqq \bar{y}_{k-1} - y_{k-1}$$
$$\gamma_k^T \coloneqq \nabla_u \bar{g}_{k-1} - \nabla_u g_{k-1}, \qquad \epsilon_k \coloneqq \bar{g}_{k-1} - g_{k-1}$$

It can be noted that the feasible region of the decision variables U_m has been reduced in order to make the corrected inequality constraint g_m closer to the one estimated from the process. Using this method, the authors reached optimum yield of recovery in a simulated batch chromatography system identified from a pilot-scale system.

The modified constraint g_m, has the following properties:

Its value is the same as the value of the constraint measured from the process if $u = u_{k-1}$.

Its first-order derivative has the same value as the one estimated from the process if $u = u_{k-1}$.

Therefore, under convergence assumptions ($u_{k-1} \to u_k$) complementary slackness and primal feasibility from the necessary conditions of optimality (NCO) of equation (3.11), have only terms measured from the process.

From the equality of the KKT conditions of the model with the process, the use of modifiers was generalized by Marchietti and coworkers (Marchetti et al., 2009a), presenting the *Modifier-Adaptation Methodology* (equation (3.12)). The authors removed the modifier b_k, justifying the use of the remaining modifiers since, under convergence assumptions, they allow matching the KKT conditions of the modified model with the ones from the process. In equation (3.12) we have neglected the model $h(\cdot)$ since the equality between y and $\bar{y}$ is no longer needed.

$$\begin{aligned} &\min_{u \in U_m} F \coloneqq f(u,\alpha) + \lambda_k^T u \\ &s.t.: \\ &G_m \coloneqq g(u,\alpha) + \gamma_k^T (u - u_{k-1}) + \epsilon_k \leq 0 \\ &U_m = [u^L, u^U] \end{aligned} \tag{3.12}$$

The KKT conditions of equation (3.12) are:

$$\begin{aligned} &\nabla_u f_k + \lambda_k^T + \mu_k^T (\nabla_u g_k + \gamma_k^T) + \xi^{U^T} - \xi^{L^T} = 0 \\ &\mu^T (g(u,\alpha) + \gamma_k^T (u - u_{k-1}) + \epsilon_k) = 0 \\ &\mu \geq 0, \qquad g(u,\alpha) + \gamma_k^T (u - u_{k-1}) + \epsilon_k \leq 0 \end{aligned} \tag{3.13}$$

Assuming that the iterative algorithm converges into a stationary point, then $u_{k-1} \to u_k \to u^\star$. If we replace this assumption in equation (3.13), it

can be easily obtained the optimality conditions the process, as equation (3.14) shows.

$$\begin{aligned} &\nabla_u \bar{f}^\star + \mu^{\star T} \nabla_u \bar{g}^\star + \xi^{U^T} - \xi^{L^T} = 0 \\ &\mu^T(\bar{g}^\star) = 0 \\ &\mu \geq 0, \qquad \bar{g}^\star \leq 0 \end{aligned} \tag{3.14}$$

Therefore, by modifying the economic optimization with λ, γ and ϵ the iterative implementation of the modifier-adaptation methodology will converge to the real optimum of the process. This algorithm has the advantage that it is not necessary to know a priori the active set of constraints in the process optimum. Also, it relaxes the conditions of reaching the process optimum: the adequacy and accuracy of the model of the process defined previously by Forbes and coworkers (Forbes & Marlin, 1994; Forbes et al., 1994) are reduced to ensure that the second derivative of the cost function with respect to the manipulated variables evaluated in the process optimum must be positive definite: $\nabla_r{}^2 f(\bar{u}^\star) > 0$, condition that does not depend on the modifier values, but only in the model, and it is necessary to ensure optimality with the second order sufficient conditions.

Even though the methodology can converge to the process optimum, it can be done following an infeasible path since there is not guarantee of feasibility in intermediate iterations. To overcome this situation, the authors suggest the use of a first order filter (equation (3.15)), in order to smooth the evolution of the solutions reaching the true optimum of the process avoiding infeasible points.

$$\begin{aligned} \lambda_k^T &= (1 - K_\lambda)\lambda_{k-1}^T + (1 - K_\lambda)\left(\nabla_u \bar{f}_{k-1} - \nabla_u f_{k-1}\right) \\ \gamma_k^T &= \left(1 - K_\gamma\right)\gamma_{k-1}^T + \left(1 - K_\gamma\right)(\nabla_u \bar{g}_{k-1} - \nabla_u g_{k-1}) \\ \epsilon_k &= (1 - K_\epsilon)\epsilon_{k-1} + (1 - K_\epsilon)(\bar{g}_{k-1} - g_{k-1}) \end{aligned} \tag{3.15}$$

Being K_λ, K_γ and K_ϵ matrices with the adequate dimensions which act as filter constants. The use of the filter from equation (3.15) also helps to the

convergence of the algorithm (Marchetti et al., 2009a, 2010). There are not methods to estimate values of K. However, in general it is recommended the use of diagonal matrices that makes the correction of each modifier independent of the others. This is the only recommendation that can be applied in practice, because convergence criteria are given only for optimization problems where the value of the process optimum is known a priory, which is a contradiction for a RTO scheme. The modifier-adaptation algorithm has been summarized in Figure 3.3.

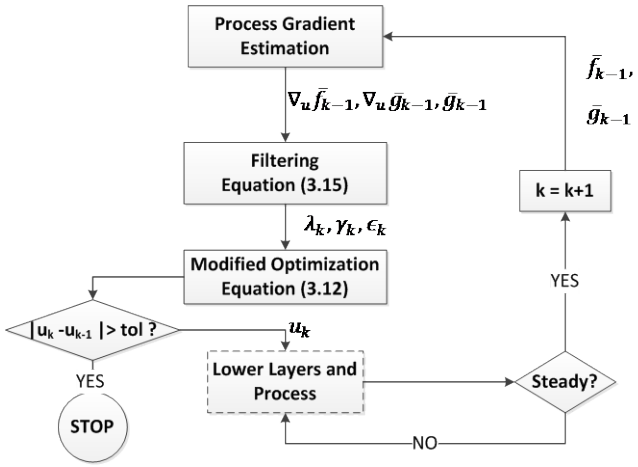

FIGURE 3.3: IMPLEMENTATION OF THE MODIFIER-ADAPTATION ALGORITHM

It can be noted that at the top of the diagram is represented the process gradient estimation step, unlike the gradient estimation of the model which is not represented. This is because the model derivatives can be obtained easily, however the process gradients can be very hard to estimate, and in fact this is the key issue in the implementation of the methodology as it will be seen in the next section where a simulation example will help to identify the challenges of the algorithm.

3.3 Implementation of the Modifier-Adaptation Methodology

The modifier-adaptation approach allows reaching the true optimum of the process in an iterative implementation, updating some modifiers with information measured or estimated directly from the process, which implies some additional efforts related with the identification of the modifiers. In particular there is an important assumption in the methodology: the capacity to estimate the process gradient in an accurate way. This is the key issue of the RTO methods based in modifiers: ISOPE, modifier gradient optimization set-point control and modifier-adaptation methodology(M. A. Brdys & Tatjewski, 2005; Gao & Engell, 2005; Mansour & Ellis, 2003; Marchetti et al., 2009b).

In chapter 4 of the book of Brdys and Tatjewski (M. A. Brdys & Tatjewski, 2005) there is an interesting discussion about the estimation of the process mapping derivatives. The authors justifies the importance of the method to estimate the process derivatives based in the fact that the time spent in their evaluation may be significantly higher than the whole remaining measurement and calculation during each iteration. That is why many attempts have been made to overcome this weak point, to find more effective realizations of the perturbation-based algorithms. There are three groups of these methodologies according of the type of information used to estimate the gradients: steady-state measurements, dynamic measurements and other attempts to add dynamic behavior to the modified RTO layer. The first two will be discussed in this section and are summarized in the work Mansour and Ellis (Mansour & Ellis, 2003), while the last one will be treated in section 3.7, since it also tries to overcome the challenges of the RTO methods related with the already commented interaction produced between the control and the optimization layers in the process decision hierarchy.

Process Derivative Mapping using steady state measurements: There are three main methods that attempt to estimate the process derivatives using the information from the measurements obtained in the steady state.

FINITE DIFFERENCES (Roberts, 1979): This is the most straightforward approach to estimate the values of the process derivatives and was presented by Roberts at the same time as the ISOPE algorithm. It consists on applying several changes around the actual operating point in order to evaluate the changes in the outputs. According to Ellis and Mansour, this method can give sufficient accuracy of the process derivatives in an acceptable time span for a small and noise-free process with reasonably rapid dynamics. However, it has been shown to be inefficient for large and slow processes because of the time needed for the estimation. In addition, the inherent noise of process signals reduces the performance of this method considerably (Mansour & Ellis, 2003).

DUAL CONTROL OPTIMIZATION (M. Brdys & Tatjewski, 1994): To calculate the gradients with this technique, it is assumed that there are as many past values of the measured variables as number of manipulated variables. Then, by using the definition of directional derivative, the partial derivative of each measurement can be estimated. In this calculus, the inversion of a matrix formed with the differences of the manipulated variables in different time instants is required; therefore, its condition number is crucial to estimate gradients accurately and with low influence of the process noise. A mixed estimation has been proposed by Gao and Engell (Gao & Engell, 2005). Marchetti and coworkers presented a work were the condition number of the difference matrix is taken into account in the optimization problem. In this work, the authors exhibit a complete description about the error sources to calculate

the gradients: truncation introduced by finite difference approximation of the derivatives and measurement noise (Marchetti et al., 2010).

BROYDEN'S APPROXIMATION (Roberts, 2000): This method is based on the Broyden family of algorithms (Luenberger & Ye, 2008) to update the previous estimate of the derivatives by using the actual measurements of the process. Even though no additional perturbations are required for this method, special care must be taken to avoid numerical errors in the update. Also, the influence of the initial guess of the gradient required in Broyden methods should be noted.

Some authors also add an additional method based in the existence of parallel units that can be perturbed at the same time in order to estimate the process gradients in few global iterations. Even when this approach reduces the total RTO iterations, the assumption of the existence of parallel units working at the same time and at the same conditions reduces its application to a few systems such as micro reactors (Srinivasan, 2007). In addition, an interesting work has been presented to overcome the imposition of identical units (Woodward et al., 2009), but the requirements of the existence of more than one unit working in parallel is mandatory. Because of the use of steady state measurements from the units at the same time, this methodology can be classified as the parallel version of the previous methods.

Process Derivative Mapping using dynamic measurements: The idea here is to identify locally a dynamic model in the transient between two RTO iterations.

DYNAMIC MODEL IDENTIFICATION METHOD (H. Zhang & Roberts, 1990): This method is based on dynamic identification in the context of optimization presented by Bamberger and Isermann (Bamberger &

Isermann, 1978). The key issue is to approximate the real process by a local dynamic model using real process information. Because of the dynamic characteristic of the model, the waiting time for the steady state is avoided and the steady-state derivatives are calculated directly from the model derived from the identified dynamic representation. The local dynamic models can be linear and non-linear, and it is not necessary to predefine its structure (Mansour & Ellis, 2003; Y. Zhang & Forbes, 2006). Applying this method, the authors ensure that the real process optimum can be achieved on-line, even if both models are very rough (the dynamic and the steady-state models). The scheme is suited for various types of constraints and nonlinearities (H. Zhang & Roberts, 1990).

The idea of using dynamic data to estimate the gradients seems to be very promising since it might not be necessary perturbing additionally the system to estimate the gradients, as steady-state methods does. However, severe difficulties can be found in the identification step that can compromise seriously the correct estimation of process derivatives. To get sufficiently accurate identification results the plant must be sufficiently excited, since it can hardly be assumed that passive identification experiment based in the measured recorded in the transient between two RTO iterations can always deliver sufficiently rich data. Hence, it is necessary planning an active identification experiment around the actual operating point adding additional dynamic perturbations into the process, falling in the same paradigm than the perturbation methods based in steady-state measurements. Moreover, these experiments are by no means easy to design (M. A. Brdys & Tatjewski, 2005), because identification methods for dynamic models are designed to get accurate information to reflect the local dynamics of the plant, not the static behavior of the system.

On the other hand, the methods based in static measurements sound as a reliable alternative to obtain an adequate mapping of the process

gradients, since they use steady-state measurements from the system. Therefore they are going to be compared in this section. In particular we will work with finite differences and with dual control optimization because of the convergence problems of the Broyden's algorithm previously commented.

The finite differences approach was presented with the ISOPE algorithm(Roberts, 1979). It is based in applying directly the definition of the finite differences into the process, perturbing the system around the actual point as many times as decision variables are. If $u_k \in \mathbb{R}^{n_u}$represents the actual operating point and $d \in \mathbb{R}^{n_u}$ is the perturbation step used to calculate the finite differences, the process derivatives of the objective and the inequality constraints can be obtained as:

$$\begin{aligned} \frac{\partial \bar{f}_k}{\partial u^i} &\approx \frac{\bar{f}_k\left(u_k^i \pm d^i\right) - \bar{f}_k\left(u_k^i\right)}{\pm d^i}, \qquad \forall i = 1 \dots n_u \\ \frac{\partial \bar{g}_k}{\partial u^i} &\approx \frac{\bar{g}_k\left(u_k^i \pm d^i\right) - \bar{f}_k\left(u_k^i\right)}{\pm d^i}, \qquad \forall i = 1 \dots n_u \end{aligned} \tag{3.16}$$

Equation (3.16) shows that it is necessary perturbing the process n_u times to estimate the process gradient and go to the next RTO iteration. This requirement can become the application of the modifier-adaptation methodology impossible for slow process and also for systems with many decision variables.

A new idea was proposed by Tatjewski and Brdys using the last $n_u + 1$ measurements to estimate the gradients, using an approximation of the directional derivatives, obtaining an estimate of the process gradient without the need of perturbations around the actual point: the *Dual Control Algorithm*.

Assuming that there is a collection of $n_u + 1$ points $u_k, u_{k-1}, \dots, u_{k-n_u}$ applied in the past RTO iterations, the vectors of differences with respect to previous points (s_{ki}) can be defined as:

$$s_{ki} = u_{k-i} - u_k, \qquad \forall i = 1 \dots n_u \tag{3.17}$$

Supposing that the vectors s_{ki} are linearly independent, it is possible to formulate a nonsingular square matrix $S_k \in \mathbb{R}(n_u \times n_u)$:

$$S_k = [s_{k1} \quad \cdots \quad s_{kn_u}]^T \tag{3.18}$$

The directional derivative $D\bar{y}_j(u_k, s_{ki})$ of the *j-th* plant output $\bar{y}_j$ at a point u_k in a direction s_{ki}, can be defined as:

$$D\bar{y}_j(u_k, s_{ki}) \coloneqq \lim_{\beta \to 0} \frac{\bar{y}_j\left(u_k + \beta \frac{s_{ki}}{\|s_{ki}\|}\right) - \bar{y}_j(u_k)}{\beta} \tag{3.19}$$

Assuming that $\bar{y}_j$ has a continuous derivative, $D\bar{y}_j(u_k, s_{ki})$ can also be rewritten as a function of its partial derivative with respect to u_k:

$$\|s_{ki}\| D\bar{y}_j(u_k, s_{ki}) = s_{ki}^T \frac{\partial \bar{y}_{j_k}}{\partial u} \tag{3.20}$$

Which can be generalized for n_u directions as:

$$S_k \frac{\partial \bar{y}_{j_k}}{\partial u} = \begin{bmatrix} \|s_{k1}\| D\bar{y}_j(u_k, s_{k1}) \\ \vdots \\ \|s_{kn_u}\| D\bar{y}_j(u_k, s_{kn_u}) \end{bmatrix} \tag{3.21}$$

On the other hand, from the definition of $D\bar{y}_j(u_k, s_{ki})$ in equation (3.19), defining $\beta = \mathscr{g}\|s_{ki}\|$ and assuming that $\|s_{ki}\| \to 0$, the limit can be removed :

$$\|s_{ki}\| D\bar{y}_j(u_k, s_{ki}) = \frac{\bar{y}_j(u_k + \mathscr{g} s_{ki}) - \bar{y}_j(u_k)}{\mathscr{g}} \tag{3.22}$$

Selecting $\mathcal{g} = 1$, the equation (3.22) can be replaced in equation (3.21). Solving for $\frac{\partial \bar{y}_{j_k}}{\partial u}$, a relation between the past measurements and the actual process derivative of the $\bar{y}_j$ measurement can be obtained.

$$\frac{\partial \bar{y}_{j_k}}{\partial u} = (S_k)^{-1} \begin{bmatrix} \bar{y}_j(u_{k-1}) - \bar{y}_j(u_k) \\ \vdots \\ \bar{y}_j(u_{k-n_u}) - \bar{y}_j(u_k) \end{bmatrix} \tag{3.23}$$

Therefore, defining $\bar{y}_1 = \bar{f}$ and $\bar{y}_{j+1} = \bar{g}_j, j = 1 \dots n_g$ it is possible to estimate the modifiers by using the actual operating point and the past n_u measurements. It can be noted that the inversion of the S_k matrix is required in order to obtain the process derivative, which implies that an additional constraint must be added in the degree of excitation of the process in order to ensure an accurate gradient guess in the next RTO iteration: the inverse of the condition number of the S_k matrix ($\delta(S_k)$), calculated with the next operating point, must be greater than a lower limit δ^L. This is the static equivalent of dual control optimization with respect to deal with active identification methods. Therefore, the *Dual Modifier-Adaptation Methodology* can be defined as equations (3.24) and (3.25) shows.

$$\begin{aligned} &\min_{u \in U_m} F \coloneqq f(u, \alpha) + \lambda_k^T u \\ &s.t.: \\ &G_m \coloneqq g(u, \alpha) + \gamma_k(u - u_{k-1}) + \epsilon_k \leq 0 \\ &U_m = [u^L, u^U] \end{aligned} \tag{3.24}$$

$$\delta(S_k) \geq \delta^L \tag{3.25}$$

The constraint from equation (3.25) represents the dual characteristic of the method: while the rest of the optimization tries to converge to the optimum of the modified model (primal objective from equation (3.24)), the dual constraint ensures that in the next RTO iteration the system will have enough energy to estimate the process gradient again (dual objective).

Equation (3.25) reduces the feasible region with respect to the original modified problem, which implies a loss of optimality in the RTO iteration that some authors justify in terms of the possibility to estimate the gradient of the process (M. Brdys & Tatjewski, 1994; M. A. Brdys & Tatjewski, 2005; Marchetti et al., 2010).

3.3.1 Application to an Oxygen Consumption Reactor

Finite differences and dual optimization was implemented in a modifier adaptation approach for a simulated reality with the aim to study the convergence to the real optimum of the process, by using a model with mismatch. Also the two-stage application was tested to justify the use of the modifiers.

The system is the CSTR reactor from Figure 3.4 that uses oxygen to perform a chemical reaction (Navia et al., 2012). In this unit, two streams (G_1 and G_2) each of them with a mixture of nitrogen and oxygen with different concentrations, are mixed and fed to a continuous stirred reactor (G_{in}). Together with the gas mixture, an aqueous solution containing sodium sulphite (Na_2SO_3) and cobalt chloride ($CoCl_2$) is sent to the reactor to be treated (F_{in}), with the objective to transform the entire sodium sulfite present in the liquid influent. The transformation is produced by means of the oxygen that comes from the gas mixture G_{in}, which can diffuse from the gas phase to the bulk of the solution and react with the Na_2SO_3, producing sodium sulfate (Na_2SO_4) with the help of $CoCl_2$ that acts as a catalyzer (equation (3.26)).

$$Na_2SO_3 + \frac{1}{2}O_2 \xrightarrow{CoCl_2} Na_3SO_4 \tag{3.26}$$

The aqueous product of the reaction is removed from the vessel in the F_{out} stream, keeping the level of the reactor constant. Because of the solubility limits of oxygen in water, only a fraction of the oxygen fed can be used in

the reactor, while the unused gas leaves the unit in the G_{out} stream that can be manipulated in order to control the pressure of the system.

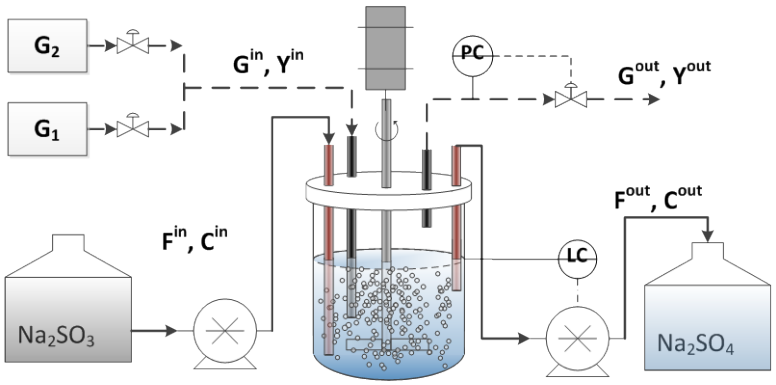

FIGURE 3.4: DIAGRAM OF THE OXYGEN CONSUMPTION REACTOR

3.3.1.1 Model of the Oxygen Consumption Reactor

To model the reactor, we have used a first order approach, considering the following assumptions:

- Because of the system is working at 1 bar, we can assume that the gases behaves ideally.
- The stirrer of the reactor makes the hydrodynamic behavior inside the vessel very similar to the CSTR[3] model.
- Zero concentration of Na_2SO_4 and O_2 in the aqueous influent
- The Nitrogen is considered as an inert in the entire system, therefore its diffusion rate can be neglected.
- The reaction from equation (3.26) is the only one that is happening inside the reactor, and no parallel reactions are verified.
- The equilibrium produced in the gas-liquid interface can be modeled using the Henry's Law because of the poor solubility of the oxygen.

[3] CSTR is the acronym for Continuous Stirred Tank Reactor

Mole Balance in the liquid phase for each compound:

$$V\frac{d}{dt}\left(C_{Na_2SO_3}^{out}\right) = F^{in}\left(C_{Na_2SO_3}^{in} - C_{Na_2SO_3}^{out}\right) - V \cdot r_{Na_2SO_3} \tag{3.27}$$

$$V\frac{d}{dt}\left(C_{Na_2SO_4}^{out}\right) = -F^{in}\left(C_{Na_2SO_4}^{out}\right) + V \cdot r_{Na_2SO_4} \tag{3.28}$$

$$V\frac{d}{dt}\left(C_{O_2}^{out}\right) = N_A \cdot V - F^{in} \cdot \left(C_{O_2}^{out}\right) - V \cdot r_{O_2} \tag{3.29}$$

Where r_{O_2} is the rate of conversion of the solubilized oxygen due to the chemical reaction, which can be calculated as:

$$r_{O_2} = K \cdot {C_{Na_2SO_3}^{out}}^{A} \cdot {C_{Co}}^{B} \cdot {C_{O_2}^{out}}^{C} \tag{3.30}$$

The relationship between the compounds involved in the chemical reaction can be obtained from the stoichiometric coefficients:

$$r_{Na_2SO_3} = r_{Na_2SO_4} = \frac{1}{2} \cdot r_{O_2} \tag{3.31}$$

Mole balance in gas phase:

$$\left(\frac{P \cdot V_g}{R \cdot T}\right)\frac{d}{dt}\left(Y_{O_2}^{out}\right) = \rho_{O_2}\left(G^{in} \cdot Y_{O_2}^{in} - G^{out} \cdot Y_{O_2}^{out}\right) - V \cdot N_A$$
$$G_1 Y_{O_2}^{1} + G_2 Y_{O_2}^{2} = G^{in} Y_{O_2}^{in} \tag{3.32}$$

$$\rho_{O_2}(G^{in} - G^{out}) - N_A \cdot V = 0 \tag{3.33}$$

The variable N_A corresponds to the oxygen flux transferred between each phase. Generally, in gas – liquid reactors, this term is a function of the gradient of the concentrations between the interface and the liquid bulk (Treybal, 1980). Assuming thermodynamic equilibrium in the interface, the oxygen concentration in the liquid that is in contact with the gas bubble, is

equal to maximum concentration that can be reached for a given temperature: the saturation concentration obtained from the Henry's Law, obtained from equations (3.34) and (3.35).

$$N_A = K_{LA} \cdot \left(C_{O_2}^{SAT} - C_{O_2}^{out}\right) \tag{3.34}$$

$$C_{O_2}^{SAT} = \left(\frac{P}{H}\right) \cdot Y_{O_2}^{out} \tag{3.35}$$

The nomenclature used in the model is listed in Table 3.1 and the superscripts "in" and "out" can be followed from Figure 3.4.

TABLE 3.1 NOMENCLATURE EMPLOYED IN THE OXYGEN CONSUMPTION REACTOR

Symbol	Meaning	Units
C	Molar concentration in liquid phase	mol/l
F	Volumetric liquid flow	l/h
G	Volumetric gas flow	l/h
H	Henry's constant	atm
K_{LA}	Volumetric mass transfer coefficient	l/h
K, A, B, C	Kinetic parameters	-
N_A	Molar flux in the interface	$mol/l\ h$
P	Pressure	atm
T	Temperature	$ºC$
R	Gas Constant	$atm\ l\ /K\ mol$
V	Volume of the liquid in the reactor	l
V_g	Volume of the gas chamber	l
Y	Mole fraction in gas phase	-
r	Conversion rate	$mol/l\ h$
ρ	Molar density	mol/l

In the system, the liquid influent F_{in} and effluent F_{out}, and also the system pressure P can be considered as known disturbances, whereas the influent gas streams G_1 and G_2 are manipulated variables that can be modified to fulfill certain operational goals.

3.3.1.2 RTO of the Oxygen Consumption Reactor

The purpose of the unit is to transform the entire incoming sulfite in sulfate in the cheapest way, finding the best combination of influents G_1 and G_2 since both streams present different purities and costs.

Because of the main purpose of the unit is to ensure the entire conversion of sulfite, enough oxygen must be fed in the reactor, meaning that the gas must be fed at least in a stoichiometric relation with the sulfite that can be viewed as the limiting reactive in equation (3.26). Defining the optimization problem in the stoichiometric point, i.e., minimizing the oxygen cost such that the outlet concentration of the dissolved oxygen and the sulfite are equal to zero, is not the best way to face the system for two main reasons: (1) in the real implementation there will be necessary the measurement of both concentrations, traduced in a more expensive installation since two composition probes are required, and (2) owing to the fact that the kinetic of the reaction under the presence of the catalyzer is very fast (Zhao et al., 2005) any disturbance will make the system infeasible. However, if we assume that an excess of oxygen must be fed to cope with possible disturbances in the feed, the constraint of transforming the entire influent sulfite in sulfate can be ensured maintaining the oxygen concentration in the liquid bulk of the reactor ($C_{O_2}^{out}$) greater than a lower bound ($C_{O_2}^{LO}$) because of the instantaneous reaction rate. This also implies that only the measurement of the oxygen concentration is necessary to check if the system is feasible. Therefore, the economic optimization problem to be solved in the RTO layer can be expressed as: minimizing the cost of oxygen, keeping the dissolved oxygen concentration at the aqueous effluent greater than a lower bound (equation (3.36)). Where C_{G_1} and C_{G_2} are the costs of G_1 and G_2 respectively. In the optimization problem $G^L$and G^U represent the lower and upper bounds of the manipulated variables.

Note that the optimization problem to be solved requires the exact model of the process in order to achieve the convergence to its optimum.

However, this situation is not realistic since there are several physical mechanisms that are not completely known: in particular we will focus our attention in the mass transfer that occurs in the gas-liquid interface and in the chemical reaction.

$$
\begin{aligned}
&\min_{G} F \coloneqq C_{G_1} G_1^2 + C_{G_2} G_2^2 \\
&s.t.: \\
&F^{in}\left(C^{in}_{Na_2SO_3} - C^{out}_{Na_2SO_3}\right) - V \cdot r_{Na_2SO_3} = 0 \\
&-F^{in}\left(C^{out}_{Na_2SO_4}\right) + V \cdot r_{Na_2SO_4} = 0 \\
&N_A \cdot V - F^{in} \cdot \left(C^{out}_{O_2}\right) - V \cdot r_{O_2} = 0 \\
&r_{O_2} = K \cdot {C^{out}_{Na_2SO_3}}^{A} \cdot {C_{Co}}^{B} \cdot {C^{out}_{O_2}}^{C} \\
&\rho_{O_2}\left(G^{in} \cdot Y^{in}_{O_2} - G^{out} \cdot Y^{out}_{O_2}\right) - V \cdot N_A = 0 \\
&G_1 Y^1_{O_2} + G_2 Y^2_{O_2} = G^{in} Y^{in}_{O_2} \\
&\rho_{O_2}(G^{in} - G^{out}) - N_A \cdot V = 0 \\
&N_A = K_{LA} \cdot \left(C^{SAT}_{O_2} - C^{out}_{O_2}\right) \\
&C^{SAT}_{O_2} = \left(\frac{P}{H}\right) \cdot Y^{out}_{O_2} \\
&\\
&g \coloneqq C^{LO}_{O_2} - C^{out}_{O_2} \leq 0 \\
&G \coloneqq [G_1 \quad G_2]^T, \qquad G \in [G^L, G^U]
\end{aligned}
\tag{3.36}
$$

In bubbled reactors, as the one used in the example, mass transfer is very often the limiting phenomenon (Quijano et al., 2010). Several factors can affect this effect: bubble diameter, power of the stirrer, temperature, pressure and gas flow among others. In general the mass transfer coefficient can be estimated as the product of the contribution of these effects using non-dimensional numbers (Treybal, 1980). For a given system, where only the gas flow can change while the rest of the factors that can affect the mass transfer can be considered constants, the value of K_{LA} can be expressed as:

$$K_{LA} = m(G^{in})^n \tag{3.37}$$

Where m and n are parameters that can be estimated experimentally or calculated from available correlations.

Regarding the chemical reaction that occurs at the bulk of the liquid phase, it can be said that the order of this reaction can change with the concentration of the reactants (Zhao et al., 2005). Therefore, kinetic parameters A, B and C from equation (3.30) are uncertain.

The previous discussion about the physical phenomena inside the reactor, give insights about the main sources of uncertainty of the system. Also, they provide the causes of more important modeling mismatches that can be expected. Therefore we can simulate the modeling mismatch expected in the real implementation using different expressions in the real process and the model used in the RTO layer, for the mass transfer and the chemical reactions. Table 3.2 summarizes the simulated modeling mismatch with the value of the parameters used in the process and the model.

TABLE 3.2 VALUE OF THE PARAMETERS USED IN THE PROCESS AND THE MODEL FOR THE OXYGEN CONSUMPTION EXAMPLE

System	n	m	A	B	C	K
Process	0.7	0.3342	1	0.5	1	3.7e9
Model	1	0.062	0	0	0	0.029

Note that the difference in the value of the parameters used in the model and the process implies a structural mismatch, because of the conversion rate in the model is constant, while in the process it depends on the concentration of the chemical compounds in the liquid phase.

The rest of the value for the parameters of the Oxygen consumption reactor that are common for the model and the process are summarized in Table 3.3.

TABLE 3.3 VALUE OF THE PARAMETERS FOR THE OXYGEN CONSUMPTION EXAMPLE

Parameter	Value
V	3
V_G	2
P	1.8
T	25
$Y_{O_2}^1$	0.7
$Y_{O_2}^2$	0.9

Since the objective function of the optimization problem only depends on the decision variables, it is not necessary to use the gradient modifier of the objective function λ, because there is no uncertainty in this part of the model. On the other hand, the inequality constraint depends on the model of the process, therefore it must be corrected with γ and ϵ. The correction with the corresponding modifiers of the optimization problem from equation (3.36) is presented in equation (3.38).

$$
\begin{aligned}
&\min_G F \coloneqq C_{G_1} G_1^2 + C_{G_2} G_2^2 \\
&s.t.: \\
&Steady - state\ model \\
&g_m \coloneqq C_{O_2}^{LO} - C_{O_2}^{out} + \gamma_k^T (G - G_{k-1}) + \epsilon_k \leq 0 \\
&G \coloneqq [G_1 \quad G_2]^T, \qquad G^L \leq G \leq G^U \\
&with: \\
&\gamma_k^T \coloneqq \left[-\frac{\partial \overline{C_{O_2}^{out}}_{\,k-1}}{\partial G} + \frac{\partial C_{O_2}^{out}{}_{k-1}}{\partial G} \right] \\
&\epsilon_k \coloneqq -\overline{C_{O_2}^{out}}_{\,k-1} + C_{O_2}^{out}{}_{k-1}
\end{aligned}
\tag{3.38}
$$

Besides the modifier-adaptation, the two-step algorithm will be applied in this example in order to analyze the effect produced by the corrections in the evolution of the algorithm. Because of the uncertainty has been focused in the physical phenomena that is produced inside the reactor, in the first step the parameters m and K will be updated using the information of the concentration of the outlet oxygen from the effluent of the process in the last operating point. Following the two-step paradigm,

with the updated model an economic optimization problem will be solved to estimate the optimal values of G. Equation (3.39) and (3.40) summarizes the two steps needed to apply this methodology.

$$\begin{aligned} &\min_{P} F_{id} \coloneqq \left(\overline{C_{O_2}^{out}}_{\,k-1} - C_{O_2}^{out}\right)^2 \\ &s.t.: \\ &Steady - state\ model\ evaluated\ at\ (G_{k-1}, P) \\ &P \coloneqq [m \quad K]^T \end{aligned} \tag{3.39}$$

$$\begin{aligned} &\min_{G} F_{eco} \coloneqq C_{G_1} G_1^2 + C_{G_2} G_2^2 \\ &s.t.: Steady - state\ model\ evaluated\ at\ (G, P_k) \\ &g_m \coloneqq C_{O_2}^{LO} - C_{O_2}^{out} \leq 0 \\ &G \coloneqq [G_1 \quad G_2]^T, \qquad G \in [G^L, G^U] \\ &P_k \coloneqq P \coloneqq [m_k \quad K_k]^T \coloneqq argmin\big(equation\ (3.39)\big) \end{aligned} \tag{3.40}$$

3.3.1.3 Results of the RTO methods in the Oxygen Consumption Reactor

The RTO methodologies, where tested in the oxygen consumption reactor, starting from three different feasible points (In the coming Figures: **A**, **B** and **C**) in order to analyze their convergence to the optimum of the process, which is represented in the figures with **R**. This point was obtained previously solving the optimization problem with the real model of the process, and mapping the entire feasible region, and lies on the constraint.

The evolutions of the decision variables are presented in Figure 3.5 for the two-step algorithm. In the figure the inequality constraint of the process $g \geq g^L$ is also represented by a discontinuous straight line, as well as the cost function.

The evolution of the two-step algorithm shows that the method converges into stationary points different than the real optimum of the process. As it was mentioned in the introduction, the interaction that is produced

between the two steps makes this method unable to identify the real optimum of the process. This interaction is also reflected in the fact that the three final points where different, which can be understood as the result of the relation that is produced between the estimation of the parameters that depends on the previous operating point of the process and the economic optimization as function of the previous model update.

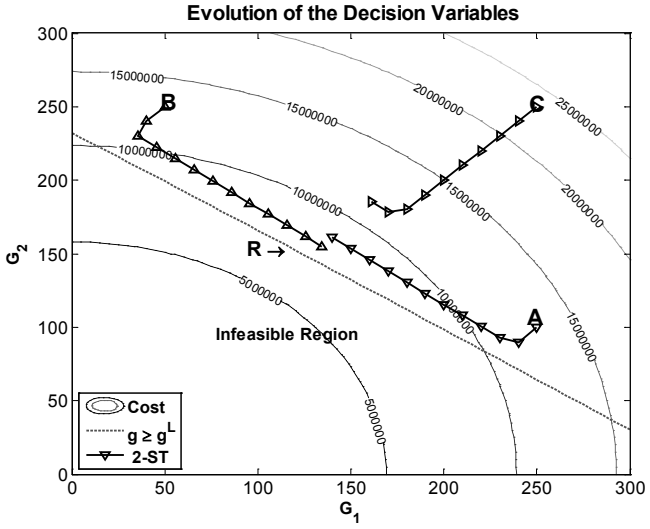

FIGURE 3.5: EVOLUTION OF THE DECISION VARIABLES WITH THE TWO-STEPS ALGORITHM IN THE OXYGEN CONSUMPTION REACTOR

Even though the final point is a function of the starting one, in the three cases there is a common pattern in the evolution of the decision variables: an off-set with respect to the real constraint. In the trajectories starting from A and B this gap very notorious since after the third iteration, the path formed with the decision variables approaches the real optimum forming a parallel manifold with the inequality constraint. On the other hand, in the initial C-point trajectory, this behavior is not so easy to see since the imaginary line $\overline{CR}$ is almost perpendicular to the inequality constraint in the optimum, implying that the trajectory described by the evolution might have the same direction; however, it can be noted how the last two

iterations changes the course of the evolution producing again a trajectory that runs in parallel with respect to $g \geq g^L$. Because of the real optimum is produced when the inequality constraint is active, the existence of this gap indicates how the two-step algorithm detects this active condition, which is quite expectable considering that the obvious solution of the optimization, even with modeling mismatch, is to supply the minimum quantity of oxygen that ensures feasibility. However, only the corrections in the parameters of the model are not enough to estimate accurately the process constraint, explaining the observed off-set.

Figures 3.6 and 3.7 present the evolution of the modifier adaptation methodology from equation (3.38), calculating the process derivatives with finite differences and dual method respectively.

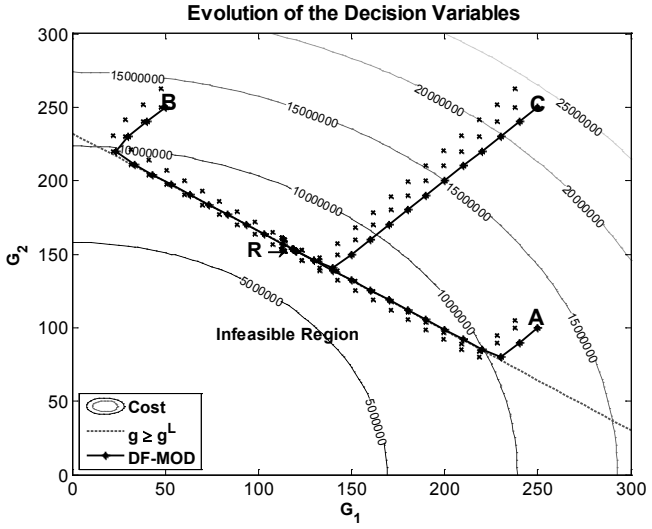

FIGURE 3.6: EVOLUTION OF THE DECISION VARIABLES WITH THE MODIFIER ADAPTATION ALGORITHM USING FINITE DIFFERENCES APPROACH IN THE OXYGEN CONSUMPTION REACTOR. DOTS REPRESENT THE PERTURBATIONS REQUIRED TO CALCULATE THE FINITE DIFFERENCES

Unlike the two-step algorithm, the correction of the curvature of the constraint of the model with process gradient information obtained from

the two methods tested, allows to estimate in a correct way the real inequality constraint. This can be observed in the fact that the evolution of the modifier-adaptation methodology from the three starting points detects the active constraint and place the path over it, until the real optimum of the process is detected.

As the gap between the model and the process has disappeared, a first order correction of the constraint is necessary in order to estimate it in an adequate way. This result is in concordance with the ones reported in the work of Marchietti and coworkers for a run-to-run batch reactor, where only a correction using ϵ is considered because of the process noise that makes the estimation of the process gradient impractical (Chachuat et al., 2009), which produces an off-set with the process optimum that is on the inequality constraints.

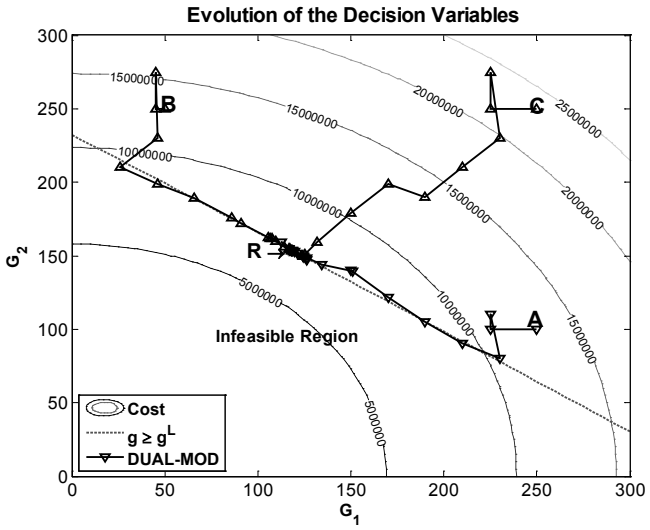

FIGURE 3.7: EVOLUTION OF THE DECISION VARIABLES WITH THE DUAL MODIFIER ADAPTATION ALGORITHM IN THE OXYGEN CONSUMPTION REACTOR

Regarding the performance of the two methods implemented to estimate the process derivatives, it can be noted that the finite differences approach

converges forming a straight line, meaning that the method estimates the optimum of the process forming an optimal path with respect to the number of iterations; however, it is necessary to take into account the necessary perturbations that must be added in each RTO point (represented with black dots in Figure 3.6), which increase the number of total RTO iterations by a factor of two. On the other hand, the evolution of the dual modifier adaptation is not so straight as the finite difference approach, noting that changes in the direction of the RTO iterations occurs before converging to the optimum of the process. Nevertheless, the number of iterations is significantly lower with respect to the implementation of the finite differences. This result is quite expected based in the idea that the method estimates the process derivatives from the previous iterations, and justify the fact that most of the authors working in the modified optimization methods, are using this strategy to estimate the process derivatives (Gao & Engell, 2005; Marchetti et al., 2010; Rodger, 2010), since it gives the less number of RTO iterations.

It is important to mention that the changes observed in the evolution of the dual modifier adaptation methodology, are produced because of the necessary excitation that the dual constraint imposes to system in order to estimate correctly the process gradient in the next iteration. At the same time, one can observe that they create some infeasibilities in the trajectory.

3.3.1.4 Remarks in the Modifier-Adaptation Implementation for the Oxygen Consumption Reactor

From the results of the implementation of the modifier-adaptation methodology in the oxygen consumption reactor, it can be said that the correction of the available model with the curvature data from the process is necessary in order to estimate in an accurate way the optimality conditions of the process. Also we can say that the method used to estimate the process gradients affects evolution of the system from the point of view of number of iterations and shape of the path formed with the decision variables.

Both remarks indicate that the implementation of the dual modifier adaptation methodology seems to be the best available alternative, in order to converge to the real optimum of the process in the less possible iterations, which is in accordance with the results from literature (M. Brdys & Tatjewski, 1994; M. A. Brdys & Tatjewski, 2005; Gao & Engell, 2005; Marchetti et al., 2008, 2009b, 2010). Taking this into account, the dual methodology will be applied throughout this thesis as a default implementation method in the modifier-adaptation algorithm.

3.4 Challenges on the Modifier-Adaptation Methodology

In previous sections we have reviewed the modifier adaptation methodology as an alternative to find the real optimum of the process, overcoming the unavoidable modeling mismatch that always exists between the real process and the abstractions used in the model-based optimization.

The implementation presented above also has reaffirmed the idea that in order to converge to the process optimum reducing the number of iterations, the dual algorithm must be implemented. However, there is a tradeoff in the number of iterations and the straightness of the path formed with the evolution of the decision variables, produced from the need to have an excited system to estimate the modifiers in an accurate way. This fact, plus the inherent condition of the modifier adaptation methodology where the convergence is ensured but not in a feasible path, can be important issues to be considered in the application of this algorithm, since feasibility is always over optimality. Taking this into account, it is necessary the application of some procedure able to cope with infeasibilities in the evolution of the system, in order to fulfill both, feasibility and optimality.

Besides, we have supposed that the convergence to the real optimum is ensured if the process derivatives can be estimated. However, this is not a trivial task, in particular for processes where the gradient can be very expensive to obtain or for processes with measurements contaminated with noise. The assumption of the gradient estimation is one of the bases of the modifier-adaptation methodology, no matter the method used to estimate the curvature of the real system. Hence, it is necessary to review this assumption in order to increase the field of application of the methods based in modifiers for the cases when the process derivatives are not available.

Finally, it has been mentioned previously the problems that appear in the interaction between the RTO and the control layers. As an alternative to get over this situation, there are several approaches, but one of the more promising ones is adding dynamic characteristics to the RTO layer (Engell, 2007; Gonzalez et al., 2001; Kadam & Marquardt, 2004; Würth et al., 2009, 2011). With this in mind, it is interesting trying to apply the concepts of modifiers in dynamic optimization, in order to handle with modeling mismatch in dynamic scenarios.

Therefore there are three main points where we focus our attention in order to increase the range of application of the modifier-based methods:

- **Handling Infeasibilities**

- **Avoiding the process gradient estimation step**

- **Appling the modifiers in dynamic optimization**

These three topics are going to be discussed in the coming sections.

3.5 Handling Infeasibilities in Dual Modifier - Adaptation Methodology

In process operation, feasibility is the most important issue. The usual practice in industrial facilities is using the degrees of freedom of the plant to ensure feasibility and the remaining ones to find improvements in the profit. This situation reflects the importance that the operators give to the operational constraints.

Safety is related with this fact, translated in ensuring that pressures, temperatures and other process variables must be always within the design boundaries. In the same way, optimality might also be affected in the case of infeasible operations of the process, for example: catalysts can be damaged if the concentrations of impurities are greater than a given value, chemical products might become useless if the formulation is not within the specifications or important fines can be applied if a given factory does not have enough stock of materials.

Modifier-adaptation methodology allows the convergence to the real optimum of the process but not always in a feasible path. Marchetti and coworkers, aware of this situation, recommend tuning the K matrix of the filter from equation (3.16) with low values in order to converge in a smooth way to the real optimum. Nevertheless, the chosen values will depend on the system and there is not a tuning procedure to estimate the values of the matrix that avoid (or at least reduce) the infeasibility occurrence (Marchetti et al., 2009a). A reduction in the constant of the filter can decrease the infeasibilities produced because of the aggressiveness of the updates of the in the modifier, nonetheless the ones produced because of the need of excitation of the system are not handled with this approach. Also it has been proposed the use of a NMPC to cope with the infeasibilities, however, the modeling mismatch assumption makes this alternative (and any other, based in models) not realistic.

In this section we propose the use of an infeasibility non-model based controller that modifies the values given by the RTO layer with the dynamic information from two consecutive RTO iterations to cope with possible infeasibilities.

3.5.1 The Infeasibility Controller

The option proposed here is to reduce the occurrence of infeasibilities in the intermediate points of the RTO evolution, by performing continuous corrections on the suggested values of the RTO layer in the transient produced between the k and the $k+1$ RTO iteration. With these corrections the idea is keeping all the time the system in the feasible side of the optimization region, waiting for the next steady state to look for optimality again.

The controller must work in the following way: if the infeasibilities are detected by a supervisory intermediate layer, the infeasibility controller is activated performing continuous corrections in the value of the decision variables at every sampling time until the system is driven into feasible conditions. The controller can be any non-model based one, such as a PI controller.

Assuming that the process constraints can be directly measured or estimated, which is a valid assumption considering that the modifiers are calculated using these variables, the infeasibility error of the process constraint g during the k RTO iteration, measured in the n^{th} sampling time of the infeasibility controller ($e_{g}^{n}{}_{k}$), can be defined as the maximum among zero and the difference between its lower (or upper) bound g^L, and the measured constrained variable at the current time ($\bar{g}_k^n$):

$$e_{g}^{n}{}_{k} \coloneqq \max\{g^L - \bar{g}_k^n, 0\} \tag{3.41}$$

The definition from equation (3.41) can be seen as an infeasibility error with respect to the model, because when the constraint is not violated the error

is equal to 0. Once the system crosses the border line of feasibility, the error becomes different to zero and corrective actions in the set-points given by the RTO layer must be applied for instance using the equation (3.42). However, any other non-model control law can be applied.

$$\begin{aligned} & u_k^{n\#} = u_k^{n-1} + K_g {e_g^n}_k \\ & with: \\ & u_k^0 \equiv u_k \end{aligned} \tag{3.42}$$

The corrections should modify the values from the RTO layer, if and only if an infeasibility is detected, whereas in other cases the solution proposed by the upper layer remains unchanged and the RTO implementation is the same than the method explained previously. Taking into account that u_k^0 is equal to the last value given by the RTO layer (u_k), if we suppose that the system remains feasible during the complete transient, taking, let us say m sampling times. Replacing these assumptions in equation (3.42) we obtain that $u_k^{m\#} = u_k^0 = u_k$, which is the solution of the RTO, meaning that during the whole transient the system tracked the set-point given by the upper layer. On the other hand, if the system detects infeasibilities the infeasibility error becomes positive and the set-points proposed by the RTO layer are updated with the proportional and the integral action of equation (3.42) until the system reaches feasibility again and the controller is deactivated.

If dual control optimization has been used as a method to estimate the process gradients, the corrections in the decision variables can affect the dual constraint from equation (3.25), which might lead to inaccurate estimation of the process gradient in the next RTO execution. Therefore, an additional controller can be added to ensure compliance with this constraint. Defining the infeasibility error of the inverse of the condition number in the n^{th} sampling time during the k RTO iteration (${e_\delta^n}_k$) as:

$$e^n_{\delta\,k} = \max\{\delta^L - \delta^{n\#}, 0\} \tag{3.43}$$

The inverse of the condition number in the actual sampling time ($\delta^{n\#}$) can be calculated as the difference between previous values recorded in the RTO layer, and the actual correction proposed by the infeasibility controller given by equation (3.42):

$$\begin{aligned} &\delta^{n\#} \coloneqq \delta\left(S_k^{n\#}\right) \\ &S_k^{n\#} = [s^n_{k0} \quad \cdots \quad s^n_{kn_u-1}] \\ &s^n_{k0} = u_{k-i} - u_k^{n\#}, \qquad \forall i = 0 \dots n_u - 1 \end{aligned} \tag{3.44}$$

Note that the variable u_{k-i} represents the values given in the i previous RTO iterations, which could also by modified by the infeasibility control layer in previous iterations. Finally, the value to be applied to the process can be calculated from the controller from equation (3.45). Therefore, we can argue that the primal feasibility of the optimization problem has been taken into account with the controller from equation (3.43), while the dual feasibility is controlled in equation (3.45).

$$u_k^n = u_k^{n\#} + K_d e^n_{d\,k} \tag{3.45}$$

Analogous to the primal infeasibility controller, the dual one has the property of preserving the last value of the RTO layer if the problem is feasible (which can be demonstrated setting $e^n_{d\,k}$ and $e^n_{g\,k}$ to zero $\forall n = 1 \dots m$), whereas this value changes to look for dual feasibility if the process is on the infeasible region.

Due to the fact that we are dealing with two different kinds of feasibilities: primal and dual, the system must have enough degrees of freedom to ensure that all the constraints are going to be fulfilled. If this is the case, then the pairing must be done following the RGA paradigm (Bristol, 1966) giving priority to the primal constraints. On the other hand, if there is not

enough degrees of freedom available, the dual constraint can be removed and replaced by the use of the finite differences approach as the alternative to estimate the process derivative when the dual constraint is violated, in the same way than Gao and Engell propose the alternation of both perturbation-based methods (Gao & Engell, 2005). The decision of using one or another method to estimate the process curvature would be taken according to the value of equation (3.44). The implementation of the infeasibility controller with the primal and dual feasibility objectives is summarized in Figure 3.8.

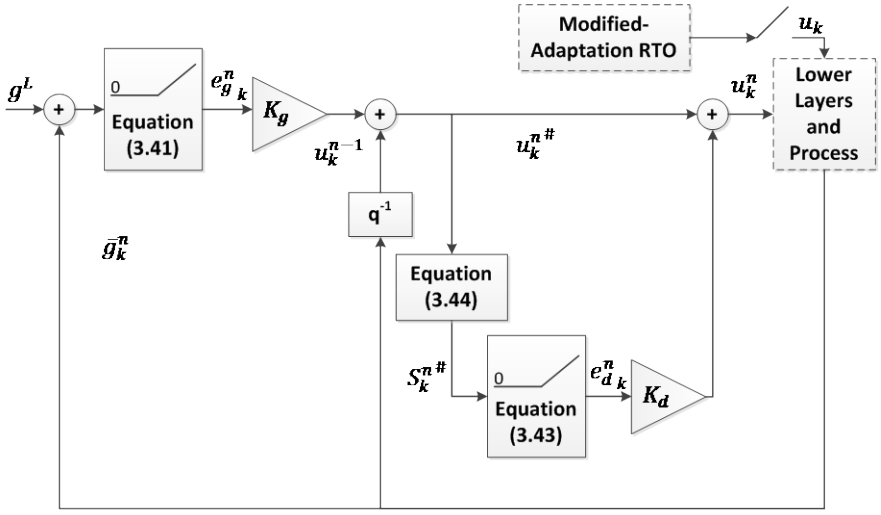

FIGURE 3.8: PRIMAL-DUAL INFEASIBILITY CONTROLLER

3.5.2 The Infeasibility Controller Implemented in the Oxygen Consumption Reactor

To test the performance of the infeasibility controller in it has been applied to the oxygen consumption reactor from Figure 3.4.

As it was commented, the best way to apply the modifier adaptation methodology in this system is using the dual control optimization approach.

This method allows converging to the real optimum of the process by means of estimating the process gradient using an approximation of the directional derivatives. However, the tradeoff between the number of iterations and the straightness of the evolution of the system, can produce additional infeasible points with respect to the ones that would be obtained if finite differences method is applied, because of the need to maintain a constantly excited system.

Figure 3.9 reproduces in more detail Figure 3.7, showing the progress of the dual modifier adaptation highlighting the infeasibilities with dashed circles. Figure 3.10 presents a close up of the evolution near to the optimum of the real process, emphasizing again the infeasibilities.

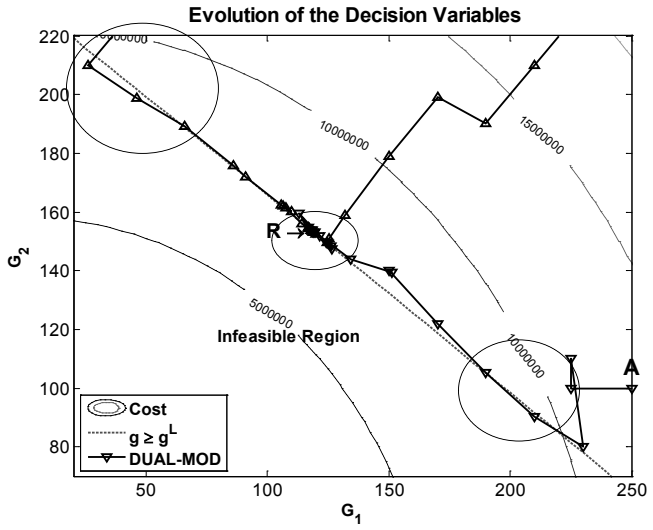

FIGURE 3.9: EVOLUTION OF THE DUAL-MODIFIER ADAPTATION WITH INFEASIBILITIES

These infeasibilities can be generated either because of an incorrect approximation of the inequality constraint or due to the need of maintaining an adequate level of excitation in the system in order to estimate accurately the process gradients. However, if we compare the evolution of the dual methodology with the one obtained with finite

differences (Figure 3.6), it can be noted that the correction in the inequality constraint is good enough to detect the active constraint in the feasible region; therefore, we can suspect that the infeasibilities are produced because of the dual constraint (equation (3.25)).

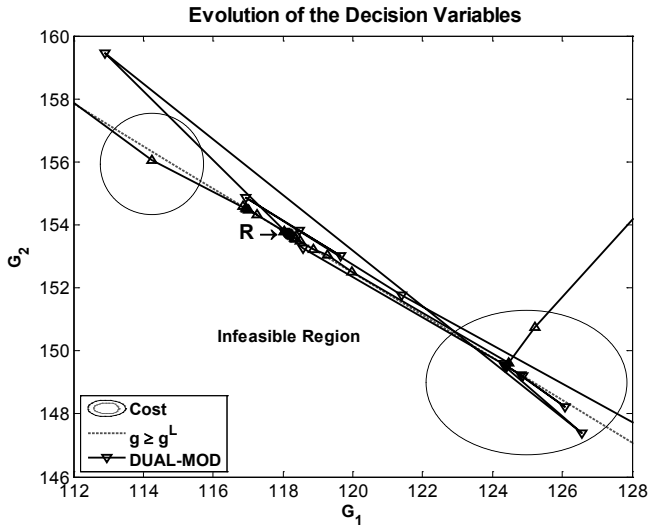

FIGURE 3.10: CLOSE-UP OF THE EVOLUTION OF THE DUAL-MODIFIER ADAPTATION WITH INFEASIBILITIES NEAR TO THE REAL OPTIMUM

As it was mentioned, decreasing the constant of the filter is recommended to make the evolution of the method as smooth as possible, reducing the aggressiveness in the update of the model. Nevertheless, the dual constraint is not affected directly by the modifiers since it only restricts the decision variables, hence no important changes in the infeasibility occurrence might be expected. Taking this into account, we propose handling the infeasibilities produced by applying the infeasibility controller from Figure 3.8.

Regarding the implementation, we can note that since the optimization problem has two decision variables and two constraints: the primal with the oxygen concentration and the dual with the inverse of the condition

number, there are enough degrees of freedom to apply completely the controller proposed. The matching between the controlled and the manipulated variables has been done giving priority to the primal constraint. This means that the oxygen concentration has been matched with the high purity stream, owing to the fact that the control action in this controlled variable must be done with less effort, implying that if an infeasibility is detected the deviation with respect to the value given by the RTO layer will be as minimum as possible to keep the constraint active. This matching implies that the dual constraint has to be controlled with the low purity stream.

Figure 3.11 and 3.12 show the optimal solution given by the dual modifier-adaptation layer and the real value applied to the process after being corrected by the infeasibility controller, denoted with "RTO" and "REAL" respectively. The difference among these values is given by the corrections made during the transient between two consecutive iterations: the path named "RTO" corresponds to the u_k decision variables from Figure 3.8, while the manifold with the name "REAL" is representing u_k^m, being m the total number of sampling periods needed to reach the steady state from the k^{th} to the $k^{th}+1$ RTO iteration.

From Figure 3.11 and 3.12 it can be noted how the system reaches the optimum of the process in an analogous way than the original dual methodology but without infeasible paths. If we focus our attention in the track that starts from C in Figure 3.11, it is observed that the evolution of the method proposed is equal to the dual methodology with no corrections, since there have not been detected infeasibilities in the primal constraint. This situation is expectable taking into account that the controller has been inactive and it was pointed previously that if this situation happens the proposed controller maintains the same value than the one given by the optimizer. On the other hand, the original paths starting from A and B both present an infeasibility occurrence at the moment when the primal constraint is detected. From Figure 3.11 it can be noted how the application

of the corrections given by the proposed controllers are sufficient to bring back the system to the feasible region. Near the optimum it can be noted from Figure 3.12 how the real time optimization layer proposes infeasible points. However, again the controller is able to manage the occurrence of these points. Therefore we can say that the infeasibilities in the oxygen concentration have been avoided completely during the whole evolution of the process. Regarding the number of iterations needed to reach the optimum, we can say that they have increased in a factor closer to 10% with respect to the original situation, concentrating this growth in the region closer to the optimum. This can be explained noting that near the convergence point, the feasible region where the controller can bring the system is very limited, therefore, in order to keep the system excited the controller must reduce the distance between two iterations. This idea is reflected in the diagram from Figure 3.13, where it has been represented the dual and the primal constraint for a general process with two decision variables.

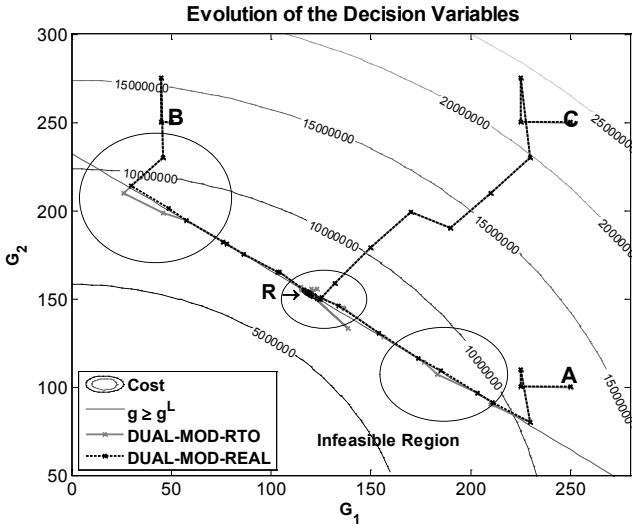

FIGURE 3.11: EVOLUTION OF THE DUAL MODIFIER ADAPTATION WITH THE INFEASIBILITY CONTROLLER

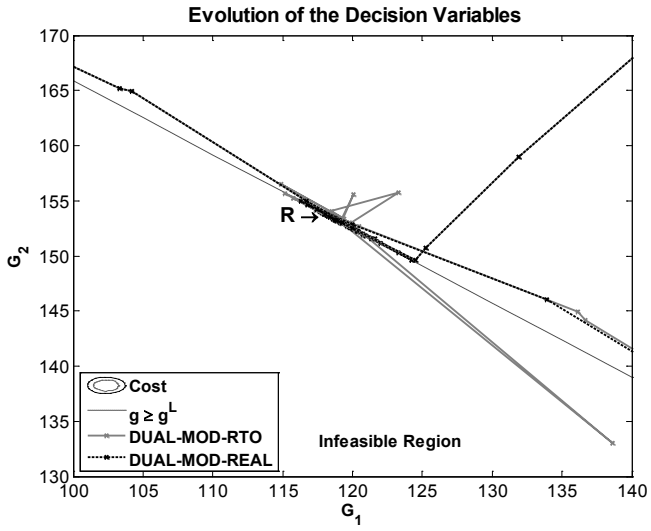

FIGURE 3.12: CLOSE-UP OF THE EVOLUTION OF THE DUAL-MODIFIER ADAPTATION WITH THE INFEASIBILITY CONTROLLER NEAR TO THE REAL OPTIMUM

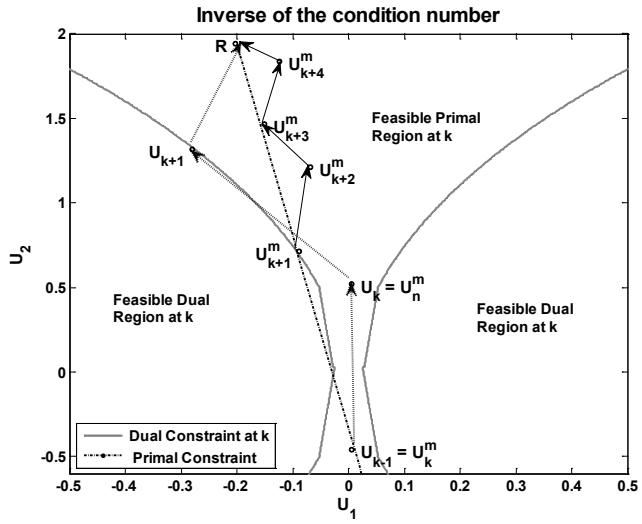

FIGURE 3.13: REPRESENTATION OF THE ACTION OF THE INFEASIBILITY CONTROLLER IN THE REGION NEAR TO THE REAL OPTIMUM. DASHED ARROWS SHOW THE EVOLUTION OF THE RTO LAYER, WHILE SOLID ARROWS REPRESENT THE ACTION OF THE INFEASIBILITY CONTROLLER

Assuming that the outcome of the RTO layer at k iteration produces the infeasible point u_{k+1}, and there is enough control action to bring the system into the feasible region, the infeasibility controller will change the location of the decision variable applying u_{k+1}^{m}, as is represented in Figure 3.13. However, this action implies that the next point in the RTO iteration would be farther than u_k with respect to the real optimum. Applying the same idea to the next RTO points, it will be an increase in the number of total iterations. Even though the number of iterations has been enlarged, the system is still able to find the real optimum of the process because of the dual controller that allows estimating the gradient in an accurate way, detecting the direction where the real optimum is.

3.5.3 Remarks about the Infeasibility Controller

The implementation of the proposed method shows that it is possible to handle the infeasibilities generated during the dual modifier-adaptation evolution. This however, is valid if and only if there is sufficient control action available to shift the controlled constraints into the feasible region, i.e., there is not saturation on the manipulated variable. This can be managed choosing the pairing between the manipulated and the controlled variables with an adequate criterion, such as the steady-state gain.

The approach proposed in this section presents some characteristics of the *self-optimizing control* methodology, proposed by the group of Professor S. Skogestad(Halvorsen et al., 2003; Skogestad, 2000a, 2000b). This method is based in solving an offline static optimization in order to identify the necessary conditions of optimality of the model, such as, active constraints and zeros gradients, and then generate controllers in charge of keeping these variables in their nominal values in spite of the disturbances. The authors also comment the possibility to run successive online RTO's with less frequency in order to evaluate if the NCO remains unchanged. In our approach, we assume that the infeasibility is produced due to the fact that the optimization layer has detected that the real optimum is closer to a

constraint (or is on the active constraint), but because of an inaccurate modification of the model and/or the excitement condition imposed by the dual constraint, the outcome of the upper layer is not in the feasible region. Then, controlling the constraints is equivalent to keep the complementary slackness conditions of the NCO of the process approximation in their nominal values. At the same time, if there are remaining degrees of freedom that have not been used for control, the directions of the zero gradients are given by the unchanged decision variables from the upper optimization layer. Therefore, besides looking for feasibility, we also can be looking for a point close to optimality around the operating point. The detection of the real optimum of the process in the oxygen consumption reactor can give a clue about this idea, since it finds the real optimum of the process in a feasible path, and with an increase in the number of RTO iterations that can be understood as the tradeoff between feasibility and optimality.

3.6 Modifier-Adaptation Methodology as a Nested Optimization Problem

Coming back to the modifier-based methods, there is an important assumption about their applicability and convergence properties: the capacity to estimate the experimental gradients. The entire modifier methods, from the ISOPE algorithm to the modifier-adaptation methodology present this assumption as a fundamental base to find the real optimum of the process.

As it was pointed out and tested with an example in section 3.3.3, the method to evaluate the gradient can modify the path of convergence. Nevertheless, no matter the chosen method the theory indicates that it will be able to find the optimum of the process under modeling mismatch assumption. Even though the idea is very promising in terms of the requirements of the model used in the optimization, i.e., model adequacy, model accuracy and set of active constraints, in terms of the real

applications, this might be a problem since the gradient can be: (1) expensive to obtain, for example in interconnected processes where the perturbations needed to calculate the gradients can affect other units, or (2) inaccurate when we are dealing with noisy process measurements. In the first case, we can try to perturb as less as possible the system in order to decrease the degree of excitation of the process; nonetheless, there is a tradeoff between this point and the quality of the estimations which becomes critical if we take into account the noise that can be present in the measurements.

As a response of this fact, one can say that the dual methodology can be used to estimate the process derivative, with a significantly lower number of additional perturbations; however, the addition of the dual constraint may force to compute a policy that guarantees a good estimation of the process gradients, changing the direction of the iterations as well as increasing its number, which, as it was pointed in previous section, can produce infeasible points.

Consequently, we can locate the applicability of the modifier-adaptation method to systems where the gradient is not expensive to obtain and the influence of the noise in the measurements can be handled correctly to obtain an accurate estimation of the gradient.

In this section we present a different way to understand the modifier-based optimization as a nested optimization problem: *the Nested modifier-Adaptation Methodology*. This method allows working with no information of the process gradient if it is not available.

3.6.1 The Nested Modifier-Adaptation

Notice that the iterative implementation of the modifier-adaptation methodology from Figure 3.3 can be summarized as follows: for a given value of the modifiers, obtained with any of the methods available to calculate the gradient of the process, an optimization problem is solved

using equation (3.11) in order to calculate and apply the next operation point to the process. Once the process has reached the steady state, apply the method to re-evaluate the gradient of the process, and with this value compute the modifiers in order to solve again the modified problem from equation (3.11), repeating this until no further changes in the decision variables are observed.

Because of the convergence assumption of the modifier-adaptation methodology states that a null change in the decision variables implies that the optimum of the process has been found, the iterative implementation already described can be viewed as a continuous update of the modifiers using a gradient based criterion, with the final goal to obtain the real optimum of the process.

In principle, any policy for updating the modifiers could be used, provided that improve the proposed optimum and respect the process constraint. With this idea in mind, one can iterate with the modifiers over the modified optimization until the optimum of the process is found, replacing the gradient estimation and the modifier calculus steps by any other method that takes into account the minimization of the cost function measured directly from the process. In particular, it is possible to implement an upper (or outer) optimization layer that uses the modifiers as decision variables to be applied over the inner modified optimization from equation (3.12) in order to obtain the decision variables to be applied into the process, and uses the cost function obtained from the process as the objective function of the upper layer.

The purpose of the upper optimization layer is to obtain the optimum of the process iterating with the modifiers. When selecting the optimization method, we can take into account that their decision variables, the gradient modifiers are not constrained, so that we can implement any unconstrained method. In particular a gradient-free algorithm can be used in this step, avoiding the need of estimating process gradients

This is the basic idea of the nested-modifier adaptation methodology presented in this section and summarized in the following algorithm:

> ***STEP 0:*** Set $\mathrm{k} = 0$ and start the algorithm with an initial guess of the modifiers α_0, γ_0 and ϵ_0.
>
> ***STEP 1:*** Once the process steady state is reached, measure the value of the process cost function $\bar{\mathrm{f}}_\mathrm{k}$ and the process constraint $\bar{\mathrm{g}}_\mathrm{k}$. Compute $\epsilon_\mathrm{k} = \bar{\mathrm{g}}_\mathrm{k} - \mathrm{g}_\mathrm{k}$.
>
> ***STEP 2:*** Update the modifiers λ_k and γ_k by using the value of $\bar{\mathrm{f}}_\mathrm{k}$ and the optimization algorithm of the upper layer. Check convergence of the upper optimization layer. If the process optimum has been found stop, if not go to step 3 (optional) or step 4.
>
> ***STEP 3 (OPTIONAL):*** Filter the modifiers given by the upper optimization layer using the equation (3.15).
>
> ***STEP 4:*** Calculate the decision variable u_k, solving the inner modified optimization problem from equation (3.11), using the modifiers λ_k and γ_k given by the upper optimization layer (which could be filtered from step 3), and the value of ϵ_k calculated directly from the available measurements of $\bar{\mathrm{g}}_{\mathrm{k}-1}$.
>
> ***STEP 5:*** Apply the decision variable u_k into the process and wait until the next steady state. Go to step 1.

It can be noted that the proposed methodology only uses the modifiers λ and γ as the decision variables, and the cost function of the process as the objective function for the upper optimization layer. The value of ϵ_k, on the other hand, is calculated in the same way than in the original modifier adaptation methodology because it can be obtained directly from the

measurements. Notice that the process constraints are not considered in the outer optimization, but in the inner one, that uses u as a decision variables.

Figure 3.14 summarizes the nested modifier-adaptation methodology described, where one can see the analogy in the implementation with respect to the original modifier-adaptation algorithm, where the process gradient estimation step has been substituted by an upper optimization block.

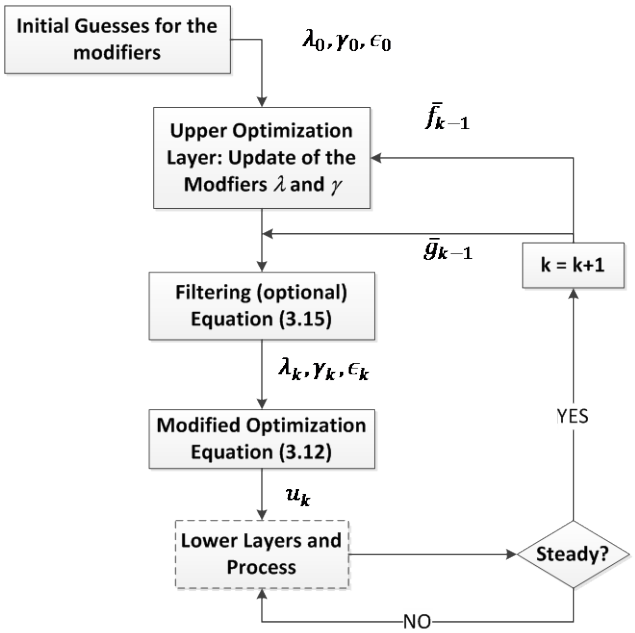

FIGURE 3.14: DIAGRAM OF THE NESTED MODIFIER-ADAPTATION METHOD PROPOSED

The termination criteria of the nested optimization is given by the convergence of the upper layer that in general, for unconstrained problems, are the relative changes of the cost function or the decision variables to be applied. These ones are equivalent to the original method, since in the modifier-adaptation methodology no changes in u_k implies that the process optimum has been found.

Regarding the characteristics of the optimum found with the proposed methodology, one can say that only local optimum can be guaranteed for a non-convex problem, in the same way than the original approach.

The upper optimization layer is continuously iterating with the modifiers with the aim of finding the optimum of the process. This implies that at every iteration of the algorithm implemented it is necessary to solve the inner modified optimization and then apply its solution to the process to obtain the value of the real cost function. If a gradient-based algorithm were implemented in the upper layer, it would be necessary to apply additional perturbations into the real system to estimate the gradient of the measured cost function, in a similar way than the original modifier-adaptation methodology does. However, if we use a gradient-free algorithm in the upper layer, there is no need to estimate the gradient of the process in order to look for the real optimum. Hence, the nested modifier-adaptation methodology can be seen as a generalization of the previous method, because the new architecture allows working with: (a) previous gradient-based method, (b) its alternative based in a gradient-free algorithm and (c) a combination of both options. This way to see the modified adaptation methodology can be very useful, since it allows implementing it when process gradients are not available and/or are difficult to obtain. Accordingly, the Nelder – Mead (NM) algorithm has been chosen for this task.

The NM algorithm works with the idea of finding the optimum of the process exploring the cost surface by means of a geometric figure with $n_\lambda + n_\gamma + 1$ vertices: the simplex. Each vertex corresponds to a set of values of the decision variables and is associated with its corresponding value of the cost function. Then, with 4 basic operations: reflection, expansion, contraction and shrinking, the algorithm iterates with the set of decision variables looking for the optimum, as Figure 3.15 shows. The main reason to choose the NM algorithm is because it is particularly parsimonious in function evaluations per iteration, since in practice it

typically requires only one or two function evaluations to construct a new iteration, while several popular direct search methods use $n_\lambda + n_\gamma$ or more function evaluations to obtain a new simplex (Lagarias et al., 1998). This property is very important considering that each function evaluation implies changing the operation point of the real process. The other reason why this method is chosen to update the modifiers is the popularity that it has in the chemical engineering field for nonlinear optimizations when the gradients are not available: experimental design, optimization of analytics tests, etc. (Walters et al., 1991).

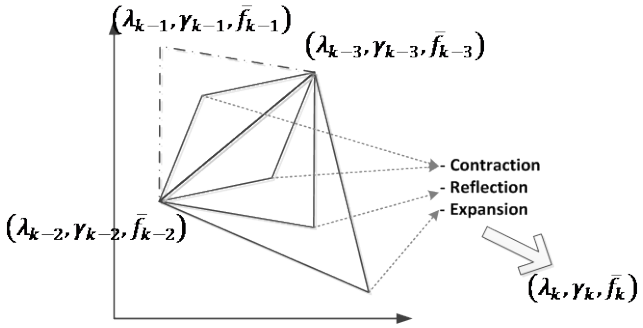

FIGURE 3.15: DIAGRAM OF THE SIMPLEX IN THE UPPER OPTIMIZATION

Regarding the initialization of the outer optimization, the easiest choice is $\lambda_0 = \gamma_0 = 0$, which corresponds to the nominal solution of the model without considering uncertainty. However, any other can be applied.

Besides the suppression of the gradient estimation step, using a direct search algorithm instead a gradient-based method, allows to obtain better results in noisy environments (Walters et al., 1991), making the entire algorithm more robust to real process conditions. Moreover, another advantage of the proposed method is the fact that one of the most sensitive parameters to tune in modifier-adaptation methodology based in the calculus of the gradients of the process is neglected: the size of the perturbation to estimate the gradient, translated in the dual methodology

as the specification of δ^L, a number that is not easy to select a priori and, as we will see in the implementation examples presented in section 3.6.3, very sensitive to the path of convergence to the real optimum of the process. Therefore, the removal of the gradient estimation step also implies a simpler way to apply the modifier approach compared with the original one.

3.6.2 Formalization of the Nested Modifier-Adaptation Methodology

The idea of the nested procedure is quite intuitive since if the system reaches a stationary point, the feasibility of the process is ensured by the definition of ϵ. Therefore, if the process cost function cannot be improved, it means that the real optimum of the constrained process (local if the problem is not convex) has been found. In this section a formal convergence proof is proposed .

Starting from the optimality conditions of the modified problem (where the dependencies of f and g have been omitted to simplify the notation):

$$\nabla_u f_k + \lambda_k^T + \mu_k^T(\nabla_u g_k + \gamma_k^T) = 0 \tag{3.46}$$

$$\mu^T(g + \gamma_k^T(u - u_{k-1}) + \epsilon_k) = 0 \tag{3.47}$$

$$\mu \geq 0, \qquad g + \gamma_k^T(u - u_{k-1}) + \epsilon_k \leq 0 \tag{3.48}$$

We can note that if a stationary point has been reached, for any value of λ and γ it is true that:

$$\begin{aligned} \epsilon &= \bar{g} - g \\ u &= u_{k-1} \end{aligned} \tag{3.49}$$

This implies that the primal and dual feasibility constraints (equation (3.48)), as well as the complementary slackness from equation (3.47), are fulfilled.

$$\mu^T \bar{g} = 0, \qquad \bar{g} \leq 0, \qquad \mu \geq 0 \tag{3.50}$$

Also, we can compute the gradient of the complementary slackness condition (equation (3.47)) with respect to the modifiers of the gradients.

$$\begin{aligned} \mu^T(\nabla_u g + \gamma_k^T)\nabla_\lambda u + (g + \gamma_k^T(u - u_{k-1}) + \epsilon_k)^T \nabla_\lambda \mu = 0 \\ \mu^T(\nabla_u g + \gamma_k^T)\nabla_\gamma u + (g + \gamma_k^T(u - u_{k-1}) + \epsilon_k)^T \nabla_\gamma \mu = 0 \end{aligned} \tag{3.51}$$

In the stationary point, independent of the value of λ and γ, condition (3.49) can be applied in equation (3.51):

$$\begin{aligned} \mu^T(\nabla_u g + \gamma^T)\nabla_\lambda u + \bar{g}^T \nabla_\lambda \mu = 0 \\ \mu^T(\nabla_u g + \gamma^T)\nabla_\gamma u + \bar{g}^T \nabla_\gamma \mu = 0 \end{aligned} \tag{3.52}$$

From equation (3.52) there are two possibilities for a given constraint $\bar{g}_i$:

$$\begin{aligned} \mu_i = 0 \rightarrow \bar{g}_i \nabla_u \mu = 0 \\ \bar{g}_i = 0 \rightarrow \bar{g}_i \nabla_u \mu = 0 \end{aligned} \tag{3.53}$$

Therefore, we can say that:

$$\bar{g}^T \nabla_u \mu = 0 \tag{3.54}$$

From the result of equation (3.54) and using the definition of ϵ and the chain rule, equation (3.52) can be rearranged as equation (3.55) shows.

$$\begin{aligned} \mu^T \nabla_\lambda \bar{g} + \mu^T(\gamma^T - \nabla_u \epsilon)\nabla_\lambda u = 0 \\ \mu^T \nabla_\gamma \bar{g} + \mu^T(\gamma^T - \nabla_u \epsilon)\nabla_\gamma u = 0 \end{aligned} \tag{3.55}$$

On the other hand, if equation (3.46) is multiplied by the gradient of u with respect to λ and γ we obtain:

$$\begin{aligned} \nabla_u f \nabla_\lambda \mu + \lambda^T \nabla_\lambda \mu + \mu^T(\nabla_u g + \gamma^T)\nabla_\lambda \mu = 0 \\ \nabla_u f \nabla_\gamma \mu + \lambda^T \nabla_\gamma \mu + \mu^T(\nabla_u g + \gamma^T)\nabla_\gamma \mu = 0 \end{aligned} \tag{3.56}$$

Combining equation (3.56) with equation (3.52) gives:

$$\begin{matrix} \nabla_u f \nabla_\lambda \mu + \lambda^T \nabla_\lambda \mu - \bar{g}^T \nabla_\lambda u = 0 \\ \nabla_u f \nabla_\gamma \mu + \lambda^T \nabla_\gamma \mu - \bar{g}^T \nabla_\gamma u = 0 \end{matrix} \tag{3.57}$$

Relation that can be simplified using the condition from equation (3.54):

$$\begin{matrix} \nabla_u f \nabla_\lambda \mu + \lambda^T \nabla_\lambda \mu = 0 \\ \nabla_u f \nabla_\gamma \mu + \lambda^T \nabla_\gamma \mu = 0 \end{matrix} \tag{3.58}$$

So far, we can say that equations (3.55) and (3.58) holds for any value of λ and γ, provided the iterative implementation of the method has reached a stationary point.

If we consider now the implementation of the upper optimization layer into the process, the necessary conditions of optimality of this system are:

$$\begin{matrix} \nabla_\lambda \bar{\mathcal{L}} = \nabla_\lambda \bar{f} + \mu^T \nabla_\lambda \bar{g} = 0 \\ \nabla_\gamma \bar{\mathcal{L}} = \nabla_\gamma \bar{f} + \mu^T \nabla_\gamma \bar{g} = 0 \end{matrix} \tag{3.59}$$

$$\mu^T \bar{g} = 0, \qquad \mu \geq 0, \qquad \bar{g} \leq 0 \tag{3.60}$$

Equation (3.60) is fulfilled for any stationary point of the nested optimization problem using the definition of ϵ, as it was mentioned previously. On the other hand, from equation (3.59) the derivative $\nabla_\lambda \bar{g}$ and $\nabla_\gamma \bar{g}$ must be zero because of the unconstrained characteristic of the upper layer. Replacing this condition in equation (3.55) we have:

$$\begin{matrix} (\gamma^T - \nabla_u \epsilon) \nabla_\lambda u = 0 \\ (\gamma^T - \nabla_u \epsilon) \nabla_\gamma u = 0 \end{matrix} \tag{3.61}$$

On the other hand, in the optimum of the nested problem $\nabla_\lambda \bar{f} = \nabla_\gamma \bar{f} = 0$, replacing this condition in equation (3.58) we have:

$$\begin{matrix}\left(\nabla_u f - \nabla_u \bar{f} + \lambda^T\right)\nabla_\lambda \mu = 0 \\ \left(\nabla_u f - \nabla_u \bar{f} + \lambda^T\right)\nabla_\gamma \mu = 0\end{matrix} \quad (3.62)$$

Provided $\nabla_\lambda u$ and $\nabla_\gamma u$ are full rank matrices, equation (3.61) and (3.62) implies that at the optimum of the nested optimization problem:

$$\gamma^T = \nabla_u \epsilon = \nabla_u \bar{g} - \nabla_u g \quad (3.63)$$

$$\lambda^T = \nabla_u \bar{f} - \nabla_u f \quad (3.64)$$

Equations (3.63) and (3.64) mean that at the optimum of the nested optimization problem, the modifiers are the ones given by the original modifier-adaptation methodology provided the gradients of u with respect to λ and γ are full rank matrices. Since we have assumed that a stationary point has been reached, we can say that the real optimum of the process coincides with the one obtained with the proposed method.

3.6.3 Examples of Implementation of the Nested Modifier-Adaptation Methodology

To illustrate the method, the proposed Nested algorithm was applied in several examples, using the *fminsearch* and the *fmincon* routines from the optimization toolbox of MATLAB(Mathworks, 2007), as optimization algorithms for the upper and the nested layers. The implementations have been compared with dual modifier-adaptation methodology for different degrees of excitation of the system. This is equivalent to modify the lower bound of the dual constraint δ^L with the aim to justify the elimination of this tuning parameter due to its sensitivity. The dual methodology was chosen to estimate the gradient of the process because of its comparative advantages already mentioned.

The first two examples: a convex optimization problem and a system with three interconnected tanks(Marchetti et al., 2009a) present parametric uncertainty. The aim of these applications is to compare the response of

the algorithm under absence and presence of noise in the process measurements. The next two examples: The Otto-Williams Reactor(Forbes et al., 1994; Williams & Otto, 1960) and a Continuous Bioreactor(Golden & Ydstie, 1989) shows the implementation of the modifier adaptation methodology under the presence of structural mismatch, being the application of the last reactor very difficult because of a change in the behavior of the system in the vicinity near the optimum of the process. Finally the last implementation is in a run-to-run batch reactor with structural uncertainty and process noise (Chachuat et al., 2009).

3.6.3.1 Convex Optimization

This example is a simulated process optimization extracted from the work of Marchetti and co-workers(Marchetti et al., 2009a) and it was used to introduce the modifier adaptation methodology. In this case, the difference between the process and the model is the value of four uncertain parameters which are not updated during the iterations of the RTO algorithm.

Consider the convex optimization problem from equation (3.65), where u are the decision variables, θ represent the parameters and g is a single inequality constraint.

$$\begin{aligned} &\min_{u} f \coloneqq (u_1 - \theta_1)^2 + 4(u_2 - 2.5)^2 \\ &s.t.: \\ &g \coloneqq (u_1 - \theta_1)^2 - 4(u_2 - \theta_4)\theta_3 - 1 \leq 0 \\ &u: [u_1 \quad u_2]^T, \qquad u \in [0, \infty) \end{aligned} \tag{3.65}$$

The target is to find the optimum of the process, represented as a simulated reality with a given set of correct parameters, using the nested modifier-adaptation methodology applied into a model that has a set of erroneous parameters (corresponding to a simulated modeling mismatch). Table 3.4 summarizes the values of the parameters considered.

TABLE 3.4 VALUE OF THE PARAMETERS USED IN THE PROCESS AND THE MODEL FOR THE CONVEX OPTIMIZATION EXAMPLE

System	θ_1	θ_2	θ_3	θ_4
Simulated Process	3.5	2.5	-0.4	1.0
Model	2	1.5	-0.5	0.5

Because of the cost and the constraint are affected with the modeling mismatch, the three modifiers λ, γ and ϵ from equation (3.11) can be applied in this example. The evolution of the dual and the nested modified-adaptation method will be compared for different strategies of adaptation of these modifiers: correction with the three modifiers, corrections only in the constraint and correction in the gradient of the objective function and in the bias of the constraint, following the same application presented in the original paper.

Regarding the implementation of both methods, the starting point was the optimum of the model. From this initial point, the first two iterations were used to estimate the gradients with the finite difference approach, because dual methodology needs at least two previous operation points to calculate the gradients as equation (3.24) shows. In the case of the nested methodology, this gradient was employed as a good initial guess of the decision variables for the NM algorithm, in order to start both in the same conditions.

Figures 3.16 and 3.17 show the evolution of the dual methodology for two different values of δ^L. In these figures there are three graphs corresponding to three different sets of values of K_λ, K_γ, and $K_{\epsilon,}$ which have been changed in order to test their influence in the detection of the real optimum of the process. In these Figures the points **M** and **R** represent the optimum of the model and the process respectively.

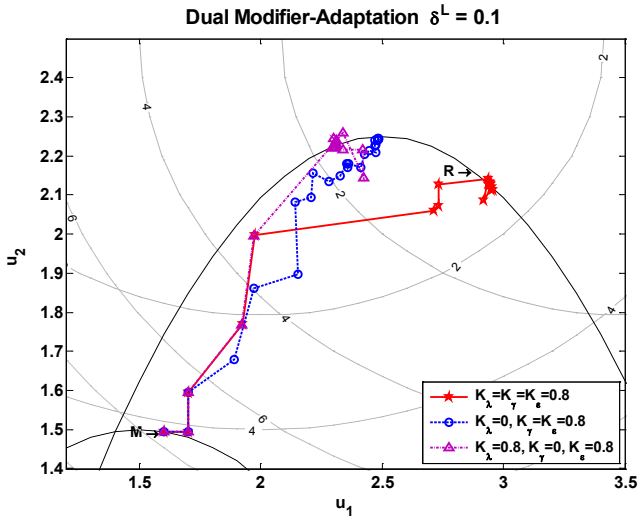

FIGURE 3.16: EVOLUTION OF THE DUAL METHODOLOGY WITH $\delta^L = 0.1$. GREY AND BLACK THIN LINES ARE THE COST AND THE CONSTRAINT OF THE PROCESS (SOLID) AND THE MODEL (DASHED)

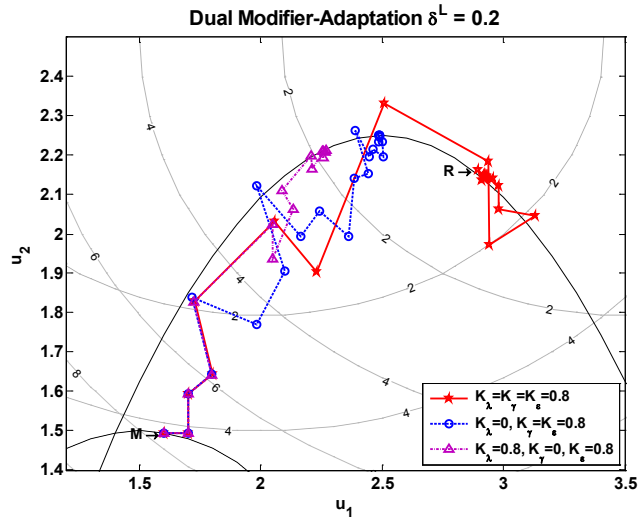

FIGURE 3.17: EVOLUTION OF THE DUAL METHODOLOGY WITH $\delta^L = 0.2$. GREY AND BLACK THIN LINES ARE THE COST AND THE CONSTRAINT OF THE PROCESS (SOLID) AND THE MODEL (DASHED)

It can be noted from Figures 3.16 and 3.17, how the use of different strategies of adaptation makes the iterative algorithm converge not to the same values, being the full adaptation policy in red (the use of the three modifiers), the only one that converges to the real optimum of the process for both values of δ^L tested. These results completely agree with the ones given previously in literature. Nevertheless, the authors do not mention the method employed to estimate the gradient of the simulated process.

The sensitivity of the dual methodology with respect to the value of δ^L can be observed comparing both figures. It can be seen how the paths obtained with $\delta^L = 0.1$ converges in 6 iterations into a neighbourhood close to the real optimum, detecting the real active constraint. On the other hand, the evolution with $\delta^L = 0.2$ takes 9 iterations to detect the active constraints, nevertheless, the convergence path is infeasible since it crosses the constraint $\bar{g}$. This behaviour indicates that the lower the value of the dual constraint, the more direct way to reach the optimum, but notice that this implies a potential worse estimation of the process gradient. Conversely, the choice of a larger value of this bound can generate infeasible points in the RTO path as can be observed. The effect in the feasible region of the dual constraint has been studied in the work of Marchetti and coworkers(Marchetti et al., 2010). For the two dimensional case, the feasible region of the dual constraint consists of two discs with the same radius centered at the same distance from the centroid of the previous two decision variables, in an orthogonal direction to the line defined with these two last values (Figure 3.18).

The size of the discs are inversely proportional with respect to δ^L, which implies that the operation points must be inside a smaller disc while the greater is the bound of the inverse of the condition number. This is translated in to a system with more energy to estimate the gradient, but with an evolution of the dual modifier adaptation that is continuously changing its direction, configuring a less direct way to reach the optimum. Because the modifiers γ and ϵ allows to estimate the real behaviour of the

process constraint only around the actual operating point (Gao & Engell, 2005), these changes on the direction of the evolution can produce a violation of the process constraint in the region where the approximation is no longer valid.

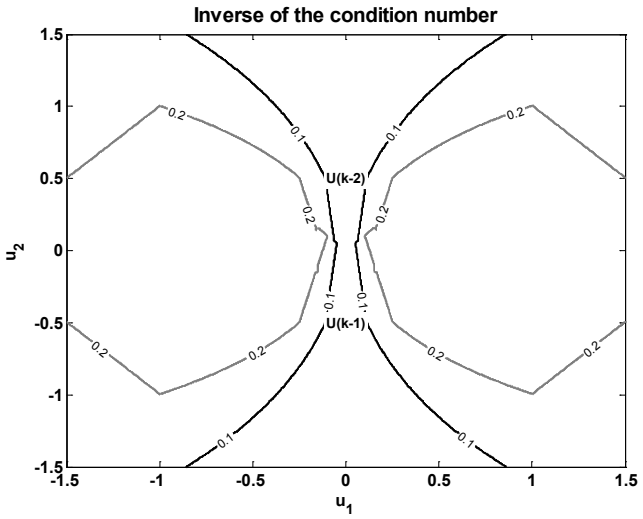

FIGURE 3.18: FEASIBLE REGION OF THE DUAL CONSTRAINT FOR DIFFERENT VALUES OF δ^L

Figure 3.19 shows the same comparison of the adaptation strategies of Figures 3.16 and 3.17 using the nested methodology proposed. Similar to the original dual method, the only policy that reaches the real optimum of the process is the full modifier-adaptation: when all the modifiers are taken into account, while the rest of the adaptation approaches get stuck in a suboptimal but feasible operation point. But using this method, the evolution of the solutions do not present the zigzag behavior of Figure 3.17 and the algorithm converges in a direct path in to a neighborhood of the process optimum, using 5 iterations to detect the active constraint, whereas the rest of the iterations attempts to refine the search of the real optimum. The evolution in the nested optimization problem is similar to the dual modifier-adaptation methodology with smaller values in δ^L, but without the need to select this parameter.

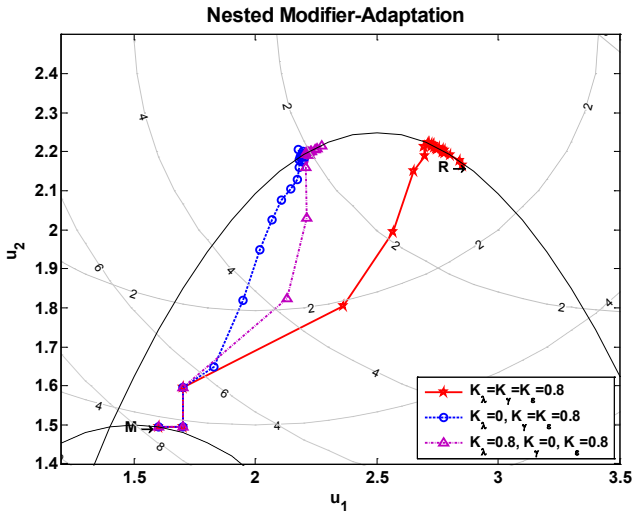

FIGURE 3.19: EVOLUTION OF THE NESTED METHODOLOGY. GREY AND BLACK THIN LINES ARE THE COST AND THE CONSTRAINT OF THE PROCESS (SOLID) AND THE MODEL (DASHED)

The reason why the nested methodology presents a more direct way to estimate the process optimum in this example is because of the dual constraint has been removed. This can be done since the search criteria is no longer the gradient of the process, but is a direct search, implying that in every iteration the algorithm is only looking for the optimum of the process, no taking care about the degree of excitation of the system.

In dual procedure, there exists an important trade-off in the choice of δ^L. As it was told previously, a smaller value of this parameter implies a more direct evolution in the solutions. However, a lower parameter means that the gradient can be poorly estimated and also the noise in the measurements can be amplified. To test the sensitivity of δ^L with respect to the noise, an additive error was simulated in the cost function and the constraint of the process. The stochastic simulated noise presents a uniform probability distribution centered in zero with amplitude of $\pm 5\%$ with respect to its expected range.

The same values of δ^L used in the noise-free scenario were employed to iterate with the dual modifier adaptation, while the finish tolerance was increased in one order of magnitude, as Brdys and coworkers recommend for the dual ISOPE methodology (M. A. Brdys & Tatjewski, 2005).

The results of dual and nested methodology, applying 15 trials for each method tested under stochastic noise conditions, are presented from Figure 3.20 to Figure 3.22.

The evolution of dual modifier- adaptation under noisy conditions, shows how the wrong election of δ^Lcan make this method more sensitive to the noise stopping convergence because of a gradient estimation problem. From Figure 3.20 it can be noted that it presents similar performance than the no noise conditions for the first 3 iterations of the dual estimation of the gradient, after this, when the solution approaches the optimum, the algorithm diverges. This can be explained because starting from a value that is far away from the optimum of the process implies significant changes in the operation points, which can be translated as an inactive dual constraint, hence a well estimate gradient at the next iteration. As the system gets closer to the process optimum, the next operating point might be close enough to the previous values which can activate the dual constraint, generating an S_k matrix more sensitive to the noise. On the other hand, the path from Figure 3.18 does not present a different behavior closer to the optimum, meaning that for a large enough value in the lower bound of the dual constraint, the algorithm converges into a neighborhood close to the real optimum. Nevertheless, the election of a big value in δ^L is not free, because it causes infeasibilities in the evolution of the algorithm in the same way than the noise free scenario. These problems are avoided with the proposed nested methodology, as the excitation requirement of the system disappear and the feasible region remains unchanged with respect to the original optimization problem, i.e., the dual constraint has been removed.

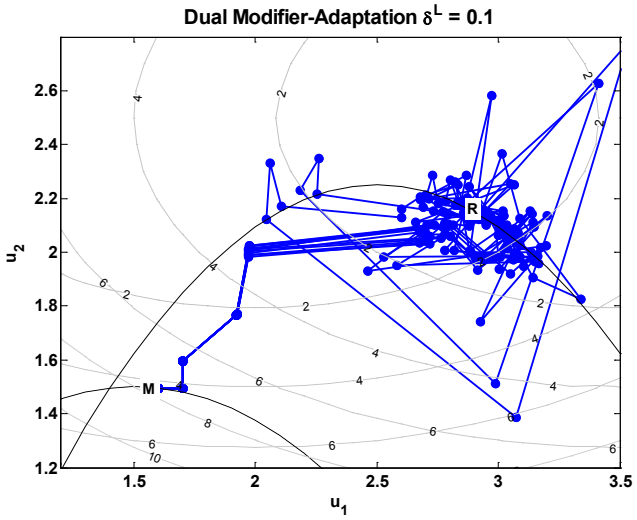

FIGURE 3.20: EVOLUTION OF THE DUAL METHODOLOGY WITH $\delta^L = 0.1$ UNDER NOISE CONDITIONS. GREY AND BLACK THIN LINES ARE THE COST AND THE CONSTRAINT OF THE PROCESS (SOLID) AND THE MODEL (DASHED)

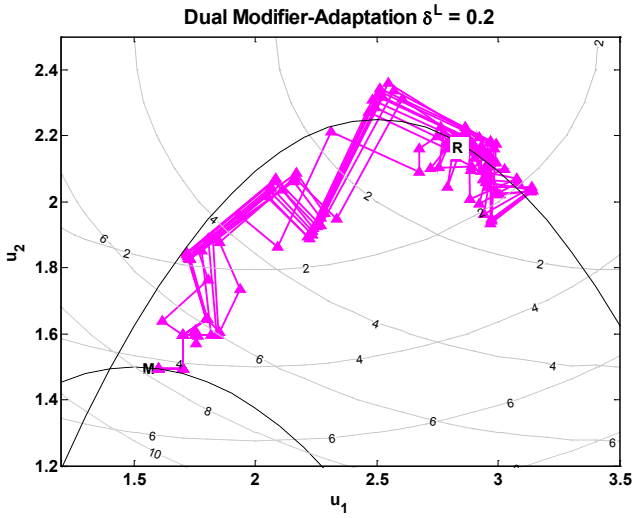

FIGURE 3.21: EVOLUTION OF THE DUAL METHODOLOGY WITH $\delta^L = 0.2$ UNDER NOISE CONDITION. GREY AND BLACK THIN LINES ARE THE COST AND THE CONSTRAINT OF THE PROCESS (SOLID) AND THE MODEL (DASHED)

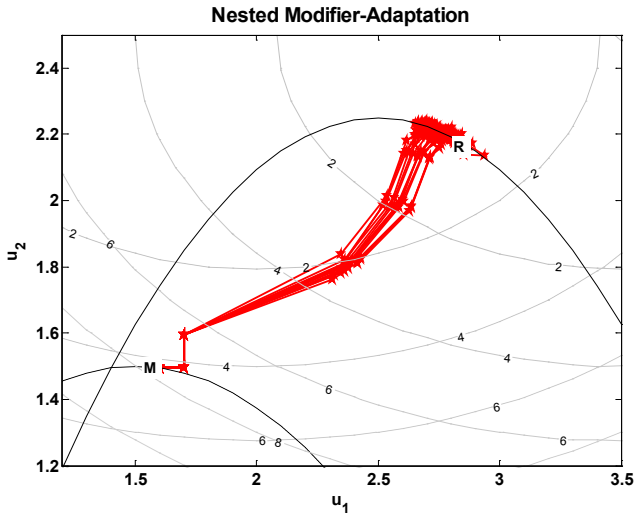

FIGURE 3.22: EVOLUTION OF THE NESTED METHODOLOGY UNDER NOISE CONDITIONS. GREY AND BLACK THIN LINES ARE THE COST AND THE CONSTRAINT OF THE PROCESS (SOLID) AND THE MODEL (DASHED)

Comparing the evolution under noise conditions of the nested methodology with the dual one, it is clear that the convergence is less sensitive to the measurement error. This is because the way in which the algorithm reaches the real optimum is quite similar to the case when there is not noise, i.e. 5 iterations to detect the active constraint and the rest to locate the optimum of the process.

As a remark in the implementation of the nested methodology in this convex optimization example, we can say that the optimum of the process is found in a similar way than the gradient based methodology, that is to say, when all the modifiers have been used, but with the advantage that no process gradient information is required. Also under noisy scenario, the adaptation strategy based in a direct search algorithm seems to be more robust than the previous gradient based method.

Regarding the infeasibilities that can be generated because of the dual constraint, it is important to mention that they can be handled with the infeasibility controller from section 3.5, nevertheless, the suppression of the gradient estimation step, allows converging in a feasible path tuning only the filter constants, because the dual constraint is not present in the nested modifier approach.

3.6.3.2 Three Interconnected Tanks

The second example used to test the proposed methodology is a system with three tanks (T_1, T_2 and T_3) from Figure 3.23. In this problem, the tanks are interconnected between them allowing the storage and the exchange of water. The complete system receives flows of water q_{p1} and q_{p2} from the pumps P_1 and P_2, while the outflows (q_1, q_2 and q_2) are the result of the fluid-dynamic potential between the surface and the bottom of each tank and the opening of the manual valves V_1, V_2 and V_3 respectively. Inside the tanks, the difference between the liquid levels (h_1, h_2 and h_3), will give the direction of the water that pass through the interconnections: q_{12} and q_{32}, whereas their flows also depend on the head loss produced by the valves V_{12} and V_{32}.

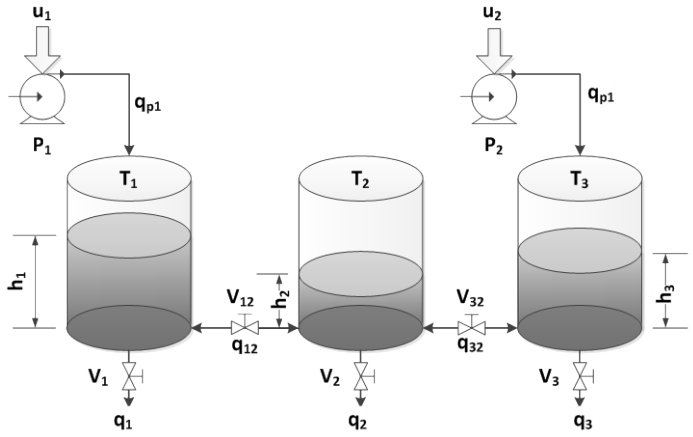

FIGURE 3.23: DIAGRAM OF THE THREE INTERCONNECTED TANKS.

This example was employed in the work of Marchetti and coworkers to test the modifier adaptation methodology in a real system (Marchetti et al., 2009a). Nevertheless, since the experimental setup is not available, the process and the model will be simulated in a similar way than the convex optimization example from section 3.6.3.1: using a model with different parameters for the process and the optimization.

A steady state model of the system can be derived using a first principle approach.

Mass balances in each tank:

$$\begin{aligned} q_{p1} - q_1 - q_{12} &= 0 \\ q_{12} + q_{32} - q_2 &= 0 \\ q_{p2} - q_3 - q_{32} &= 0 \end{aligned} \tag{3.66}$$

To describe the flow that passes through the valves it can be used the Torricelli's law that relates the flow with the liquid level and the hydrodynamic losses in the valves.

$$\begin{aligned} q_i &= Aa_i\sqrt{h_i}, \qquad i = 1,2,3 \\ q_{j2} &= sign(d_{j2})Aa_{j2}\sqrt{|d_{j2}|}, \qquad j = 1,3 \\ d_{j2} &\coloneqq h_j - h_2, \qquad j = 1,3 \end{aligned} \tag{3.67}$$

The operational goal of the system is keeping the liquid inside the tanks within a safe rage, consuming the lowest energy in the pumps. If the voltage that needs to be applied in each pump is denoted as u_1 and u_2 for P_1 and P_2 respectively, this operational task must be reached solving the optimization problem from equation (3.68). In this equation it is assumed that the water flow pumped is proportional to the signal to the pumps and the energy consumed is also proportional to the square of that signal, being w_j the proportional constant of the influent flow from the pumps.

$$\begin{aligned}
&\min_{u} f := u_1^2 + u_2^2 \\
&s.t.: \\
&steady - state\ model\ from\ equation\ (3.66) and\ (3.67) \\
&q_{pj} = w_j u_j, \qquad j = 1,2 \\
&u = [u_1 \quad u_2]^T, \qquad u \in [u^L, u^U] \\
&h_j \in [h^L, h^U]
\end{aligned} \tag{3.68}$$

A summary of the nomenclature used in the three tanks example is presented in Table 3.5

TABLE 3.5 NOMENCLATURE EMPLOYED IN THE THREE INTERCONNECTED TANKS EXAMPLE

Parameter	Meaning	Units
A	Transversal area of each tank	cm^2
a_i	Proportional constant of the valve V_i	$cm^{0.5}/min$
a_{j2}	Proportional constants of the valve V_{j2}	$cm^{0.5}/min$
d_{j2}	Difference on the liquids heights from tank T_j with respect to tank T_2	cm
h_i	Liquid height in tank T_i	cm
h^L, h^U	Bounds for h	cm
q_i	Effluent from the bottom of tank T_i	l/min
q_{p_j}	Influent water from pump P_j	l/min
q_{j2}	Flow interchanged between tanks T_j and tank T_2	l/min
u_j	Voltage applied to the pump P_j	V
w_j	Proportional constants of pump P_j	$l/minV$

The values of the parameters used from equation (3.66) to equation (3.68) are summarized in Table 3.6. They correspond to the original calibration coefficients obtained by Marchetti and coworkers, after an identification stage with the original experimental setup.

As it was mentioned previously, to simulate the modeling mismatch different values in the parameters for the model and the process for equation (3.67) must be used. The difference between the process and the model corresponds to a clogging situation in the valves V_{12} and V_{32}, which can be translated in a reduction on the flow coefficient:

$$a_{j2}^{M} = \frac{a_{j2}^{P}}{2}, \qquad j = 1,3 \tag{3.69}$$

In equation (3.69), the superscript P represents the value of the parameter used in the simulated process and is the same than the one presented in Table 3.6, whereas the parameter with subscript M denotes the one used in the model of the process with mismatch.

TABLE 3.6 VALUE OF THE PARAMETERS USED THREE INTERCONNECTED TANKS EXAMPLE

Parameter	Value	Parameter	Value
A	154	w_1	13.22
a_1	0.1203	w_2	14.96
a_2	0.0613	h^L	5
a_3	0.1141	h^U	30
a_{12}	0.0381	u^L	0
a_{32}	0.0285	u^U	8

Following the same procedure of the example from section 3.6.3.1, dual and nested modifier- adaptation methodologies were applied starting from the optimum of the model, while the first two iterations were used to estimate the gradient of the process using the finite difference approach. In this example, only the modifiers of the constraints of the liquid levels, i.e.: $\gamma$and ϵ where updated because the constraints that restrict the value of u_j as well as the cost function only depends on the decision variables and there is no uncertainty on them.

Figure 3.24 shows a comparison of the evolution of the decision variables u_1 and u_2 for the dual and the nested modifier-adaptation method with full modification (γ and ϵ). In the figure, the dual modifier-adaptation has been tested for different degrees of excitation of the system, which is reflected in an implementation with different values for δ^L. As in previous example, **M** and **R** highlights the location of the optimum of the model and the process respectively. As it is expected, it can be noted that the mismatch changes

the location of these two points, nevertheless both for the model and for the process lies on the active constraint $h_2 \geq h^L$. On the other hand, Figure 3.25 shows a comparison on the evolution of the objective function with the dual and the nested methods implemented.

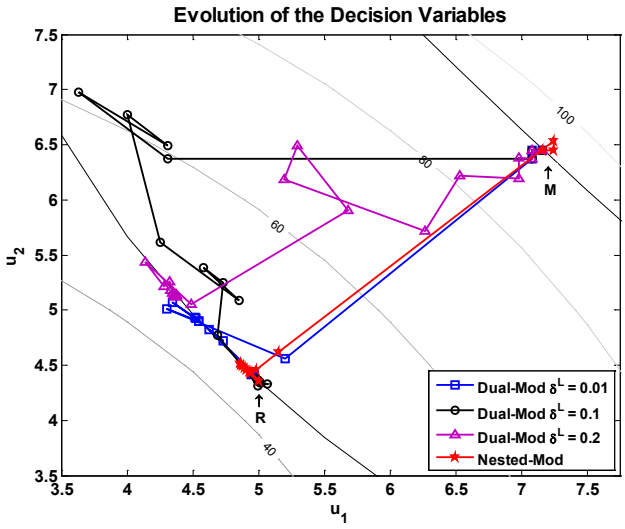

FIGURE 3.24: EVOLUTION OF THE DECISION VARIABLES FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE THREE TANKS EXAMPLE. GREY LINE IS DE COST FUNCTION, BLACK LINES ARE THE CONSTRAINT $h_2 \geq h^L$ FOR THE PROCESS(SOLID) AND MODEL (DASHED)

From Figures 3.24 and 3.25 it can be noted how the evolution of the decision variables of the dual methodology is affected by the value of the lower bound of the dual constraint, following different paths in the convergence to the real optimum. In a similar way than the example from section 3.6.3.1, a lower value of this parameter implies a more rapid convergence to the real optimum due to the fact that the use of larger values of δ^L reduce the feasible region of the dual constraint (Figure 3.18). The performance of the dual methodology can be observed more clearly in the evolution of the objective function (Figure 3.25), where the more tight constraint needs only 5 iterations to find a region close to the optimum (coinciding with the detection of the active constraint), while the other two

values of δ^L tested needs 11 and 15 iterations to detect this region. On the other hand, the evolution of the nested approach shows a convergence similar to the dual case with the lower value of δ^L, converging in 5 iterations into the closer region of the real optimum.

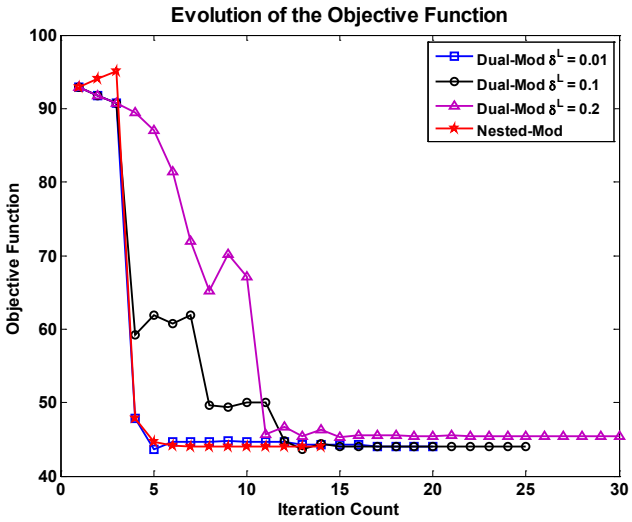

FIGURE 3.25: EVOLUTION OF THE OBJECTIVE FUNCTION FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE THREE TANKS EXAMPLE.

As we can see, in the unlikely situation of a process without noise, the convergence of the nested methodology is comparable with the dual algorithm with lower values in δ^L in terms of optimality. Nevertheless, as can be seen next, the trade-off between a larger feasible region and a system with low excitation levels gives an additional advantage to the use of the proposed methodology in terms of its robustness. To show this, an additive noise was applied in the measurements of the liquid levels of the process, using a uniform probability distribution function centered in zero with a range of 0.1 cm. Figure 3.26 and 3.27 shows the evolution of the decision variables and the cost function respectively, for the dual and the nested approach for 15 trials for every method under noisy condition.

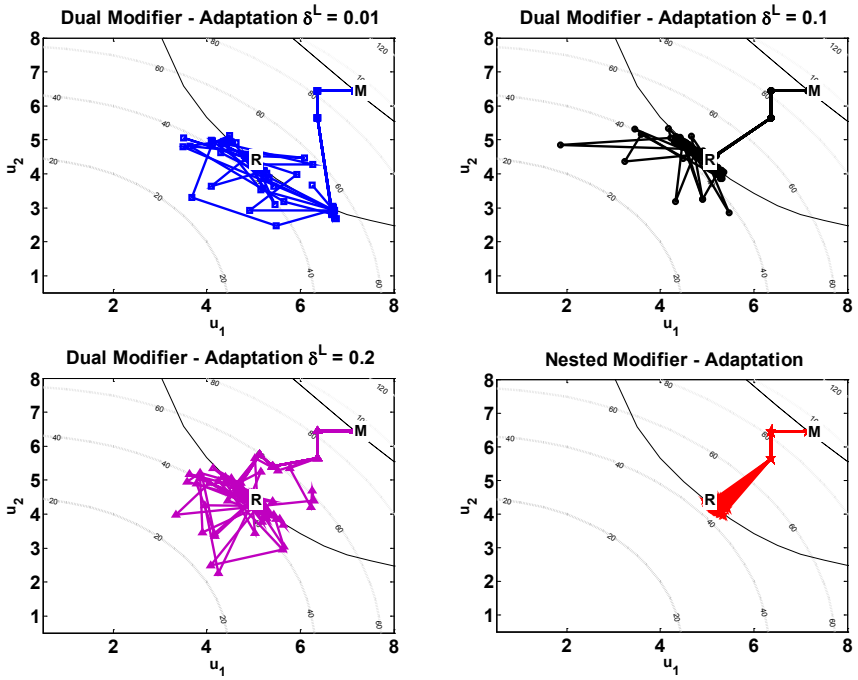

FIGURE 3.26: EVOLUTION OF THE DECISION VARIABLES FOR THE DUAL AND NESTED MODIFIER-ADAPTATION IN THE THREE TANKS EXAMPLE. GREY LINE IS DE COST FUNCTION, BLACK LINES ARE THE CONSTRAINT $h_2 \geq h^L$ FOR THE PROCESS (SOLID) AND MODEL (DASHED) , UNDER NOISY CONDITIONS

As it can be expected, the presence of noise affects the convergence to the real optimum of the gradient-based modifier methodology, being the effect on measurement disturbances more critical when a lower excitation in the system is set. This is equivalent to the poor performance observed for the lower value of δ^L tested. In fact, it can be seen from Figure 3.27 that the evolution of the dual methodology with $\delta^L = 0.01$ presents the largest dispersion in the evolution of the cost function of the process, and even the stagnation in different points for some trials tested. Increasing the value of this bound implies a better estimation of the process gradient which can be observed in the fact that the other two dual methodologies converges into the real optimum ($\delta^L = 0.1$ and 0.2). However, an increase on the energy

of the system implies the decrease of the size of the dual feasible region, meaning that if δ^L is overestimated it can be produced infeasible points, as we already mentioned and shown in the oxygen consumption example from section 3.5.2 and the convex optimization example from section 3.6.3.1, observing also this phenomenon in the three tanks example in the path formed with $\delta^L = 0.2$. Unlike the dual modifier – adaptation, Figure 3.26 shows again how the evolution of the nested methodology proposed in this work seems to be less sensitive to the measurement noise, converging in a similar way than the noise free scenario. For all the cases tested, it can be noted from Figure 3.27 just how the algorithm spends the same 5 first iterations to converge into a region closer to the optimum, corresponding to the detection of the active constraint in Figure 3.26, while the rest of the iterations are used to refine the search, as the same as the dual case with larger values of δ^L, but with the advantage of removing a non-convex dual constraint, keeping the original feasible region of the modified problem.

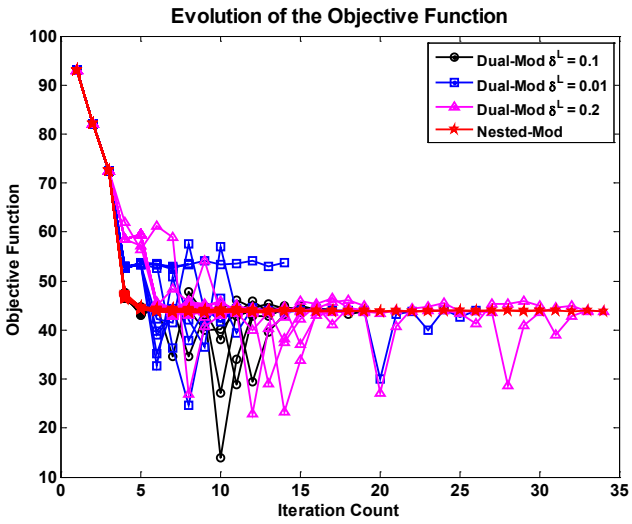

FIGURE 3.27: EVOLUTION OF THE OBJECTIVE FUNCTION FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE THREE TANKS EXAMPLE, UNDER NOISY CONDITIONS.

As a remark in the implementation of the nested methodology in the three tanks example, we can mention that for this example the gradient-free approach converges in a similar way than the previous modifier-adaptation scheme with a wide feasible region of the dual constraint, but with robustness comparable with a system that has enough excitation to estimate the gradient of the process adequately, even in the presence of process noise.

Since, a tighter dual feasible region is required to get over the process noise and ensure the convergence to the optimum of the process in real life conditions, the nested approach also helps reducing the occurrence of infeasibilities, since it has been neglected the constraint of persistent excitation. These infeasibilities could also be treated with the infeasibility controller from section 3.5; however, as we pointed in previous example, if the dual constraint is no longer needed, they can be handled tuning the filter constants in a correct way.

3.6.3.3 Otto-Williams Reactor

This example, tries to show and compare the behavior of the nested approach when the uncertainty is in the form of structural mismatch, that is when the difference between the model and the process includes its structure and not only a change in the value of some parameters.

The Otto Williams Reactor is a CSTR that has been used widely in the literature to study the performance of different RTO approaches with modeling mismatch. (Forbes et al., 1994; Marchetti, 2009; Roberts, 1979; Rodger, 2010; Yale Zhang & Fraser Forbes, 2000). The system consists of a continuous reactor that is fed with two sources of raw material A and B, by means of the streams F_A and F_B respectively. Inside the vessel, three parallels reactions take place forming 4 new compounds: C, E, G and P, as equation (3.70) shows. These compounds, along with the unused reactive, leave the reactor from the bottom of the vessel in a single stream F_R. Figure 3.28 schematizes the Otto-Williams Reactor, where X_i represents the mass

fraction of the i compound inside the reactor and T_R is the reactor temperature.

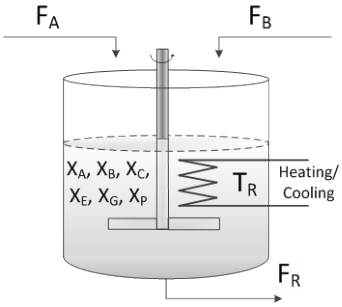

FIGURE 3.28: DIAGRAM OF THE OTTO-WILLIAMS REACTOR.

$$A + B \overset{k_1}{\rightarrow} C, \qquad B + C \overset{k_2}{\rightarrow} P + E, \qquad C + P \overset{k_3}{\rightarrow} G \tag{3.70}$$

The system can be described using a first principles approach. The mass balance for each compound in the vessel can be defined:

$$\begin{aligned}
V_R \frac{dX_A}{dt} &= F_A - F_R X_A - V_R r_1 \\
V_R \frac{dX_B}{dt} &= F_B - F_R X_B - V_R r_1 \frac{\mathcal{M}_B}{\mathcal{M}_A} - V_R r_2 \\
V_R \frac{dX_C}{dt} &= -F_R X_C + V_R r_1 \frac{\mathcal{M}_C}{\mathcal{M}_A} - V_R r_2 \frac{\mathcal{M}_C}{\mathcal{M}_B} - V_R r_3 \\
V_R \frac{dX_E}{dt} &= -F_R X_E + V_R r_2 \frac{\mathcal{M}_E}{\mathcal{M}_B} \\
V_R \frac{dX_G}{dt} &= -F_R X_G + V_R r_3 \frac{\mathcal{M}_G}{\mathcal{M}_C} \\
V_R \frac{dX_P}{dt} &= -F_R X_P + V_R r_2 \frac{\mathcal{M}_P}{\mathcal{M}_B} - V_R r_3 \frac{\mathcal{M}_P}{\mathcal{M}_C} \\
F_R &= F_A + F_B
\end{aligned} \tag{3.71}$$

Where $\mathcal{M}_i$ represents the molecular weight of the compound i, and r_j is the molecular reaction rate of the chemical reaction j defined with respect to its limiting reactive.

Since we are dealing with pseudo-compounds it is necessary to define the relation among their molecular weight. This can be obtained assuming that $\mathcal{M}_A = \mathcal{M}_B = \mathcal{M}_P$. Under this consideration the ratios from equation (3.71) are:

$$\frac{\mathcal{M}_B}{\mathcal{M}_A} = \frac{\mathcal{M}_P}{\mathcal{M}_B} = 1, \quad \frac{\mathcal{M}_C}{\mathcal{M}_A} = \frac{\mathcal{M}_C}{\mathcal{M}_B} = \frac{\mathcal{M}_E}{\mathcal{M}_B} = 2, \quad \frac{\mathcal{M}_G}{\mathcal{M}_C} = 1.5, \quad \frac{\mathcal{M}_P}{\mathcal{M}_C} = 0.5 \tag{3.72}$$

Regarding the reaction rate, it can be calculated as the product of the concentration of the compounds involved in the reactions(Levenspiel, 1999):

$$\begin{aligned} r_1 &= k_1 X_A X_B \\ r_2 &= k_2 X_B X_C \\ r_3 &= k_1 X_C X_P \end{aligned} \tag{3.73}$$

Where k_j is the kinetic constant of the reaction j that can be obtained using an Arrhenius expression.

$$k_j = k_j^0 \exp\left(-\frac{E_{A_j}}{T_R}\right) \tag{3.74}$$

Being E_{A_j} the activation energy from reaction j.

The operational goal for this system is to maximize the steady-state profit of the reactor, which can be expressed as a function of the flow rate of the compounds (equation (3.75)).

$$f_{Otto} \coloneqq F_R(X_P P_P + X_E P_E) - F_A X_A C_A - F_B X_B C_B \tag{3.75}$$

Being P_j the price of produced compound j and C_i the cost of the raw material i.To achieve the operational objective, the system can be operated changing the flow rate F_B and also the reactor temperature T_R by means of the heating/cooling system.

The mass fraction of the product C is one order of magnitude below the rest of the compounds. Therefore, a common choice in a gross representation of the process is considering only the other five species, with the corresponding modeling mismatch. Forbes and Marlin(Forbes & Marlin, 1996), defined a model of the Otto Williams example to be used in the model-based optimization, neglecting the existence of the product C and considering only two parallel reactions inside the reactor (equation (3.76)). In these reactions, the compound C is not present because this is an intermediate product. This fact implies that normally it will not be detected by laboratory measurements in steady-state; therefore, it seems to be a reasonable source of mismatch.

$$A + 2B \overset{\tilde{k}_1}{\rightarrow} P + E, \qquad A + B + P \overset{\tilde{k}_2}{\rightarrow} G + E \tag{3.76}$$

With this given source of modeling mismatch, the steady state model to be used in the RTO layer is the following:

$$\begin{aligned}
&F_A - F_R X_A - V_R \tilde{r}_1 - V_R \tilde{r}_2 = 0 \\
&F_B - F_R X_B - 2V_R \tilde{r}_1 - V_R \tilde{r}_2 = 0 \\
&-F_R X_E + 2V_R \tilde{r}_1 = 0 \\
&-F_R X_G + 3\tilde{r}_2 = 0 \\
&-F_R X_P + V_R \tilde{r}_1 - V_R \tilde{r}_2 = 0 \\
&F_R = F_A + F_B \\
&\tilde{r}_1 = \tilde{k}_1 X_A (X_B)^2 \\
&\tilde{r}_2 = \tilde{k}_2 X_A X_B X_P \\
&\tilde{k}_j = \tilde{k}_j{}^0 \exp\left(\frac{\widetilde{E_{A_j}}}{T_R}\right), \qquad j = 1,2
\end{aligned} \tag{3.77}$$

Where the tilde represents the parameters used in the model with mismatch.

Hence, the model based optimization can be summarized as: finding the decision variables F_B and T_R inside a feasible space, such that they

maximize the profit, subject to a model that takes into account only five compounds and two chemical reactions, corresponding to the simulated modeling mismatch:

$$\begin{aligned} &\min_{F_B, T_R} -f_{Otto} \\ &s.t.: \\ &steady-state\ model\ with\ mismatch\ from\ equation\ (3.77) \\ &F_B \in [F_B^L, F_B^U] \\ &T_R \in [T_R^L, T_R^U] \end{aligned} \tag{3.78}$$

The nomenclature used in the Otto-Williams Reactor is presented in Table 3.7.

TABLE 3.7 NOMENCLATURE EMPLOYED IN THE OTTO-WILLIAMS REACTOR

Variable	Meaning	Units
F_A, F_B	Mass inlet flow of A and B	kg/s
F_R	Mass outlet flow	kg/s
T_R	Reactor Temperature	^{o}C
X_i	Mass fraction of the i compound	Adimensional
V_R	Reactor holdup	kg
r_j	Molecular reaction rate of reaction j in the process	$kmol/l\ s$
$\tilde{r}_j$	Molecular reaction rate of reaction j in the model with mismatch	$kmol/l\ s$
k_j	Kinetic constant of reaction j in the process	$kmol/l\ s$
$\tilde{k}_j$	Kinetic constant of reaction j in the model with mismatch	$kmol/l\ s$
k_j^0	Kinetic constant in the Arrhenius expression for reaction j in the process	$kmol/l\ s$
$\tilde{k}_j{}^0$	Kinetic constant in the Arrhenius expression for reaction j in the model with mismatch	$kmol/l\ s$
E_{A_j}	Energy constant in the Arrhenius expression for reaction j in the process	^{o}C
$\widetilde{E_{A_j}}$	Energy constant in the Arrhenius expression for reaction j in the model with mismatch	^{o}C
F_B^L, F_B^U	Bounds for F_B	kg/s
T_R^L, T_R^U	Bounds for T_R	^{o}C

The values of the parameters utilized from equation (3.71) to equation (3.78) are summarized in Table 3.8.

TABLE 3.8 VALUE OF THE PARAMETERS EMPLOYED IN THE OTTO-WILLIAMS REACTOR

Parameter	Value	Parameter	Value
F_A	1.8725	$\widetilde{E_{A_1}}$	-8077.6
V_R	2105	$\widetilde{E_{A_2}}$	-12438.5
k_1^0	1.6599×10^{6}	F_B^L	3
k_2^0	7.2177×10^{8}	F_B^U	6
k_2^0	2.6745×10^{12}	T_R^L	70
E_{A_1}	-6666.7	T_R^U	100
E_{A_2}	-8333.3	P_P	1143.38
E_{A_3}	-11111	P_E	25.92
$\tilde{k}_1{}^0$	2.611×10^{12}	C_A	76.23
$\tilde{k}_2{}^0$	1.655×10^{8}	C_B	114.34

The nested, along with the dual modifier-adaptation methods, where tested starting from two different points: (**A**) the optimum of the model and (**B**) a boundary of the feasible region. Because the system presents only inequality constraints in the decision variables, to modify the optimization it is necessary only the corrector of the gradient of the objective function λ. As in the previous examples, the first two iterations were used to estimate the gradients with the finite difference approach, which have been used as a starting point for the NM algorithm in the upper optimization layer of the nested methodology.

Figures 3.29 and 3.30 present the comparison of the evolution of the decision variables and the objective function, for the nested and the dual modifier – adaptation method, changing the lower bound of the dual constraint in the last case and starting from (**A**). In these figures **R** represents the optimum of the process.

The results of the modifier-adaptation methodology from Figures 3.29 and 3.30 show that the nested methodology converges into the real optimum of the process in the minimum number of iterations compared with the dual approach: starting from the first two iterations, where the finite difference approach was implemented to give good initial iterators, the nested

method converged in 8 iterations. On the other hand, the convergence of the dual approach was a function of the level of excitation of the system, noting that for a gradient estimated with $\delta^L = 0.05$ the path converged into a point different to the required one, unlike the manifolds formed with $\delta^L = 0.1$ and 0.15 which were able to find the optimum of the process, in 17 and 57 iterations respectively. As it was discussed in previous examples, an increase in the excitation level on the dual methodology improves the estimation of the gradient, which is translated in the detection of the real optimum of the process; nevertheless, this is not free because the number of iterations must growth as a consequence of the reduction in the feasible region of the dual modified problem.

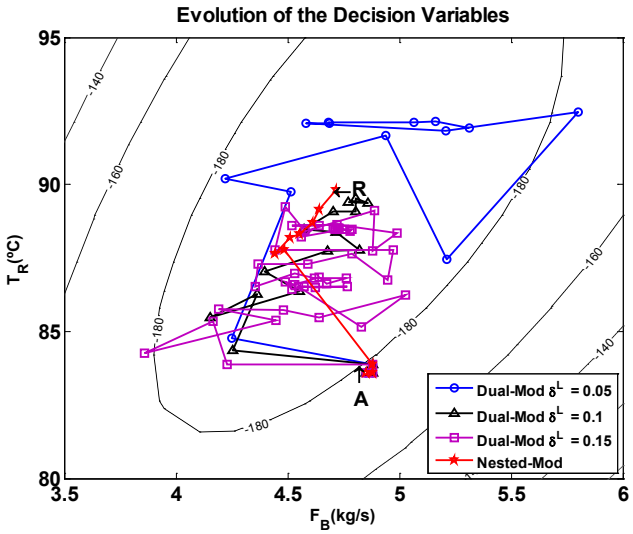

FIGURE 3.29: EVOLUTION OF THE DECISION VARIABLES FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE OTTO-WILLIAMS REACTOR STARTING FROM (A). BLACK DASHED LINE IS DE COST FUNCTION OF THE PROCESS

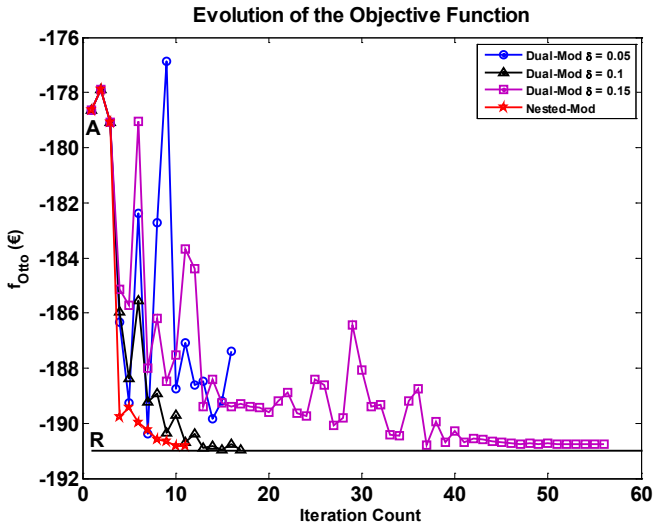

FIGURE 3.30: EVOLUTION OF THE OBJECTIVE FUNCTION FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE OTTO-WILLIAMS REACTOR STARTING FROM (A)

Comparing now the path formed with the two methodologies, it can be noted from previous figures that the proposed algorithm finds the optimum of the process in a more direct way, because the method does not need estimating the gradient of the process to update the modifiers. Instead of this it uses directly the cost function measured from the process and a geometric method based in previous points, which has no limitations about forming linear combinations with the next decision variables. This result is in concordance with the discussions from previous examples.

Now we are interested in evaluating the performance of the algorithm starting from a point which is farther from the real optimum of the process. Figure 3.31 and 3.32 present the same comparison from Figures 3.29 and 3.30, but starting from (**B**). In these figures **R** represents the optimum of the process.

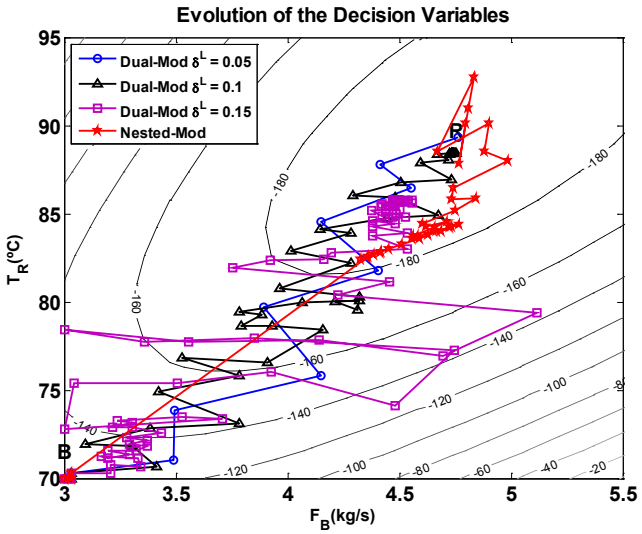

FIGURE 3.31: EVOLUTION OF THE DECISION VARIABLES FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE OTTO-WILLIAMS REACTOR STARTING FROM (B). BLACK DASHED LINE IS DE COST FUNCTION OF THE PROCESS

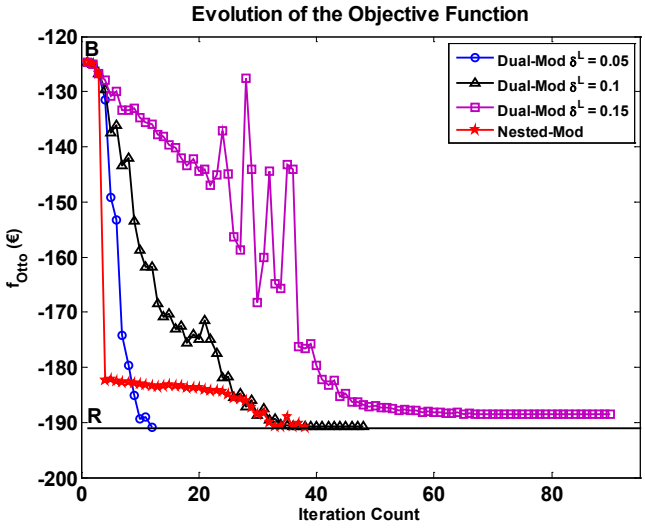

FIGURE 3.32: EVOLUTION OF THE OBJECTIVE FUNCTION FOR THE DUAL AND THE NESTED MODIFIER ADAPTATION IN THE OTTO-WILLIAMS REACTOR STARTING FROM (B)

The results from Figure 3.31 and 3.32 do not agree with the ones starting from (**A**), noting that in this case the best way to apply the modifier adaptation methodology, in terms of convergence and number of iterations, is using the dual approach with a degree of excitation equivalent to $\delta^L = 0.05$ converging in only 13 iterations, indicating that the process gradient has been estimated accurately and the feasible region is large enough to no interfere in the evolution of the algorithm. On the other hand, the nested methodology again is able to locate the real optimum of the process; nevertheless its convergence takes a number of iterations three times larger. The path formed by this method shows an important reduction in the objective function in the first iterations because of the good initial points given, after this, the gradient-free approach iterates with the modifiers changing the shape of the simplex, forming a zone where the iterations are very close between them explaining the growth in the number of iterations. Its performance is comparable with the path obtained with the dual methodology for $\delta^L = 0.1$, which finds the optimum in a similar number of iterations. In contrast, the dual methodology with $\delta^L = 0.15$ fails in finding the optimum of the process because of the reduction on the dual feasible region.

The fact that the optimum of the process is reached in less iterations with the lowest value of the dual bound starting from (**B**), but it was not able to find the process optimum if the starting point is (**A**) indicates that the dual constraint was not active in all the iterations from Figure 3.31 and 3.32. This can explain the correct detection of the curvature of the process (confirmed with the detection of the process optimum), and also the lower number of iterations needed for convergence. This condition can be very complicated in the implementation of the dual methodology, since an accurate gradient estimation of the process must be ensured in all the cases within the feasible region, implying that even when the process optimum has been found in a few iterations with $\delta^L = 0.05$ starting from (**B**), the more robust way to apply the dual approach for he Otto-Williams reactor is using dual

bound larger. Taking this into account, the proposed gradient-free methodology produces similar results than the dual approach in terms of the ability to find the process and the number of iterations to converge.

As a remark in the implementation of the nested approach in the Otto-Williams example, we can say that the proposed methodology can detect the optimum of the process in a similar way than the gradient-based approach for a problem with structural mismatch, as the same as in previous examples with parametric uncertainty, without the need to estimate the process gradients. The convergence of the suggested method is comparable with the dual approach with enough energy to estimate accurately the gradients of the process. Nevertheless, it is important to mention that because of the excitation of the system can be removed its application seems to be easier than the gradient-based algorithm.

3.6.3.4 Continuous Bioreactor

In the same line as the previous Otto-Williams reactor, this example tries to test the nested modifier-adaptation methodology in a system with structural modeling mismatch, but with the additional difficulty of a change in the behavior of the process in a region closer to the optimum as we will see next.

The example is a bioreactor that has been used previously to test RTO methodologies (Y. Zhang & Forbes, 2006), and was presented in the context of the Adaptive Extremum Control (Golden & Ydstie, 1989). The system contains a continuous culture that grows inside a bio-CSTR. With the adequate constant aeration (Air), mixing conditions, substrate concentration at the feed (S_0) and pH ($NaOH$), the dilution rate D of the reactor can be modified in order to change the concentration of the cells (or biomass) at the effluent of the reactor (X). To be more precise, the real manipulated variable of the system is the flow rate of the feed (F); nevertheless, if the volume of the reactor (V) is constant, these two variables are equivalent ($D = F/V$). Figure 3.33 schematizes the system.

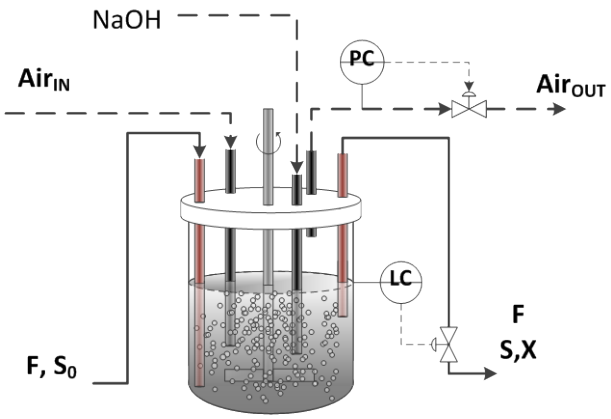

FIGURE 3.33: DIAGRAM OF THE CONTINUOUS BIOREACTOR.

The continuous bioreactor can be described using a first principle approach that will be used as a simulation of the real process (Golden & Ydstie, 1989), following a dynamic fermentation process.

A mass balance at the cells can be written assuming that there is no biomass in the influent F and the reactor volume remains constant:

$$\frac{dX}{dt} = (\mu - D)X \tag{3.79}$$

Where μ is the specific growth rate of the microorganisms that can be modeled using the Monod equation.

$$\mu = \mu_{max} \frac{S}{K_S + S} \tag{3.80}$$

Being μ_{max} and K_S constants of the model, and S the concentration of the substrate inside the vessel, which can be obtained with a mass balance:

$$\frac{dS}{dt} = -\frac{\mu X}{Y} + D(S_0 - S) - m_c X \tag{3.81}$$

In equation (3.81) S_0 is the concentration of the biomass in the influent, Y is a yield that takes into account the proportion of substrate that is transformed into biomass, while m_c is a constant that takes into account the relative quantity of the substrate that the biomass uses for its internal metabolism.

The operational goal of the system is to maximize the production of cells in steady state. This can be expressed as a product of the dilution rate and the concentration of the microorganisms:

$$f_{Bio} \coloneqq XD \tag{3.82}$$

The concentration of biomass in steady-state (equation (3.83)), can be obtained combining equations (3.79) to equation (3.81), setting the time derivatives to zero.

$$X = \underbrace{\frac{YD}{m_c + D}}_{Y_{obs}} \underbrace{\left(S_0 - \frac{K_S D}{\mu_{max} - D}\right)}_{\Delta S} \tag{3.83}$$

Note that Equation (3.83) becomes singular when $D = \mu_{max}$.

Equation (3.83) can be viewed as a product between the consumed biomass (ΔS)) and an observed yield (Y_{obs}) that depends on the dilution rate. In order to test the modifier methods using an incorrect model, we will suppose that the observed yield of X is constant. Equation (3.84) describes the approximation employed in the model of the process.

$$X = \tilde{Y}_{obs}\left(S_0 - \frac{\tilde{K}_S D}{\tilde{\mu}_{max} - D}\right) \tag{3.84}$$

The parameters empathized with a tilde have the same meaning than the ones described previously, but they have been used in the model with different values. Note that the equation (3.84) in addition to the parametric

uncertainty, presents a structural mismatch with respect to the equations used to model the process since the observed yield does not depend on other variables.

As a summary, we can express the optimization problem to be solved in the continuous bioreactor example as:

$$\begin{aligned} &\min_{D} -f_{Bio} \\ &s.t.: \\ &X = \tilde{Y}_{obs}\left(S_0 - \frac{\tilde{K}_S D}{\tilde{\mu}_{max} - D}\right) \\ &D \in [D^L, D^U] \end{aligned} \tag{3.85}$$

The value of the upper bound of D must be chosen below $\tilde{\mu}_{max}$ since it can produce a singularity in the model.

The nomenclature for the continuous bioreactor is summarized in Table 3.9, whereas the values of the parameters are presented in Table 3.10.

TABLE 3.9 NOMENCLATURE EMPLOYED CONTINUOUS BIOREACTOR

Variable	Meaning	Units
D	Dilution Rate	$1/h$
S	Substrate concentration	g/l
X	Biomass concentration	g/l
S_0	Substrate concentration at the influent	g/l
μ	Specific growth rate of the biomass	$1/h$
D^L, D^U	Bounds of D	$1/h$
K_S	Half-saturation constant for the process	g/l
$\tilde{K}_S$	Half-saturation constant for the model	g/l
Y	Yield coefficient of the biomass with respect to the available substrate	Adimensional
$\tilde{Y}_{obs}$	Observed yield of the biomass with respect to the consumed substrate.	Adimensional
m_c	Maintenance coefficient	$1/h$
μ_{max}	Maximum specific growth rate of the biomass for the process	$1/h$
$\tilde{\mu}_{max}$	Maximum specific growth rate of the biomass for the process	$1/h$

TABLE 3.10 VALUE OF THE PARAMETERS EMPLOYED IN THE CONTINUOUS BIOREACTOR

Parameter	Value
D^L	0
D^U	0.42
S_0	5
m_c	0.025
K_S	0.09
Y	0.5
μ_{max}	0.35
$\tilde{K}_S$	0.19
$\tilde{Y}_{obs}$	0.4
$\tilde{\mu}_{max}$	0.42

The mismatch of the bioreactor produces an interesting effect with respect to the real behavior of the system. Figure 3.34 shows a comparison of the objective function f_{Bio} and the biomass concentration X between the process and the model for the entire feasible region of D.

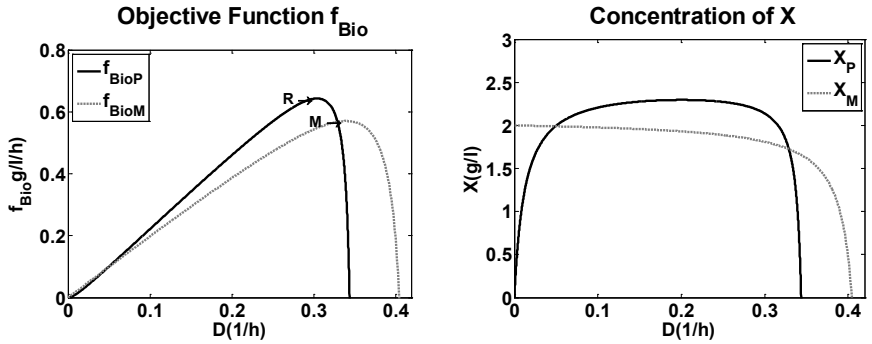

FIGURE 3.34: STEADY-STATE MAPPING OF THE PROCESS AND THE MODEL FOR THE OBJECTIVE FUNCTION AND THE CONCENTRATION OF BIOMASS. SUBSCRIPTS "P" AND "M" DENOTES THE PROCESS AND THE MODEL RESPECTIVELY.

Because of the modeling mismatch, the optimum of the process (**R**) does not match with the one from the model (**M**), situation that is quite expectable. However, from Figure 3.34 it can be observed that the dilution rate that optimizes the performance index of the approximated model of

the reactor causes a washout in the steady state of the real process. Washout is a peculiar phenomenon for continuous culture systems. It consists in a complete removal of the cells inside the reactor ($X = 0$) when the dilution rate is greater than the cell growth rate. Taking this into account, it is clear that the model of the process is a poor representation of the behavior of the process.

Because of the inequality constraints in equation (3.85) are on the decision variables, it is only necessary the use of the corrector of the gradient of the objective function λ. Regarding the washout phenomenon, the feasible region on each iteration must be reduced in order to avoid this situation as equation (3.86) shows, where a trust region equation has been added.

$$\begin{aligned}
&\min_{D} -f_{Bio} \\
&s.t.: \\
&steady - model\ with\ mismatch\ from\ equation\ (3.84) \\
&D_k \in [D_k^L, D_k^U] \\
&D_k^L \coloneqq \max(D^L, D_{k-1} - \Delta D/2) \\
&D_k^U \coloneqq \min(D^U, D_{k-1} + \Delta D/2)
\end{aligned} \tag{3.86}$$

For the gradient based modifier-adaptation, the finite differences approach has been used in order to avoid the dual constraint that can cause problems with the washout phenomenon. Regarding the nested approach, the first iteration with finite differences has been used as a starting point for the upper optimizer.

The evolution of the decision variables and the objective function for the gradient-based methodology are summarized in Figure 3.35. In this implementation, the size of the local feasible region was $\Delta D = 0.01$. This value is extremely sensitive in terms of the convergence, noting that if it is increased, the algorithm does not converge and the decision variable bounces around the optimum of the process.

The convergence to the real optimum of the process in the gradient-based approach takes around 40 iterations. It can be noted that the path formed with the decision variables is very straight, unlike the dual modifier adaptation method that in general generates a zigzag manifold. This result is similar to the one obtained in the oxygen consumption reactor from section 3.3.1, and, as it was mentioned previously, can be explained in terms of the reduction of the feasible region because of the dual constraint.

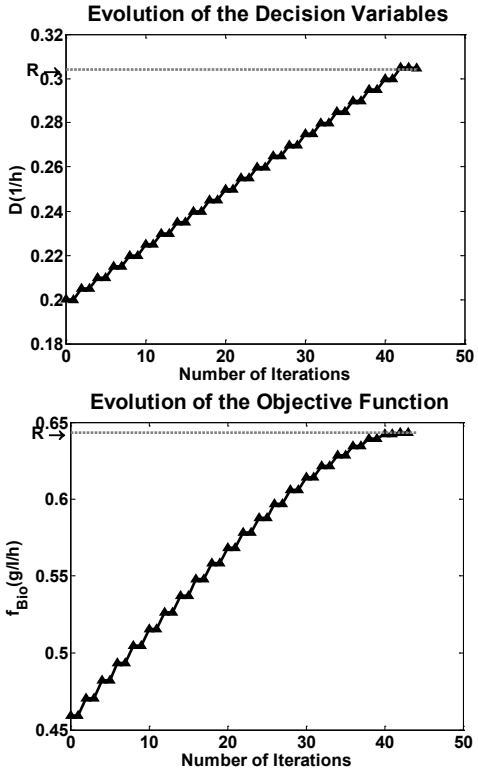

FIGURE 3.35: EVOLUTION OF THE MODIFIER ADAPTATION METHODOLOGY WITH FINITE DIFFERENCES FOR THE CONTINUOUS BIOREACTOR. GREY DASHED LINE SHOWS THE OPTIMUM OF THE PROCESS.

Regarding the detection of the optimum of the process, the shrinking of the local feasible region, as a consequence of changing the bounds in the

decision variables in each iteration, allows to converge to the desired point. The decrease in the feasible region avoids the washout phenomenon that produces important instabilities in the algorithm. The problems in the convergence can be explained noting that if the decision variable is approaching the real optimum from below, the gradient of the process is more positive than the derivative from the model. As result of this, the value of decision variable will increase. If the feasible region is not constrained, the next operating point could pass the optimum of the process (not detecting it) and fall in the region closer to the washout phenomenon. If this is the case, now the process derivative will be more negative than the gradient of the model, moving the system again into the region before the optimum of the process, repeating this behavior over and over and explaining the observed bouncing for larger values of ΔD. The problem in the change on the sign of the modifiers, can be viewed also as a consequence of an inaccurate correction of the objective function using a first order approximation in the region close to the optimum of the process.

Regarding the implementation of the nested methodology, the *fminsearch* function might present problems in the optimization when the optimum of the process is near a discontinuous region (Mathworks, 2007). If the algorithm is applied directly as in previous examples, the nested methodology finds the optimum of the process, but it still searches in the entire region. Figure 3.36 shows the implementation of the Nelder-Mead algorithm in the continuous bioreactor. The outcome of the NM algorithm was the real optimum of the process, as the same as the gradient-based approach. In spite of this, the evolution of the method shows that the algorithm does not stop in this point and continues iterating on the rest of the feasible region. This behavior can be explained based in the fact that there is an important change in the rate of change of the objective function on both sides of the process optimum. As a consequence of this, if the system starts from a point below the optimum of the process, the size of the simplex will be bigger than the one required when the process pass the

real optimum, because of the rate of change is increased after this point. This implies that the next RTO iteration will be farther from the required point. In order to avoid this situation, we have tried working with smaller local feasible regions with no positive results.

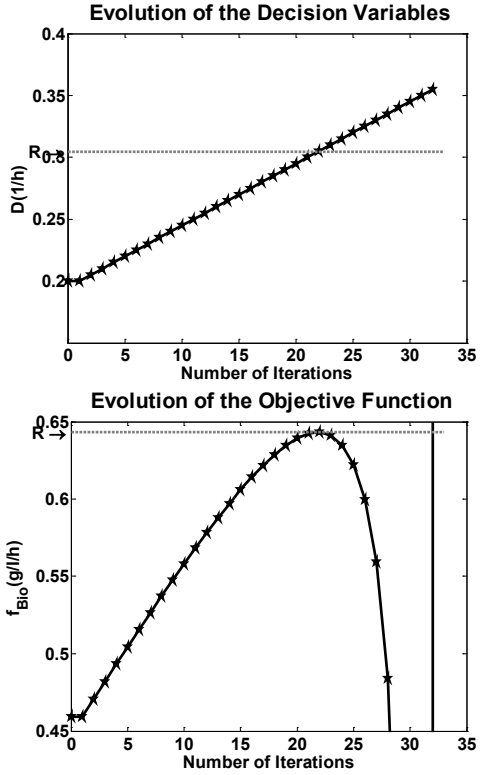

FIGURE 3.36: EVOLUTION OF THE NESTED MODIFIER-ADAPTATION METHODOLOGY FOR THE CONTINUOUS BIOREACTOR. GREY DASHED LINE SHOWS THE OPTIMUM OF THE PROCESS.

With the purpose of making the nested methodology applicable in this case, we have used the previous knowledge of the system to stop the algorithm when a situation of possible optimum is detected. Even when the point of washout is not known a priori, these kinds of systems exhibit this phenomenon for some value of the dilution rate. Furthermore, starting

from a suboptimal operation point (on the left side of the process optimum in Figure 3.34) an increase on the dilution rate will rise the concentration of the cells inside the vessel because there is additional available subtract to produce more biomass. This increase is maintained until the maximum growth rate (an unknown parameter) is exceeded. After this point, the cells inside the reactor have not enough residence time to create more microorganisms, decreasing the concentration of the biomass and producing the washout. As this is an expectable behavior of the continuous cell cultures, we can try to detect the point when the change in the performance of the reactor is produced, postulating it as a point (or a region) close to the real optimum. As a criterion to detect a candidate for process optimum, it has been used the change in the tendency of the objective function.

If we are interested in refining the search, we can implement a gradient based methodology before the gradient free algorithm, following the analogy that some global optimization solvers, like the SSm GO (Egea et al., 2007), implement in their routines. As it was commented in section 3.6.1, the possibility of mixing the two approaches (gradient free with gradient-based) is one of the contributions that the reinterpretation of the modifiers presents, giving an additional degree of adaptation to the modifier method depending on the particular characteristics of the system.

The implementation of the mixed (nested and gradient-based) modifier-adaptation methodology is summarized in Figure 3.37.

The outcomes obtained applying the previous idea, are summarized in Figure 3.38. We have tested the early detection of a possible optimum for the two scenarios: not refining the search and using the gradient based methodology to find the optimum of the process, corresponding to the upper and the lower row of the Figure 3.38 respectively. The gradient of the process in the last case has been estimated with the finite differences approach.

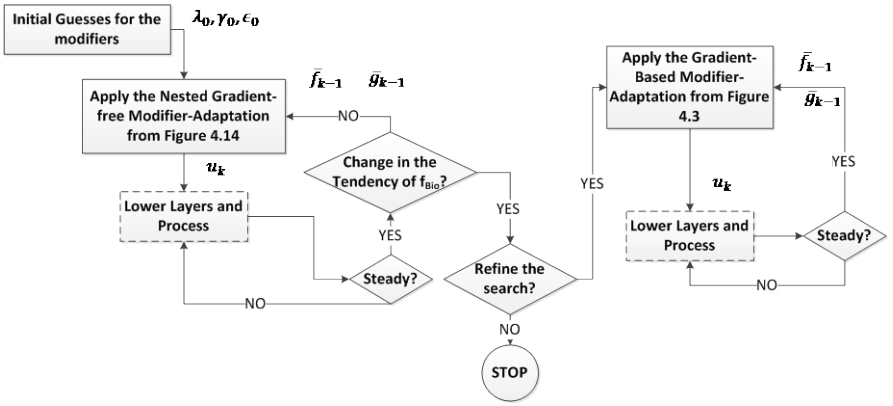

FIGURE 3.37: IMPLEMENTATION OF THE MIXED MODIFIER-ADAPTATION

For the case of no refining in the search of the optimum of the process (upper row of Figure 3.38), the nested methodology with the early detection of the process optimum is able to stop once the change in the tendency of the objective function is detected. As a result of this, the washout phenomenon is no longer observed and the iterations finish in a neighborhood of the process optimum. The last decision variable proposed by the nested methodology is not exactly the real optimum of the process but an operating point that is slightly above this value, producing a mild deterioration of the objective function corresponding to a 0.35% with respect to the real value. This worsening in the objective function can be understood as a tradeoff with respect to the number iterations obtained with the gradient-based algorithm, and it can be a good indicator regarding if it is necessary the refining of the search using the gradient-based methodology or not.

Even when the worsening obtained in the objective function can be acceptable for a real application, taking into account the reduction in the number of iterations of 50%, we have refined the search applying the gradient-based methodology to the outcomes from the NM algorithm in order to test the complete method. The evolution of the mixed

methodology can be observed in the lower row of Figure 3.38, where the starting point of the refinement iterations is highlighted with a black arrow.

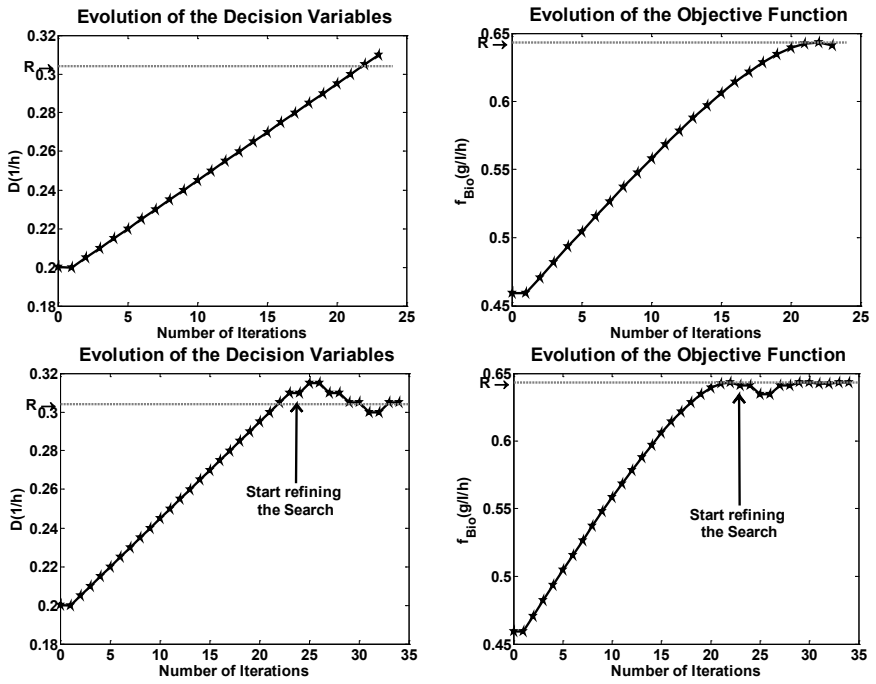

FIGURE 3.38: EVOLUTION OF THE NESTED MODIFIER-ADAPTATION METHODOLOGY FOR THE CONTINUOUS BIOREACTOR WITH TERMINATION FUNCTION. UPPER ROW GREY DASHED LINE SHOWS THE OPTIMUM OF THE PROCESS

As it can be observed from the lower row of Figure 3.38, the mixed approach finds the optimum of the process by adjusting the results given by the NM algorithm with early the optimum detection, using a gradient based algorithm. Starting from the end of the gradient-free methodology it was necessary 11 iterations to converge to the real optimum of the process. That is to say: to reduce a 0.35% in the objective function measured from the process, it was necessary to increase the number of iterations in 45% with respect to apply only the gradient-free scheme. Even though the application of the gradient-based method seems to be unjustified to find the optimum in this example, taking into account the effort in the number

of iterations and the almost no additional improvement in the cost function that was obtained, in real application the optimum of the process is not known a priory. Therefore, when the NM algorithm stops the operator does not know how far is from this point, but only the progresses in the performance index with respect to the starting point. Hence, the decision of applying the refinement or using directly the outcomes from the NM algorithm cannot be taken based in distance with respect to the final goal, and it must be based in the criterion of the operator and in its previous knowledge of the system considering the progress achieved. In any case, the nested methodology with the early stop criterion only is able to find a region where the optimum could be according to the expected behavior around this point.

As a remark in the implementation of the Nested methodology in the continuous bioreactor example, we can say that for a problem with structural mismatch where the optimum of the process is closer to a region that can produce important changes in the operation mode of the system with respect to the nominal mode, the implementation of the proposed methodology by itself is not recommended, requiring additional information about the process behavior to early detect a region where the process optimum could be. With this scheme we can avoid falling in problematic regions of operation that can originate important damages in the system, but not detecting the optimum of the process. Depending on the progress observed from the start of the algorithm to the early stopping criterion, we can apply additional perturbations into the system in order to improve the objective function of the process using a gradient-based methodology. For both updates criteria: gradient-based and direct search, special care must be taken in the size of the feasible region for each RTO iteration with the intention of remaining in the nominal operation region during the entire evolution of the algorithm.

3.6.3.5 Run to Run Semi-Batch Reactor

Previous two examples have shown the expected evolution of the proposed nested approach when structural mismatch is observed. On the other hand, the examples of convex optimization and the three interconnected tanks have tested the methodology under the presence of process noise with parametric uncertainty. Putting aside the particularities in the application of the nested-modifier adaptation for each example, it is observed in all the cases that the algorithm converges to the real optimum (or into a region close to it), with almost no information about the process derivatives. The example presented in this section deals both with structural mismatch and process noise. In addition, it considers the process dynamics.

The system was presented as a test example in the work of Chachuat and coworkers in the context of a comparison among different ways to look for the real optimum of a process in RTO (Chachuat et al., 2009). It consists in a semi-batch reactor (Figure 3.39) that is used to produce 2-acetoacetyl pyrrole (C) from pyrrole (A) and diketene (B). Initially, the vessel contains a solution rich in A, plus some amount of the other compounds and sub-products from parallel reactions that can occur: dehydroacetic acid (D), Oligomers (E) and other undesired by-products (P). From the beginning until the end of the batch ($t_f = 250\,min$), an inlet solution (F) with constant concentration of diketene (C_B^{in}) can be fed into the reactor, being possible regulate its flow rate during the whole period in order to change the concentrations of the products and the end of the batch.

The reactions that take place in the reactor can be summarized as:

$$\begin{aligned} &A + B \overset{k_1}{\rightarrow} C \\ &2B \overset{k_2}{\rightarrow} D \\ &B \overset{k_3}{\rightarrow} E \\ &C + B \overset{k_4}{\rightarrow} P \end{aligned} \tag{3.87}$$

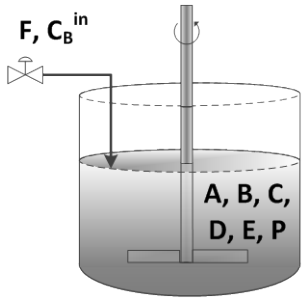

FIGURE 3.39: DIAGRAM OF THE SEMI-BATCH REACTOR

A first-principles model can be proposed to describe the process based in mass balances:

$$\begin{aligned}
\frac{dC_A}{dt} &= -r_1 - \frac{F}{V}C_A \\
\frac{dC_B}{dt} &= -r_1 - 2r_2 - r_3 - r_4 + \frac{F}{V}\left(C_B^{in} - C_B\right) \\
\frac{dC_C}{dt} &= r_1 - r_4 - \frac{F}{V}C_C \\
\frac{dC_D}{dt} &= r_2 - \frac{F}{V}C_D \\
\frac{dV}{dt} &= F
\end{aligned} \tag{3.88}$$

Being C_i the molar concentration of the i compound, V the reaction volume and r_j the molar concentration rate from reaction j, which can be calculated as the product of the concentrations of the reactants:

$$\begin{aligned}
r_1 &= k_1 C_A C_B \\
r_2 &= k_2 C_B^2 \\
r_3 &= k_3 C_B \\
r_4 &= k_4 C_B C_C
\end{aligned} \tag{3.89}$$

The operational goal of the system is finding the optimal profile of the flow rate of feed F, such that the production of C at the final time is maximized,

keeping C_B and C_D lower than some upper bounds, solving the following optimization problem:

$$\begin{aligned}
&\min_{F(t)} J_{Batch} \coloneqq -C_C(t_f)V(t_f) \\
&s.t.: \\
&Process: equations\ (3.88) and\ (3.89) \\
&C_B(t_f) \leq C_B^U \\
&C_D(t_f) \leq C_D^U \\
&F(t) \in [F^L, F^U], \qquad t \in [t_0, t_f]
\end{aligned} \tag{3.90}$$

Regarding the model of the process, as a source of mismatch, we have neglected the existence of the last two reactions from equation (3.87), since they are parallel reactions that in general present conversion rates with orders of magnitude lower than the first two. This is equivalent to set k_3, k_4 equal to zero.

Therefore, we can summarize the model-based optimization to be implemented in the RTO layer as:

$$\begin{aligned}
&\min_{F(t)} J_{Batch} \coloneqq -C_C(t_f)V(t_f) \\
&s.t.: \\
&Model\ with\ only\ two\ reactions \\
&C_B(t_f) \leq C_B^U \\
&C_D(t_f) \leq C_D^U \\
&F(t) \in [F^L, F^U], \qquad t \in [t_0, t_f]
\end{aligned} \tag{3.91}$$

In the optimization problem from equation (3.91), we are interested in obtaining a dynamic trajectory to be applied during the entire batch that minimizes the cost function and fulfill the inequality constraint at the end of the run. Even when this is not a steady-state problem, we can implement the modifier-adaptation methodology in a similar way than the steady problem, taking into account that the process measurements used to estimate the modifiers will be calculated after the batch is finished, and the

decision variable calculated in the model based optimization will be applied in the next batch.

The nomenclature for equations (3.88) to (3.91) is summarized in Table 3.11, while the values of the parameters used in the simulated process as well as the initial states, are summarized in Table 3.12 (Chachuat et al., 2009)

TABLE 3.11 NOMENCLATURE EMPLOYED IN THE SEMI-BATCH REACTOR

Variable	Meaning	Units
C_i	Molar concentration of compound i	mol/l
C_B^{in}	Molar concentration of B in the inlet solution	mol/l
F	Flow rate of the inlet solution	l/min
V	Volume of the reactor	l
r_j	Conversion of reaction j	$mol/l\ min$
k_j	Kinetic constant of reaction j	$l/mol\ min$ or $1/min$
F^L, F^U	Bounds of F	l/min
C_i^U	Upper bound of concentration of i	mol/l
t_0	Initial time of the batch	min
t_f	Final time of the batch	min

TABLE 3.12 VALUE OF THE PARAMETERS AND INITIAL STATES EMPLOYED IN THE SEMI-BATCH REACTOR

Parameter	Value	States	Value
k_1	0.053	$C_A(t_0)$	0.72
k_2	0.128	$C_B(t_0)$	0.05
k_3	0.028	$C_C(t_0)$	0.08
k_4	0.001	$C_D(t_0)$	0.01
C_B^{in}	5	$V(t_0)$	1
t_0	0		
t_f	250		
F^L	0		
F^U	2×10^{-3}		
C_B^U	0.025		
C_D^U	0.15		

As it was mentioned previously, the simulated mismatch consists of neglecting the last two reactions from equation (3.87) setting their corresponding kinetic constants to zero. This, far from being a parametric uncertainty, reflects a structural mismatch since physical dependences among the variables have been omitted because of partial knowledge of the real process.

The dynamic optimization from equation (3.90) must be transformed into a finite dimensional one, parameterizing the variables to solve it using standard NLP techniques. In general, there are two approaches to do this: simultaneous and sequential (Biegler Lorenz, 2010). In the simultaneous approach, both state and decision variables are discretized using some colocation strategy. On the other hand, in sequential approach only the decision variables are parameterized, while the state variables are calculated in an intermediate step using a DAE solver (as it was already presented in Chapter 2). To solve equation (3.87) we have applied the sequential approach.

The optimal trajectory of the real process was obtained solving equation (3.90) using two different control vector parameterizations as Figure 3.40 shows. The first implementation (dashed grey line) was obtained with a piece-wise constant parameterization dividing the batch time in 100 constant-length intervals. It can be noted that the solution can be characterized by three arcs: (1) from t_0 to t_s the feed is at its upper bound, (2) from t_s to t_m we have assumed for simplicity that the feed remains practically constant in an intermediate flow rate and (3) from t_m to t_f the flow rate of the feed is equal to zero. Taking this into account, a new parameterization was used dividing the time horizon in three intervals with different lengths, where in the first one $F = F^L$, in the second $F = F_s$ and in the third $F = 0$. The decision variables in this case are the length of the first interval (t_s), the intermediate value of F in the second arc (F_s) and the length of the third interval (t_m). The solution of this second parameterization is in solid black line in Figure 3.40. The difference in both

parameterizations is a degradation in the cost function obtained in the second parameterization from $J_{Batch} = 0.5081$ to $J_{batch} = 0.5079$. The trajectories of the final constrained states are presented in Figure 3.41, where it can be noted that for both parameterizations the optimal solution of the process is when the inequality constraints over C_B and C_D are active. The optimum of the process is $t_s^\star = 7.91(min)$, $F_s^\star = 1.1 \times 10^{-3}(l/min)$ and $t_m^\star = 229.14(min)$.

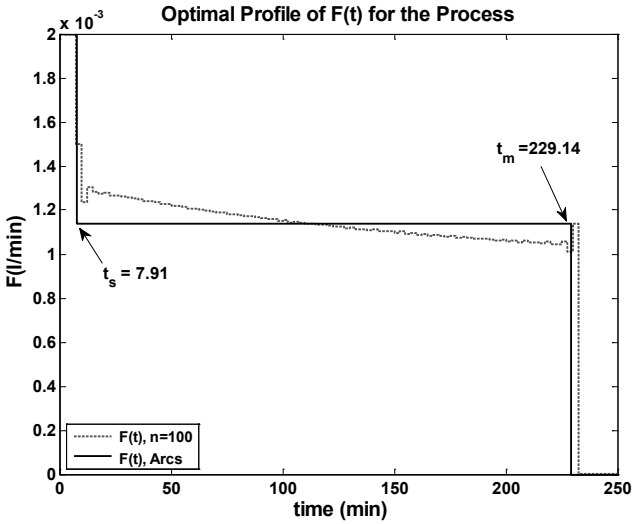

FIGURE 3.40: OPTIMAL PROFILE OF F(T) FOR THE PROCESS OF THE SEMI-BATCH REACTOR

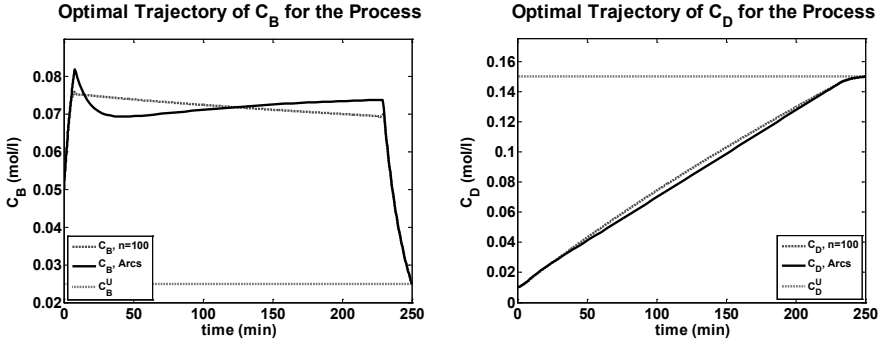

FIGURE 3.41: OPTIMAL TRAJECTORY OF C_B AND C_D FOR THE PROCESS OF THE SEMI-BATCH REACTOR

The same two parameterizations were used to obtain the optimum of the model solving equation (3.91). The optimal trajectory of $F(t)$ with 100 constant-length values and three arcs is presented in Figure 3.42, as can be seen the optimal trajectory of the model with mismatch presents the same structure than the ones computed with the real process: three arcs with constant flow rate of the feed, being the optimum of the model: $t_s^\star = 5.67(min)$, $F_s^\star = 0.7 \times 10^{-3}(l/min)$ and $t_m^\star = 205.2(min)$.

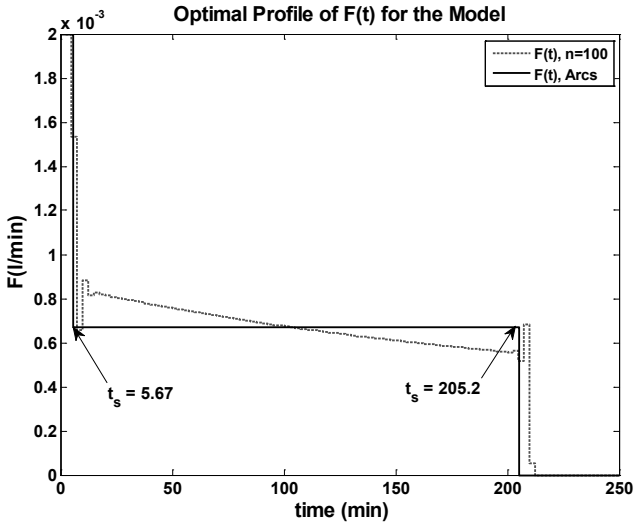

FIGURE 3.42: OPTIMAL PROFILE OF $F(t)$ FOR THE MODEL OF THE SEMI-BATCH REACTOR

If the optimal trajectory obtained with the solution arcs of the model is implemented in the process in an open loop scheme, the objective function decreases to $J_{Batch} = 0.3874$, implying a degradation in the objective function of 23.72%.

The trajectories of the constrained variables of the model and the process applying the open loop solution from the model are presented in Figure 3.43 in gray dashed and black solid lines respectively besides their upper values. It can be noted that in the model, the optimum is produced when the inequality constraints of the state variables are active. However, if this

trajectory is applied to the process the inequality constraints at final time are inactive, explaining the loss of optimality as a consequence of the modeling mismatch.

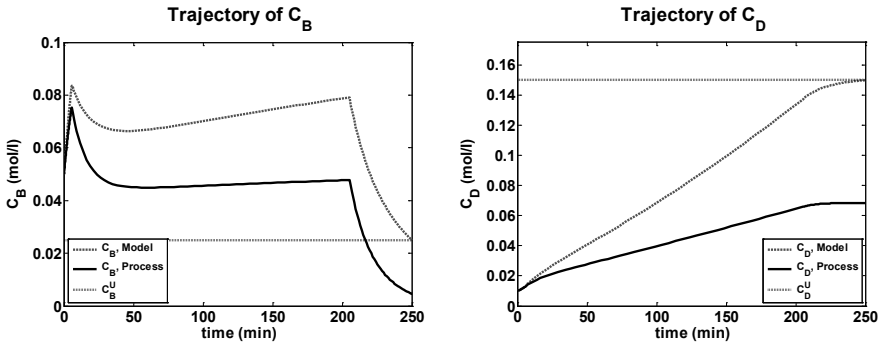

FIGURE 3.43: TRAJECTORY OF C_B AND C_D FOR THE PROCESS AND THE MODEL OF THE SEMI-BATCH REACTOR, APPLYING THE OPTIMAL TRAJECTORY FROM OF THE MODEL

The dual methodology as well as the proposed nested approach has been implemented in this example using the complete set of the modifiers, since the uncertainty is present in both the objective function and the inequality constraints of the state variables. Even if in the original paper the authors only use the bias corrector of the inequality constraint ϵ (Chachuat et al., 2009), we have preferred using λ and γ also in order to compare the performance of the adaptation strategies, taking into account that if only ϵ is updated dual and nested methodologies are the same.

Starting from the optimum of the model, the evolution in the decision variables and the objective function for the dual and the nested method are presented in Figure 3.44. As in previous examples, we have tested different degrees of excitation of the system for the gradient-based approach, which is represented in the graphs for different values of δ^L. Regarding the nested implementation, unlike previous examples, in this case the initial guesses of the modifiers for the upper layer were chosen equal to zero, in order to let the NM algorithm search around the optimum of the model with the simplex in the first iteration.

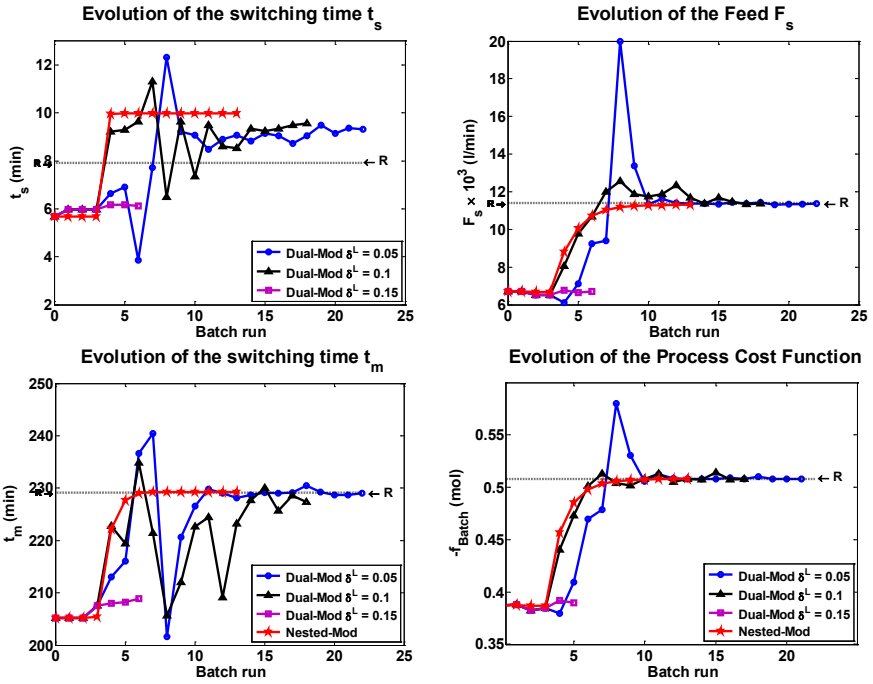

FIGURE 3.44: EVOLUTION OF THE DUAL AND THE NESTED METHODOLOGY FOR THE SEMI-BATCH REACTOR

Notice that because of the example is implemented in a run-to-run approach, each iteration of the RTO layer is a batch run. The comparison on the evolution of the cost function from Figure 3.44, shows that for both methodologies the RTO layer is able to find the optimum of the process. About the evolution of the decision variables the ones from the nested optimization shows a smother and a faster evolution than the dual approach, where the speed of convergence depends on the value of δ^L. Also it can be noted that no strategy was able to locate the real value of t_s, while F_s and t_m were detected successfully. The deterioration in the objective function as a consequence of the incorrect detection of t_s is 0.0056% in the nested approach, implying that the optimal value of the cost function from the process is approximately no sensitive with respect to this decision variable.

The trajectories of the decision and the constrained variables from the process, obtained applying the outcomes from the dual and the nested methods are compared with the expected optimal trajectories from the process in Figure 3.45 and 3.46 respectively.

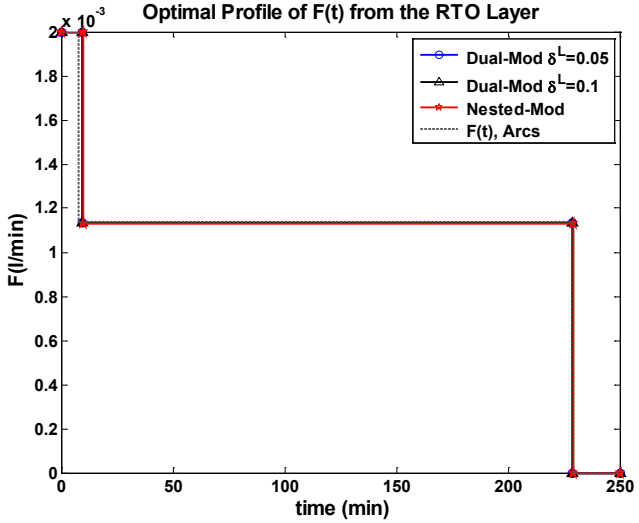

FIGURE 3.45: COMPARISON OF THE OPTIMAL TRAJECTORIES OF $F(t)$ FROM THE RTO LAYER

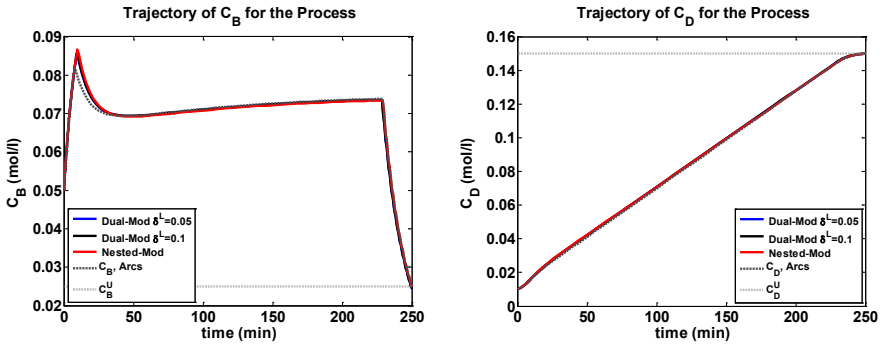

FIGURE 3.46: COMPARISON OF THE TRAJECTORIES OF C_B AND C_D FROM THE PROCESS, APPLYING THE OUTCOMES FROM THE RTO LAYER

It can be noted from previous figures that the only difference in the trajectory of the constrained variables is a slight growth in the

concentration peak of C_B (at $t = 9.7\ (min)$), nevertheless at t_f C_B and C_D are in their upper bounds. Therefore, we can say that, the RTO approaches were able to converge to the optimum of the process.

An additive measurement noise was simulated in the state variables C_B, C_C and C_D, using a Gaussian zero mean distribution function with a 99.5% of confidence interval equivalent to the 10% of the expected range of the molar concentrations. The evolution of the objective function as well as the decision variables for the dual approach are presented in Figures 3.47 and 3.48 for different values of δ^L, while the nested algorithm is shown in Figure 3.49. In each figure, it is represented the evolution of the consecutive batches for 10 different realizations of the stochastic noise.

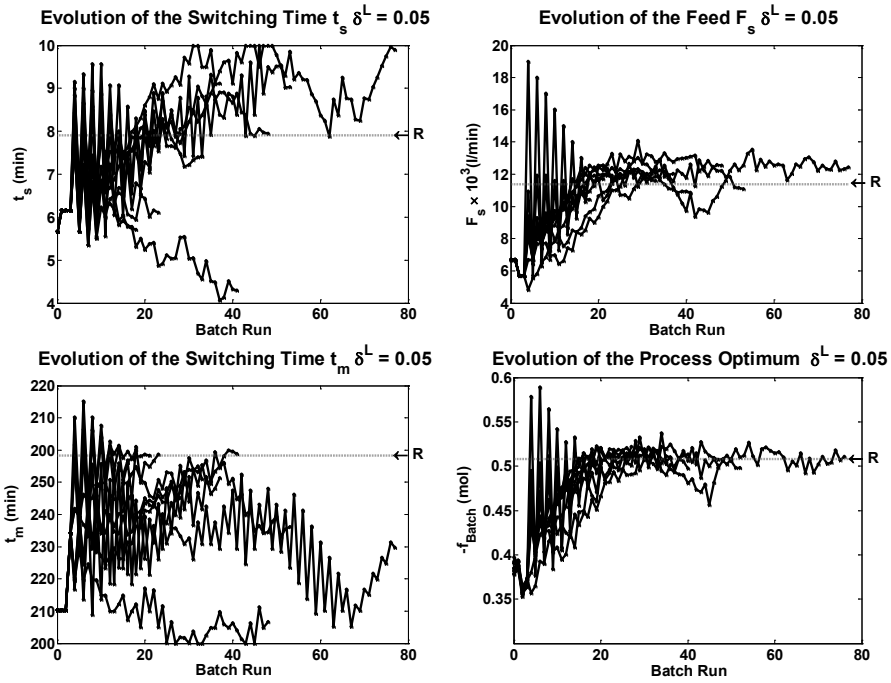

FIGURE 3.47: EVOLUTION OF THE DUAL METHODOLOGY FOR THE SEMI-BATCH REACTOR WITH $\delta^L = 0.05$, UNDER THE INFLUENCE OF PROCESS NOISE

As it is expected, in the gradient-based methodology the evolution of a system poorly excited presents more sensitivity with respect to the measurement noise, observing differences in the capability to converge to the real process depending on the energy to estimate its gradients.

The evolution of the decision variables in the case of $\delta^L = 0.05$ shows how the random nature of the measurements is reflected in the path formed by the dual approach producing a chaotic progression. Regarding the evolution of the objective function, it can be noted that in some cases the system is able to converge into a region close to the optimum of the real system but in a noisy way. The poor ability of the methodology to estimate the process optimum as well as the dispersion in its evolution is the consequence of a bad estimation of the experimental derivatives.

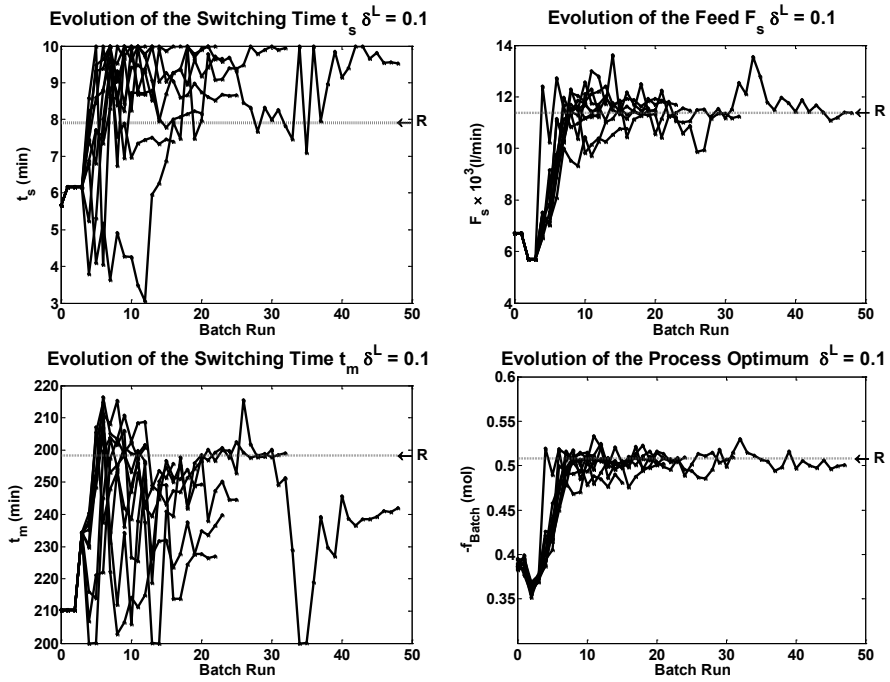

FIGURE 3.48: EVOLUTION OF THE DUAL METHODOLOGY FOR THE SEMI-BATCH REACTOR WITH $\delta^L = 0.1$, UNDER THE INFLUENCE OF PROCESS NOISE

If the bound of the dual constraint is increased, the process growth its capability to estimate accurately the real gradients, which implies a diminution in the dispersion of the path formed by the decision variables and the objective function with $\delta^L = 0.1$. In this case, the system is able to detect the process optimum in all of the times using less iterations to reach a region close to the convergence point and following a more regular trajectory, as a product of the improvement in the estimation of the experimental derivatives.

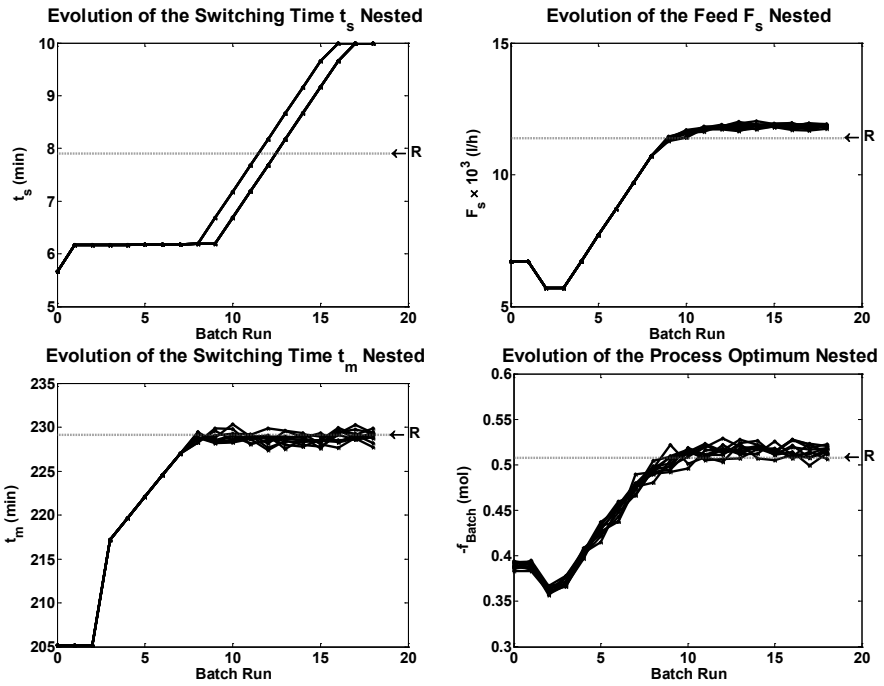

FIGURE 3.49: EVOLUTION OF THE NESTED METHODOLOGY FOR THE SEMI-BATCH REACTOR UNDER THE INFLUENCE OF PROCESS NOISE

Regarding the performance of the nested approach, the influence of the noise in the path formed by the decision variables and the objective function is minimal in comparison of the gradient-based methodology, noting that the shape of the evolution of the algorithm is quite similar to

the noise-free scenario. The algorithm takes about 7 iterations to converge into a region close to the optimum of the process, and after this, the remaining 23 iterations are used to improve the objective function. This result is quite expectable, considering that the volume of the simplex proposed by NM algorithm decreases as the objective function is improved, which means that when the process is near to its optimum the size of the simplex is more affected by the noise of the objective function, worsening the convergence rate to the desired point. This can also explain the small no influence of the noise during the first iterations.

As in the noise free scenario, both dual and nested methodologies were not able to detect the real value of t_s. However, as it was commented previously, the poor influence of this decision variable over the value of the optimum makes this fact less important considering the improvements observed in the objective function applying the RTO layer.

As a remark in the implementation of the nested methodology in the semi-batch reactor, we can say that it was able to find the real optimum of the process under structural modeling mismatch and noise-free scenario in a similar way than the gradient-based methodology, but in less batch runs. If the measurements are contaminated with random noise, the nested approach was capable to estimate a region close to the optimum of the process in a similar way than the noise-free scenario, unlike the gradient based methodology that was more sensitive to the process noise.

3.6.4 Remarks about the Nested Modifier-Adaptation

The basic ideas of the modifier-adaptation methodology suggest a convergence of the RTO layer to the real optimum of the process, correcting the gradients of the model using the information about gradient of the process. Following this notion, it is mandatory to estimate in an accurate way the experimental gradients to achieve the convergence to the desired point. As it was commented, there are some methods to estimate

the process derivatives that in general need a system with enough energy to calculate the modifiers in an appropriate way, being the dual approach the most recommended from the point of view of the tradeoff between accuracy and number of perturbations needed. The last point is a key issue in its application to processes, since it is not desirable perturbing the system because of its collateral effects.

Even if dual methodology is the best alternative to estimate the process gradients, we have shown in the examples that there are important tradeoffs between the energy of the system and the size of the feasible region because of the reduction in this area as a consequence of the dual constraint. The bound of this constraint, and by extension the size of the feasible region, depends on the system and of course on the reliability of the process measurements. That is to say: it must be tuned correctly in order to give enough energy to the system to estimate the gradients in a good way in spite of possible errors in the measurements from the process, but taking care about the feasible region in order avoid blocking the path of convergence to the desired point. As it was shown in previous examples, if this bound is well tuned, the dual methodology is able to detect the optimum of the process. Nevertheless, its convergence also depends in this bound, obtaining more or less iterations depending on its value.

Previous description, gives some insights about the difficulty of applying this method in real life, not only because of the need to estimate the gradient and the continuous perturbations on the process, but also because of the tradeoff that exists with the size of the feasible region of an artificial constraint (artificial in the sense that it is not related with the process itself) and the strong relation that it has with respect to the process noise.

To remove these implementation issues related to the estimation of the process curvature, in this section we have proposed another way to see the modifier adaptation algorithm, with a methodology that intuitively finds the optimum of the process directly in the space of the gradient modifiers. In

our proposal the decision variables to be applied in the process must be calculated as the outcomes of an inner modified optimization problem with the same structure as the one solved in the original modifier-adaptation scheme. Nevertheless, the update law of the modifiers is implemented in an upper optimization layer that uses as a performance index the cost function measured from the process. The assumption of convergence of the modified model implies that the best feasible point to be reached corresponds to the optimum of the process. Regarding now that the feasibility is considered in the nested approach updating the bias of the inequality constraints, which are measured directly from the process, implies that the best feasible point to be the reached when the process cost function is minimized in the space of the gradient modifiers must be also a feasible point that is a minimizer of the process. A formal proof of the convergence of the nested approach has been proposed.

The advantage of understanding the gradient modifiers as decision variables of an outer optimization problem, lies in the fact that it is possible to explore different alternatives to update the modifiers according to the characteristics of the process itself. With this in mind, we can select as a convenient update law a gradient-free algorithm in order to avoid the problems related with the estimation of the experimental gradient.

We have applied this idea into a set of examples that were used previously in the literature of process optimization with different kind of mismatch: parametric and structural. Moreover, we have tested the influence of additional random perturbations in the measurements considering noise-free and noisy scenarios. In all the studied cases, our proposal was capable of finding a region close to the real optimum of the process as the former gradient-based methodology did, regardless of the nature of the mismatch. However, this point was reached without the need to estimate the process derivatives during its evolution, traduced in most of the cases in a reduction in the number of iterations needed to converge. In spite of the number of iterations, the fact that the nested methodology detected the real optimum

of the process is a very important result, due to the fact that we have overcome the modeling mismatch neglecting the main assumption of the modifier adaptation strategy: the estimation of the process derivatives.

For the cases where the direct search approach is not recommended in certain regions of the process, we have shown in the example of the continuous bioreactor that it is possible converging to the real optimum, combining both gradient-based and gradient-free approaches taking the advantage of previous knowledge of the process, with an important reduction in the number of estimations of the process derivatives, confirming the idea of applying the best update law depending of the system.

From the point of view of the implementation of the nested methodology, it seems to be easier to apply it comparing with the gradient-based approach since it is not necessary to: take care about the degree of excitation of the process and impose additional constraints into the model.

With respect to the performance of the nested approach under noisy scenarios, it can be observed from the tested examples that its evolution towards the optimum of the process seems to be almost insensitive, unlike the gradient based approach that is affected in more or less magnitude depending on the degree of excitation of the system. This result is quite expected considering that if the process measurements are contaminated with noise, their gradients can amplify the noise from the process depending on the size of the change employed, making critical the selection of an adequate level of excitation of the system. On the other hand, an algorithm based in direct search uses directly the measurements to update the decision variables, meaning that the size of the process noise remains unaltered, making the nested methodology more insensitive to the contaminations of the measurements. These ideas also can be viewed from the point of view of selecting the best update method depending on the system: for those kinds of systems where the measurements are

contaminated with noise, it is recommended the use of a direct search algorithm or a gradient based one with enough energy to estimate accurately the process gradients.

As a summary, we have presented a method with a formal proof that is able to converge to the real optimum of the process in the context of RTO with modeling mismatch, in a similar way than the modifier adaptation would do, but without the need to take care about the excitation of the system and the estimation of the experimental gradient, assuming the convergence of the gradient-free algorithm implemented in the upper layer.

3.7 Modifier Adaptation Methodology in the Context of Dynamic Optimization

As it was mentioned in the introduction of this chapter, the hierarchical structure given by the separation of the problem according to their temporal scales is a very useful way of thinking in terms of ordering the different goals that an industrial facility must fulfill.

Nevertheless, there are emerging problems derived by the fact that using different time scales also implies using not the same models in each layer in terms of details and scope of application. Focusing our attention in the interaction among the RTO and the control layer, these problems can be summarized as inconsistency of both models and suboptimal operations due to the steady-state assumption to perform the optimization. In terms of the inconsistency of both models, it can be said that as a product of this fact, suboptimal and even infeasible points can appear by the incapability of the control scheme to follow the references given by the upper layer. On the other hand, the steady-state assumption can generate also suboptimal operating points, but now because of the possible disturbances that can affect the system during the transient, which can change the location of the real optimum, being necessary waiting until the next steady state to detect these changes and take the correcting actions.

To face these problems, several authors coincide in the idea of giving dynamic characteristics to the process optimization layer, transforming it in a DRTO stage (D from dynamic) (De Prada, 2004; Engell, 2007, 2009; Helbig et al., 2000; Kadam et al., 2002; Würth et al., 2009, 2011). The differences among the applications of the DRTO fundamentally lay in its frequency of execution, which is translated in keeping the two layers of optimization or merging them, for the cases of low and high frequencies respectively.

In any of the cases, it is important to study the effect of the modeling mismatch in terms of the ability to converge to the real optimum of a process in the DRTO layer. To do this, in this section we have tried to apply the ideas of the modifier adaptation method into a dynamic optimization problem.

3.7.1 Modified Dynamic Optimization

This is not a new concept and it was already presented in the work of Becerra and coworkers in the context of the DISOPE algorithm (Becerra, 1994; Becerra & Roberts, 1996; Becerra et al., 1998; Roberts & Becerra, 1999). In these works, the authors use a procedure analogous to the ISOPE method to solve a complicated NMPC problem by using an easier equivalent. To do this, it is necessary to use some modifiers in the objective function that take into account the differences in the partial derivatives between the complicated model and the simplified one.

Taking this into account, and considering that the infinite dimensional dynamic optimization of the process from equation (3.92), can be transformed into an NLP equivalent by parameterizing the decision variables and using a sequential approach, the implementation of the modifiers in the finite dimensional equivalent is quite immediate as it was shown in the example of the semi-batch reactor from section 3.6.3.5.

$$\begin{aligned}
&\min_{u(t)} F \coloneqq \bar{f}\left(u, x, y, t_f\right) \\
&s.t.: \\
&\bar{h}(\dot{x}, x, y, u, t) = 0, \qquad x(t_0) = x_0 \\
&\bar{g}\left(u, x, y, t_f\right) \leq 0 \\
&u(t) \in U
\end{aligned} \tag{3.92}$$

Where the bar "‾" indicates the quantities from the process t_0 and t_f are the initial and final time and the rest of the variables have the same meaning than the ones already presented previously. After the discretization step in the decision (and the state) variables, the NLP equivalent of equation (3.92) can be expressed as equation (3.93)

$$\begin{aligned}
&\min_{p \in P} F \coloneqq \bar{f}\left(p, t_f\right) \\
&s.t.: \\
&\bar{g}\left(p, t_f\right) \leq 0 \\
&P = [p^L, p^U]
\end{aligned} \tag{3.93}$$

Being p the time independent parameters used in the discretization of u. The KKT conditions of this equation are analogous to equation (3.14), therefore the correctors needed to modify the NLP equivalent from equation (3.93) have the same definitions than the steady-state modifier adaptation methodology.

Even when the application of the modifiers is quite similar to the steady-state case, their implementation implies that process variables must be measured at t_f, that is to say, it is necessary to know the value of $\bar{f}\left(t_f\right)$ and $\bar{g}\left(t_f\right)$. This requirement implies that if we are interested in the iterative implementation of the DRTO layer using a receding horizon approach, it is necessary to measure or estimate the process variables at the end of the prediction horizon to know the value of the objective function and the inequality constraints, which is a very restrictive assumption. Therefore, the methodology shown in this section can be applied only in a very small set of cases that present this condition.

If we define the iterative implementation of the model-based DRTO layer in its k^{th} iteration applied every Δt_{DRTO} as:

$$\begin{aligned} &\min_{p} F \coloneqq f\left(p, \alpha, t_{(k+1)}\right) \\ &s.t.: \\ &g\left(p, \alpha, t_{(k+1)}\right) \leq 0 \\ &p \in P, \qquad t \in [t_k, t_{k+1}) \end{aligned} \tag{3.94}$$

Being $\Delta t_k \coloneqq t_{k+1} - t_k$ the prediction horizon and α the parameters of the model.

The requirement of measuring the process variables at final time, implies that $\Delta t_{DRTO} = \Delta t_k$, which is one of the limitations of the proposed approach. Taking this into account, we can apply the modifiers to the model-based optimization as:

$$\begin{aligned} &\min_{p} F_m \coloneqq f\left(p, t_{(k+1)}\right) + \lambda_k^T p \\ &s.t.: \\ &G_m \coloneqq g(p, \alpha, t_{k+1}) + \gamma_k (p - p_{k-1}) + \epsilon_k \leq 0 \\ &p \in P, \qquad t \in (t_k, t_{k+1}] \\ &\\ &with: \\ &\lambda_k \coloneqq \frac{\partial \bar{f}_k}{\partial p} - \frac{\partial f_k}{\partial p} \\ &\gamma_k \coloneqq \frac{\partial \bar{g}_k}{\partial p} - \frac{\partial g_k}{\partial p} \\ &\epsilon_k \coloneqq \bar{g}_k - g_k \end{aligned} \tag{3.95}$$

Where the notation used in equation (3.95) can be understood as $y_k \coloneqq y(t_k)$ and $\bar{y}_k \coloneqq \bar{y}(t_k)$, being $y = \{f, g\}$.

3.7.2 Estimation of the Process Gradient

The gradients of the process can be estimated using a perturbation based method in order to quantify the effect that a variable has on the process measurements at t_k. Considering the time evolution of the process variable

$\bar{y} \in \mathbb{R}$ when the parameter p has been perturbed in Δp units from Figure 3.50, an approximation of the experimental gradient of the process would be calculated as equation (3.96) shows, assuming enough excitation on the process.

$$\frac{\partial \bar{y}_k}{\partial p} \approx \frac{\bar{y}_k - \bar{y}_{k-1}}{p_{k-1} - p_{k-2}} = \frac{\Delta \bar{y}_k}{\Delta p_k} \tag{3.96}$$

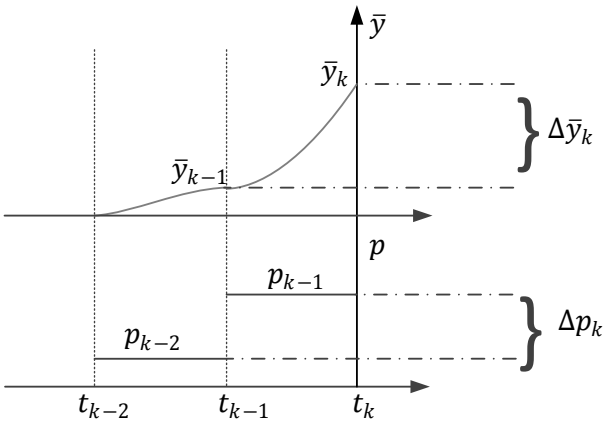

FIGURE 3.50: DIAGRAM OF THE EVOLUTION OF THE PROCESS VARIABLE y FOR A CHANGE IN p

Equation (3.96) corresponds to the finite differences implementation of the perturbation based method used in the steady-state approach. It assumes that the gradient of y with respect to p can be estimated as the ratio of the finite change in these variables keeping constant the rest of the independent variables. However, in a dynamic system the trajectory of the dependent variables, and by addition their values at discrete times, do not depend only on the decision variables but also on their natural dynamic response. This idea is very easy to understand considering the extreme case when $\Delta p = 0$. Under this scenario, and supposing that the process has not reached the steady-state, $\bar{y}_{k-1} \neq \bar{y}_k$ indicating that there is another dependence that must be considered in equation (3.96).

For a period Δt_k small enough to ensure $\dot{x}(t_i) \cdot \dot{x}(t_j) > 0, \forall t_i \neq t_j$, $\{t_i, t_j\} \in \Delta t_k$, a given system is not in steady-state when the state variables $\bar{x}_{k-2} \neq \bar{x}_{k-1}$ for $\Delta p = 0$. Taking this into account, we can understand the natural dynamic response of a system in the time point t_k as a dependence on the change of its initial states.

Therefore, the evolution of $\bar{y}(t_k)$ can be stated as:

$$\bar{y}_k = \bar{\phi}_k \coloneqq \bar{\phi}(p_{k-1}, \bar{x}_{k-1}) \tag{3.97}$$

To estimate an expression for the gradient of the dynamic process with respect to the parameter p we can apply a Taylor expansion to equation (3.97) around $\bar{y}_{k-1}$.

$$\bar{y} = \bar{y}_{k-1} + \frac{\partial \bar{\phi}_{k-1}}{\partial p}(p - p_{k-2}) + \frac{\partial \bar{\phi}_{k-1}}{\partial x}(\bar{x} - \bar{x}_{k-2}) + o^2 \tag{3.98}$$

Replacing equation (3.98) with the measured and applied values at t_k, the gradient of the process can be obtained neglecting the higher order terms, assuming that the value of Δt_k is small enough to represent the dependence of ϕ with respect to p and x as a lineal equation.

$$\begin{aligned} &\bar{y}_k \approx \bar{y}_{k-1} + \frac{\partial \bar{\phi}_{k-1}}{\partial p}(p_{k-1} - p_{k-2}) + \frac{\partial \bar{\phi}_{k-1}}{\partial x}(\bar{x}_{k-1} - \bar{x}_{k-2}) \\ &\Longrightarrow \frac{\partial \bar{\phi}_k}{\partial p} \approx \frac{\partial \bar{\phi}_{k-1}}{\partial p} \approx \underbrace{\frac{\bar{y}_k - \bar{y}_{k-1}}{p_{k-1} - p_{k-2}}}_{A} - \underbrace{\frac{\partial \bar{\phi}_{k-1}}{\partial x}\left(\frac{\bar{x}_{k-1} - \bar{x}_{k-2}}{p_{k-1} - p_{k-2}}\right)}_{B} \end{aligned} \tag{3.99}$$

Equation (3.99) can be viewed as an extension of the finite-difference approach used in the steady state RTO implementation, since the right hand side term named “A” is the same than equation (3.96), whereas the term “B” can be understood as a dynamic correction of the gradient as a

consequence of the changes in the states because of the unsteady behavior of the system.

Assuming that the states can be measured, it is necessary to estimate $\frac{\partial \phi_{k-1}}{\partial x}$ in order to calculate the derivatives with respect to the decision variables, and therefore, the modifiers. This value must be calculated using the variation of the process measurements obtained when $\Delta p = 0$. Figure 3.51 shows the trajectory of the output and the state for $t \in (t_{k-1}, t_k]$. Over this period p remains constant, therefore the changes in y are only due to the evolution in the states of the system.

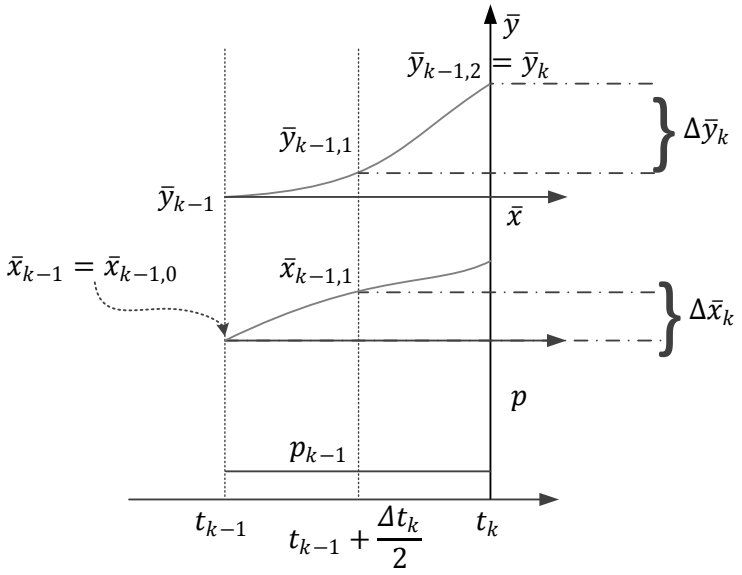

FIGURE 3.51: DIAGRAM OF THE EVOLUTION OF THE PROCESS VARIABLE y AND THE STATE x WHEN p IS CONSTANT

For a process with one state, the diagram from Figure 3.51 suggests that if it is possible measuring the output and the state variable at an intermediate point between t_{k-1} and t_k, an estimation of the gradient $\frac{\partial \phi_{k-1}}{\partial x}$ can be

calculated as the double of the ratio among the change in $\bar{y}$ over the variation of $\bar{x}$, when p remains constant as equation (3.100) shows.

$$\frac{\partial \bar{\phi}_{k-1}}{\partial x} \approx 2\frac{\bar{y}_k - \bar{y}_{k-1,1}}{\bar{x}_{k-1,1} - \bar{x}_{k-1}} = 2\frac{\Delta \bar{y}_k}{\Delta \bar{x}_k} \tag{3.100}$$

It is necessary multiply the ratio by 2 because when Δt_k is split in a half to measure the intermediate point is equivalent to estimate $\frac{\partial \bar{\phi}_{k-1/2}}{\partial \bar{x}}$.

If the system presents n states, the equation (3.100) can be generalized cutting Δt_k in $n+1$ parts, estimating $\frac{\partial \bar{\phi}_{k-1/n+1}}{\partial x}$ using the directional derivatives approach used in the steady-state RTO presented in section 3.3, applied into the dynamic problem as:

$$\begin{aligned} &\frac{\partial \bar{y}_j{}_k}{\partial x} = (S_k)^{-1}\begin{bmatrix} \bar{y}_{k,n} - \bar{y}_k \\ \vdots \\ \bar{y}_{k,1} - \bar{y}_k \end{bmatrix} \\ &S_k = [s_{k,1} \quad \cdots \quad s_{k,n}]^T \\ &s_{k,i} = \bar{x}_{k-1,n+1-i} - \bar{x}_{k-1} \end{aligned} \tag{3.101}$$

It can be noted that, as the dual approach, it is necessary to ensure a good condition number of matrix S_k for equation (3.101), which is in the same direction of estimating accurately the gradients with respect to p imposing enough excitation on the system. In addition we can complement the possibilities to estimate the gradients with respect of the states using local dynamic model identification with the measured data in the intermediate points.

3.7.3 Nested Modifier-Adaptation in Dynamic Optimization

The nested approach presented in section 3.6 can also be extended for the case of dynamic optimization. Due to the fact that there is no need to estimate the process gradients with respect to the decision variables, it

seems to be a good alternative in this context. As it was presented in example 3.6.3.5, its application is immediate in terms of modify the NLP equivalent of the dynamic optimization. However, the assumption of measuring the cost function from the process at the end of the time horizon cannot be removed owing to the fact that this value is the performance index to update the modifiers.

In RTO we have shown that the improvement in the performance index measured from the process can be used as a criterion to change the value of the modifiers. This idea, has the implicit assumption that the changes in the process cost function are produced mainly because of the variation in the value of the decision variables applied, which, as it was explained in previous section, is a consequence of the decision variables proposed by the upper layer. Therefore, implicitly there is a one to one relation between the modifiers and the value of the cost function. In spite of important disturbances, this is true for every steady state reached by the process. Nevertheless, as it was pointed in section 3.7.2, under dynamic considerations the evolution of the system is also a consequence of its natural dynamic response. Therefore, the performance index to update the modifiers in the upper layer must reflect the effect that the value of the modifiers has on this variable. To obtain this, we can focus our attention in equation (3.99), where the evolution of the process variable $\bar{y}_k$ can be viewed as a contribution of two factors: p_{k-1} and $\bar{x}_{k-1}$, which also means that the change in $\bar{y}_k$ with respect to $\bar{y}_{k-1}$ is a function of the same contributions:

$$\Delta\bar{y}_k \coloneqq \bar{y}_k - \bar{y}_{k-1} \approx \frac{\partial \phi_{k-1}}{\partial p}(p_{k-1} - p_{k-2}) + \frac{\partial \phi_{k-1}}{\partial x}(x_{k-1} - x_{k-2}) \qquad (3.102)$$

Being the first term on the RHS of equation (3.102) the contribution of a change in p and the second term the influence of the natural response. Consequently, the expected change in $\bar{y}$ as a result of a modification in p can be obtained as:

$$\Delta\bar{y}_k^p = \Delta\bar{y}_k - \frac{\partial\phi_{k-1}}{\partial x}(x_{k-1} - x_{k-2}) \tag{3.103}$$

Equation (3.103), as well as equation (3.99) in the gradient-based approach, is an extension of the performance evaluation of the process in steady-state. In addition, the objective of the optimization problem is minimizing $\bar{f}$; therefore, this expression also must be evaluated in the performance index to be used in the upper layer. Taking these two objectives into account, we can define the value of the cost function that the upper layer must use to update the modifiers as:

$$\bar{f}_k^{upper} = \omega\bar{f}_k + \Delta\bar{f}_k^p \tag{3.104}$$

Being $\Delta\bar{f}_k^p$ the variation of the cost function measured from the process due to the change on p and ω a weighting coefficient. The term $\Delta\bar{f}_k^p$ can be calculated from equation (3.103).

Correcting the cost function in equation (3.104), implies the use of the gradients with respect to $\bar{x}_k$ that can be obtained as it was pointed previously.

Regarding the modifiers given by the upper layer in the iterative implementation of a receding horizon approach, it is important to note that we are looking for a trajectory of the modifiers, not for a fixed optimal value as in steady state approach. Considering that a direct search algorithm gives only static values to be evaluated in the objective function (the process), makes necessary parameterizing the modifiers as a curve in time, being the decision variables given by the upper layer the time invariant parameters for this trajectory, represented as P_λ and P_γ in equation (3.86).

$$\lambda(t) = f_\lambda(P_\lambda, t)$$
$$\gamma(t) = f_\gamma(P_\gamma, t) \tag{3.105}$$

The proposed implementation of the Nested approach in DRTO is summarized in Figure 3.52.

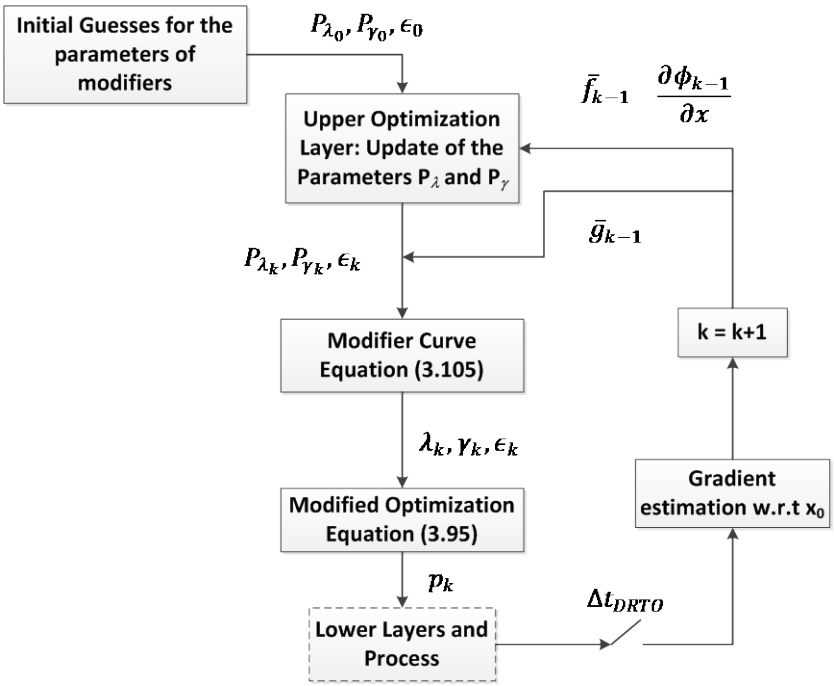

FIGURE 3.52: DIAGRAM OF THE IMPLEMENTATION OF NESTED APPROACH IN DRTO

3.7.4 Application example

The proposed implementation of the modifiers in the context of a receding horizon problem has been applied in a very simple example.

It consist in the tank of Figure 3.53, with two sources of inlet water flow F_{in}^1 and F_{in}^2 and an outlet flow F_{out}. The idea is keeping the water level h in a reference value h_{SP}. To achieve this goal we can implement a model-based optimization that minimizes the distance between the actual liquid level

and its reference, manipulating the inlets flow F_{in}^1 changing the voltage (u) applied to the pump P_1. Equation (3.106) shows the corresponding optimization problem, where we have used a first principle approach to model the system.

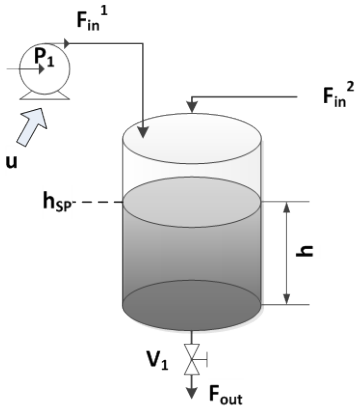

FIGURE 3.53: DIAGRAM OF THE LEVEL CONTROL IN THE TANK

$$
\begin{aligned}
&\min_{u(t)} f \coloneqq e(t_{k+1}) \\
&A\frac{dh}{dt} = F_{in}^1 + F_{in}^2 - F_{out}(t), \qquad h(t_k) = h_0 \\
&\frac{de}{dt} = (h(t) - h_{SP})^2, \qquad e(t_k) = 0 \\
&F_{out}(t) = a_V\sqrt{h(t)} \\
&F_{in}^1 = wu(t) \\
&t \in (t_k, t_{k+1}] \\
&u \in [u^L, u^U]
\end{aligned}
\tag{3.106}
$$

Here e is the accumulated error of the controlled variable, A corresponds to the transversal area of the tank, a_V is the constant of the outlet valve V_1 and w is the constant of the pump to convert the applied voltage into flow. To solve the infinite dimensional optimization problem from equation (3.106), we have parameterized the decision variable with a zero order function: $F_{in}^1(t) = p_F$ being p_F a constant value.

The nomenclature used in this example, as well as the value of the parameters are summarized in Table 3.13 and 3.14 respectively.

TABLE 3.13 NOMENCLATURE EMPLOYED IN THE DYNAMIC TANK

Variable	Meaning	Units
F	Volumetric flow	l/min
A	Transversal area of the tank	cm^2
a_V	Constant of the valve V	$cm^{2.5}/min$
h	Liquid height	cm
h_{SP}	Reference of the liquid height	l/min
e	Error of the liquid height w.r.t. its reference	cm^2
w	Pump constant	$l/minV$
u	Voltage applied to the pump	V
u^L, u^U	Bounds of u	V

TABLE 3.14 VALUE OF THE PARAMETERS EMPLOYED IN THE DYNAMIC TANK

Parameter	Value
A	154
a_V	0.1203
F_{in}^2	10
w	13.22
u^L	0
u^U	8

The simulated mismatch to be applied in this example corresponds in a change in the value of the constant of the valve of the process, while the model-based optimization uses its nominal value from Table 3.14.

$$a_V^{Process} = \frac{a_V}{\beta} \tag{3.107}$$

Two extreme values of β has been used to illustrate the effect of the modeling mismatch on the controlled process. Figure 3.54 shows the evolution of the system starting from $h(t_0) = 5\ cm$, applying the outcomes

from the model based optimization with no corrections. It can be noted that the system does not reach the optimum of the process ($h = h_{SP}$).

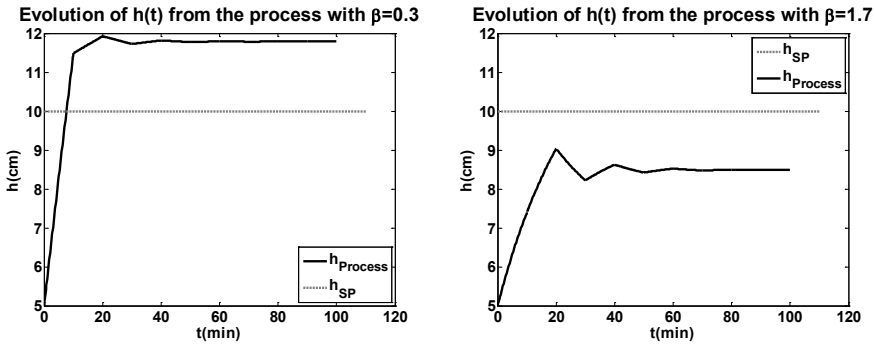

FIGURE 3.54: EVOLUTION OF THE LIQUID HEIGHT FROM THE PROCESS APPLYING THE MODEL BASED OPTIMIZATION WITH MISMATCH

The gradient-based modifier-adaptation methodology has been applied in the example, correcting the gradient of the objective function using λ. The evolution of the liquid height starting from $h(t_0) = 5cm$ is presented in Figure 3.55, for the same values of β already used in Figure 3.54.

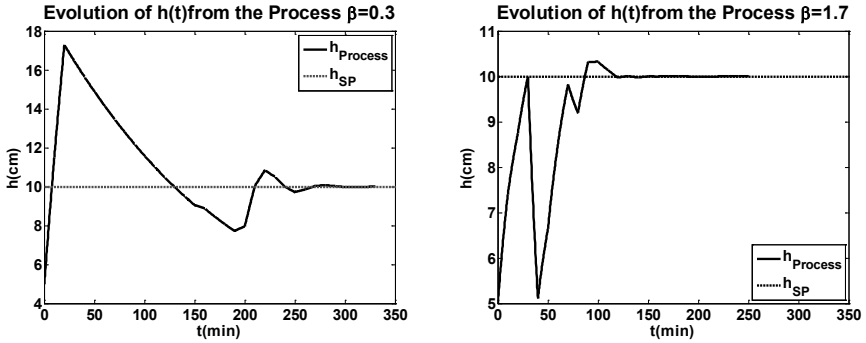

FIGURE 3.55: EVOLUTION OF THE LIQUID HEIGHT FROM THE PROCESS APPLYING THE MODIFIER-ADAPTATION METHODOLOGY

Figure 3.55 shows how the method is able to find the optimum of the process, iterating with the gradient of the process, for the two simulated mismatches. For the case of $\beta = 0.3$ it can be noted that the first two

iterations (from $t = 0$ to $t = 20$) the liquid height follows the tendency of the case with no correction, since for these points the value of λ has been set to 0 and the first gradient has been calculated. After these points, the system changes its direction reaching the desired set point. Similarly, in the evolution of $h(t)$ for $\beta = 1.7$ it can be noted that after the third iteration, the system presents corrective actions in order to find the optimum of the process with the process derivatives.

As it was mentioned in section 3.7.3 in dynamic implementation, the modifiers can be viewed as a trajectory, since the effect that they have depends on the time when they were applied. Figure 3.56 shows the path drawn by the modifiers in time, noting that they converge into a stationary point when the algorithm has detected the optimum of the process.

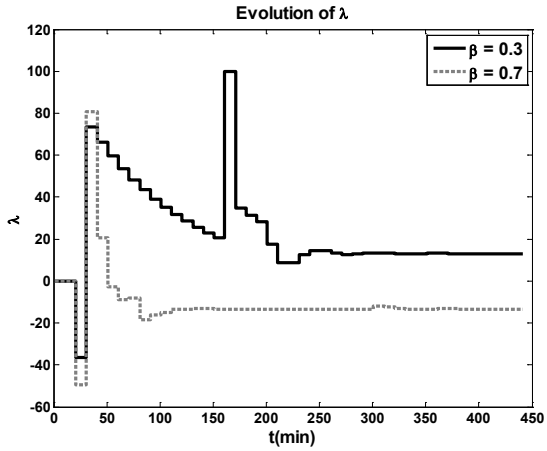

FIGURE 3.56: EVOLUTION OF λ IN THE MODIFIER-ADAPTATION METHODOLOGY

Starting from the same initial point $h(t_0)$, the nested approach has been implemented. Bearing in mind the evolution of λ from Figure 3.56 a first order dynamics has been implemented for f_λ (equation (3.108)).

$$\lambda(t_k) = \frac{P_\lambda^2}{P_\lambda^1} + \left(P_\lambda^3 - \frac{P_\lambda^2}{P_\lambda^1}\right)\exp\left(-\left|P_\lambda^1\right|t_k\right) \tag{3.108}$$

The reason why we use this function is because it converges into the stationary point $\frac{P_\lambda^2}{P_\lambda^1}$ when $t_k \to \infty$ and we can interpret the parameters P_λ^i as: the stationary point of λ, the expected time constant of the evolution of the modifier and its initial value. In spite of this, other functions might also be used provided their convergence into a stationary point.

The evolution of the controlled liquid height when the nested modifier-adaptation approach is applied in the example of the dynamic tank is represented in Figure 3.57.

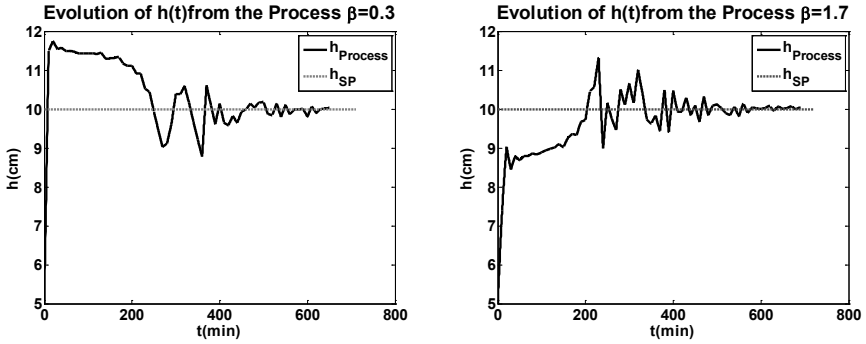

FIGURE 3.57: EVOLUTION OF THE LIQUID HEIGHT FROM THE PROCESS APPLYING THE NESTED MODIFIER-ADAPTATION METHODOLOGY

It can be noted that applying the corrections on the gradient of the model using the outer optimization, the system has converged into the real optimum for both cases tested. Regarding the time that the method requires to detect the process optimum, we can say that it has increased by a factor of two with respect to the gradient-based approach. This can be explained taking into account that the upper optimization has augmented the number of decision variables with respect to the number of modifiers,

which implies that the gradient-free algorithm needs more iterations to look for the optimum of the process.

Regarding the trajectory of the modifiers, it can be noted from Figure 3.58 that the path has reached at the end of the optimization the same values than the ones obtained with the gradient-based implementation.

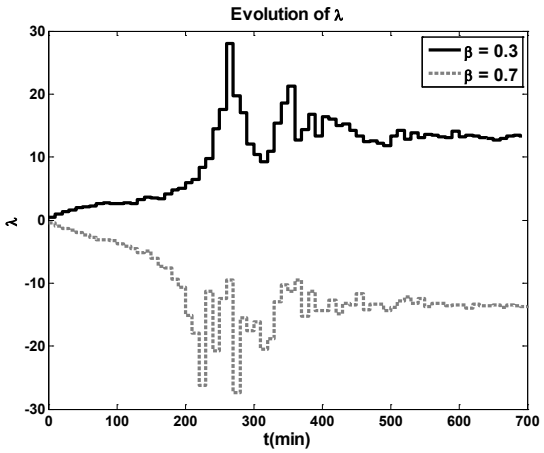

FIGURE 3.58: EVOLUTION OF λ IN THE NESTED MODIFIER-ADAPTATION METHODOLOGY

To test the behavior of the modified optimization under an incipient mismatch: we have increased the flow rate of F_{in}^{2} in the process following a step and a ramp function starting from t_D. With respect to the model, this value has remained unchanged at its nominal value (Table 3.14). Figure 3.59, shows the evolution of the liquid height for the gradient-based and the nested modifier adaptation. In the first column it is presented the response of the process under the step change, while the second column shows the same but for the ramp case. In the figure it is also presented the changes in the incoming flow in the last row.

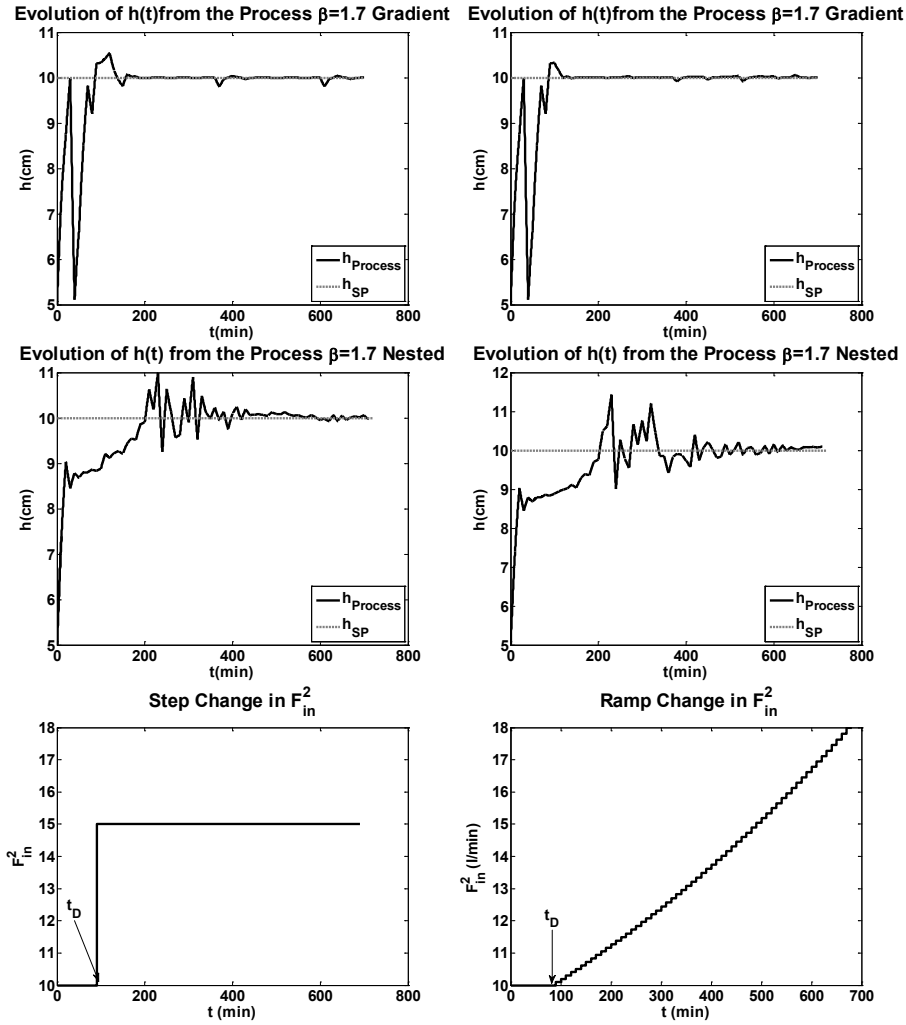

FIGURE 3.59: EVOLUTION OF THE PROCESS UNDER CHANGES IN F_{in}^2 FOR THE GRADIENT-BASED AND THE NESTED MODIFIER-ADAPTATION METHODOLOGY

As in the nominal values of F_{in}^2, we can note that the modifier adaptation methodology is able to detect a point closer to the optimum of the process both for the gradient-based and the nested approach.

Concerning the gradient-based methodology, the evolution of the liquid level changes with respect to the nominal case from Figure 3.55, just after the change in the influent is produced, being more notorious the effect of the step change. Nevertheless, the system immediately updates the corrections with the measurements and the process is driven into its optimum in a similar way than the previous example.

On the other hand, the nested methodology converges into a region closer to the optimum of the process, noting the presence of an optimality gap in the ramp scenario. This is produced because of the decrease in the size of the simplex when the objective function is improved, which is equivalent to a reduction in the search region as time passes and with this the possibilities to adapt it to new conditions. The effect of the disturbances in the evolution of the liquid height can be viewed as an increase in the perturbations that the NM algorithm uses to evaluate the new edges of the simplex after the occurrence of the disturbance.

Even though in both cases the nested optimization reached a region close to the real optimum of the process, it can be pointed out that its convergence is very sensitive with respect to its initial values, noting that if the sign of the initial iterant of P_{λ}^{2} is different to the optimal one, the optimization was not able to detect the optimum and it got stuck in a sub-optimal point.

3.7.5 Remarks on the Dynamic Implementation of the Modifier-Adaptation Methodology

The assumption of measuring the process quantities at the final time of each iterative implementation is one of the most restrictive constraints of this proposal because of the problems related with the stability that can be experienced. On the other hand, the need to estimate the gradients of the process with respect to the state variables, can also be a very difficult problem both from the point of view of implementation it (in general, the state variables cannot be measured) and the state estimation for systems

with many states, which requires enough energy to ensure either a well-conditioned matrix S or an accurate identification experiment. This can produce important problems in the stability of the system that might be unjustifiable with respect to the profit earned with the DRTO method.

Anyhow, we can say that for the system studied, with only one decision variable in the optimization, waiting until the final time to run another DRTO iteration, and with only one state the proposed approach seems to work in an adequate way, being possible to overcome the modeling mismatch tested and also adapting the system for changing uncertain conditions produced when the value of the disturbances are modified. This is based in the fact that the process has been driven from a feasible to its optimal point using only the corrections in the gradient of the cost function.

Regarding the nested implementation, the sensitivity of the upper layer with respect to the initial points, makes necessary to study other strategies to calculate the profit index to update the modifiers. On the other hand, the idea of iterating over the trajectory needs to be more developed in terms of its characteristic as well as the number of parameters required to define it because of the growth in the number of iterations required to converge.

Therefore, we can argue that for a very particular kind of system, we have presented a proposal to apply the modifiers ideas in the context of dynamic optimization. Far from being a general strategy to be implemented as the steady-state approach, the ideas presented in this section can be used as a starting point for future and more general applications.

3.8 Conclusions

Modifier-adaptation technique is a strong proposal for RTO to deal with the modeling mismatch that can be produced in the modeling stage. Since this is a fact that cannot be avoided because of the inherent partial knowledge of the process and/or the necessary simplifications that must be done in order to have a solvable model to be used in optimization, it is mandatory

its application (or the application of the other methods to overcome uncertainty in the model).

Even though this methodology allows converging into the real optimum of a process using an inaccurate model, there are open issues to be exploited. In this work we have tried to face three of them providing results that contribute to the advance in this field.

Regarding the violation of the constraints that can be produced in the evolution of the modifier method, we have presented the application of an intermediate stage between the RTO and the control layer that handles the occurrences of infeasibilities. With this methodology we can cope with them assuming enough control availability. Also, we have shown a second controller to take into account the dual constraint when the gradient of the process is estimated using the approximation of the directional derivatives, ensuring, if possible, an accurate estimation of the process derivatives. About the pairing of the manipulated set point and the inequality constraint using the steady-state gains seems to be a good idea, bearing in mind that there are two kinds of inequality constraints: primal and dual, being the first ones more important to keep them in the feasible region. With the application of the proposed controller we have avoided successfully the occurrences of infeasibilities generated as a consequence of the reduction of the original feasible region as a product of the dual constraint in a test example.

Under the assumption that the infeasibility is produced because of the RTO layer has detected that the optimum of the process lies on the inequality constraint, but the modification of the model is not good enough to estimate in a correct way the process constraint, the controller can also be understood as a way to follow the complementary slackness condition of the process.

With respect to the gradient estimation, we have shown an alternative approach, with the purpose of increasing its robustness and simplify its implementation. The idea behind is to substitute the errors in the gradients by a direct correction given by an upper gradient-free optimization layer, that can operate without constraints.

Regarding the comparison of the proposed nested methodology with the dual modifier adaptation, it can be said that, for the set of examples tested, the convergence to the real optimum is quite similar to the case when a relaxed dual constraint is employed for a noise free scenario. On the other hand, when the process measurements are contaminated with noise, the method behaves in a more robust way. Also it can be said that the implementation of the nested methodology is easier than the dual one, since one of the most sensitive parameters, the degree of excitation of the system to estimate accurately the gradient (the lower bound of the dual constraint) is neglected, as well as the additional non convex dual constraint that has no physical meaning in the real system.

Since the proposed methodology has found the real optimum of the process in a robust way, it makes the nested modifier-adaptation method an attractive alternative for real-time implementation.

Also, we can conclude that the reformulation of the modifier adaptation methodology as a nested optimization problem, allows understanding in a better way the concept of modifiers. This permits looking for additional alternatives to update the modifiers, with the idea of finding the best update algorithm that can be applied into the system, i.e., if the system has not important influences in the process noise and the process gradient is cheaper to obtain, then a gradient based methodology can be applied. However, if there is noise in the system and/or the gradient of the process is not available or is expensive to obtain, the use a direct search algorithm can be a better alternative, working in a similar way than the field of model

optimization does, where the best optimization algorithm depends on the type of the problem to be solved.

A proof of the equivalence of the optimality conditions between the gradient based modifier-adaptation and the nested modifier methodology has been provided.

Finally, concerning with the implementation of the modifier-adaptation technique in the context of dynamic optimization, we have proposed a very restricted approach to converge to the real optimum of the process in a receding horizon implementation for the gradient based and the nested approach. The main assumption to apply these ideas is the possibility to measure the process variables at the end of the time horizon, which can produce instability issues in many applications.

Anyway, for the example tested, we have shown that if we are able to measure the objective function at the end of the prediction horizon, using an estimation of the gradient of the objective function corrected with the natural evolution of the system it is possible to detect the real optimum of the process modifying the gradients of the objective function. On the other hand, the implementation of the nested approach correcting the performance index with the evolution of the system and the definition of the modifier trajectory, allows finding the optimum of the process, nevertheless, since the path of the modifiers must be parameterized it produces an increase in the number of decision variables, translated in a slow convergence rate to the stationary point.

3.9 References of the Chapter

Bamberger, W., & Isermann, R. (1978). Adaptive on-line steady-state optimization of slow dynamic processes. Automatica, 14, 223-230.

Becerra, V. M. (1994). Development and Applications of novel Optimal Control Algorithms. Unpublished PhD Thesis, City University, London.

Becerra, V. M., & Roberts, P. D. (1996). Dynamic integrated system optimization and parameter estimation for discrete time optimal control of nonlinear systems. International Journal of Process Control, 63, 257-281.

Becerra, V. M., Roberts, P. D., & Griffiths, G. W. (1998). Novel developments in process optimisation using predictive control. Journal of Process Control, 8, 117-138.

Biegler Lorenz, T. (2010). Nonlinear programming : concepts, algorithms, and applications to chemical processes: Society for Industrial and Applied Mathematics.

Biegler, L. T., Grossmann, I. E., & Westerberg, A. W. (1985). A note on approximation techniques used for process optimization. Computers & Chemical Engineering, 9, 201-206.

Bieker, H. P., Slupphaug, O., & Johansen, T. A. (2006). Real-Time Production Optimization of Offshore Oil and Gas Production Systems: A Technology Survey. In E. Society of Petroleum (Ed.), Intelligent Energy Conference and Exhibition.

Brdys, M., Chen, S., & Roberts, P. D. (1986). An extension to the modified two-step algorithm for steady-state system optimisation and parameter estimation. International Journal of Systems Science, 17, 1229-1243.

Brdys, M., Ellis, J. E., & Roberts, P. D. (1987). Augmented integrated system optimisation and parameter estimation technique: derivation, optimality and convergence. Control Theory and Applications, IEE Proceedings D, 134, 201-209.

Brdys, M., & Roberts, P. D. (1986). OPTIMAL STRUCTURES FOR STEADY-STATE ADAPTIVE OPTIMIZING CONTROL OF LARGE-SCALE INDUSTRIAL-PROCESSES. International Journal of Systems Science, 17, 1449-1474.

Brdys, M., & Roberts, P. D. (1987). Convergence and optimality of modified two-step algorithm for integrated system optimization and parameter estimation. International Journal of Systems Science, 18, 1305-1322.

Brdys, M., & Tatjewski, P. (1994). An algorithm for steady-state optimizing dual control of uncertain plants. In IFAC (Ed.), 1st IFAC Workshop on New Trends in Design of Control Systems. Smolenice, Slovakia: IFAC.

Brdys, M. A., & Tatjewski, P. (2005). Iterative algorithms for multilayer optimizing control. London: Imperial College Press.

Bristol, E. (1966). On a new measure of interaction for multivariable process control. Automatic Control, IEEE Transactions on, 11, 133-134.

Chachuat, B., Srinivasan, B., & Bonvin, D. (2009). Adaptation strategies for real-time optimization. Computers & Chemical Engineering, 33, 1557-1567.

Darby, M. L., Nikolaou, M., Jones, J., & Nicholson, D. (2011). RTO: An overview and assessment of current practice. Journal of Process Control, 21, 874-884.

De Prada, C. (2004). El Futuro del Control de Procesos. Revista Iberoamericana de automática e informática industrial, 1, 5-14.

Egea, J. A., Mar\, \#237, Rodr\, a., guez-Fern\, \#225, ndez, Banga, J. R., & Mart\, R. (2007). Scatter search for chemical and bio-process optimization. J. of Global Optimization, 37, 481-503.

Engell, S. (2007). Feedback control for optimal process operation. Journal of Process Control, 17, 203-219.

Engell, S. (2009). Online Optimizing Control: The Link Between Plant Economics and Process Control. In C. A. O. d. N. Rita Maria de Brito Alves & Evaristo Chalbaud Biscaia, Jr. (Eds.), Computer Aided Chemical Engineering (Vol. Volume 27, pp. 79-86): Elsevier.

Forbes, J. F., & Marlin, T. E. (1994). Model Accuracy for Economic Optimizing Controllers: The Bias Update Case. Industrial & Engineering Chemistry Research, 33, 1919-1929.

Forbes, J. F., & Marlin, T. E. (1996). Design cost: a systematic approach to technology selection for model-based real-time optimization systems. Computers & Chemical Engineering, 20, 717-734.

Forbes, J. F., Marlin, T. E., & MacGregor, J. F. (1994). Model adequacy requirements for optimizing plant operations. Computers & Chemical Engineering, 18, 497-510.

Gao, W., & Engell, S. (2005). Iterative set-point optimization of batch chromatography. Computers & Chemical Engineering, 29, 1401-1409.

Gattu, G., Palavajjhala, S., & Robertson, D. B. (2003). Are Oil Refineries Ready for Non-Linear Control and Optimization? In International Symposium on Process Systems Engineering and Control. Mumbai, India.

Golden, M. P., & Ydstie, B. E. (1989). Adaptive extremum control using approximate process models. AIChE Journal, 35, 1157-1169.

Gonzalez, A. I., Zamarreño, J. M., & De Prada, C. (2001). Non-linear model predictive control in a batch fermentator with state estimation. In European Control Conference. Porto, Portugal.

Halvorsen, I. J., Skogestad, S., Morud, J. C., & Alstad, V. (2003). Optimal selection of controlled variables. Industrial & Engineering Chemistry Research, 42, 3273-3284.

Helbig, A., Abel, O., & Marquardt, W. (2000). Structural concepts for optimization based control of transient processes. In F. Allgöwer & A. Zheng (Eds.), Nonlinear Model Predictive Control Workshop - Assessment and Future Directions (Vol. 26). Ascona, Switzerland: Birkhäuser Verlag, Basel.

Kadam, J. V., & Marquardt, W. (2004). SENSITIVITY-BASED SOLUTION UPDATES IN CLOSED-LOOP DYNAMIC OPTIMIZATION. In DYCOPS 7. Cambridge, USA.

Kadam, J. V., Schlegel, M., Marquardt, W., Tousain, R. L., van Hessem, D. H., van den Berg, J., & Bosgra, O. H. (2002). A two-level strategy of integrated dynamic optimization and control of industrial processes--a case study. In G. Johan & S. Jan van (Eds.), Computer Aided Chemical Engineering (Vol. Volume 10, pp. 511-516): Elsevier.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence Properties of the Nelder--Mead Simplex Method in Low Dimensions. SIAM J. on Optimization, 9, 112-147.

Levenspiel, O. (1999). Chemical reaction engineering (3rd ed.). New York: Wiley.

Lin, J., Chen, S., & Roberts, P. D. (1988). MODIFIED ALGORITHM FOR STEADY-STATE INTEGRATED SYSTEM OPTIMISATION AND PARAMETER ESTIMATION. IEE Proceedings D: Control Theory and Applications, 135, 119-126.

Lin, J., Hendawy, Z. M., & Roberts, P. D. (1988). EXTENSION OF INTEGRATED SYSTEM OPTIMIZATION AND PARAMETER-ESTIMATION TO HIERARCHICAL CONTROL OF STEADY-STATE SYSTEMS WITH OUTPUT-DEPENDENT CONSTRAINTS. International Journal of Control, 47, 413-431.

Luenberger, D. G., & Ye, Y. (2008). Linear and nonlinear programming. New York, NY: Springer.

Mansour, M., & Ellis, J. E. (2003). Comparison of methods for estimating real process derivatives in on-line optimization. Applied Mathematical Modelling, 27, 275-291.

Marchetti, A. (2009). Modifier-Adaptation Methodology for Real-Time Optimization. École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Marchetti, A., Chachuat, B., & Bonvin, D. (2008). Real-time optimization via adaptation and control of the constraints. In B. Bertrand & J. Xavier (Eds.), Computer Aided Chemical Engineering (Vol. Volume 25, pp. 393-398): Elsevier.

Marchetti, A., Chachuat, B., & Bonvin, D. (2009a). Modifier-Adaptation Methodology for Real-Time Optimization. Industrial & Engineering Chemistry Research, 48, 6022-6033.

Marchetti, A., Chachuat, B., & Bonvin, D. (2009b). Real-time optimization with estimation of experimental gradients. In S. Engell & Y. Arkun (Eds.), 7th IFAC International Symposium on Advanced Control of Chemical Processes (Vol. 7). Koç University Campus, Turkey: IFAC.

Marchetti, A., Chachuat, B., & Bonvin, D. (2010). A dual modifier-adaptation approach for real-time optimization. Journal of Process Control, 20, 1027-1037.

Mathworks, I. (2007). MATLAB Optimization Toolbox 6 user's guide. Natick, Mass.: The MathWorks.

Michalska, H., Ellis, J. E., & Roberts, P. D. (1985). JOINT COORDINATION METHOD FOR THE STEADY-STATE CONTROL OF LARGE-SCALE SYSTEMS. International Journal of Systems Science, 16, 605-618.

Navia, D., Marti, R., Sarabia, D., & De Prada Moraga, C. (2012). Laboratory Plant of a Gas Distribution Network. In P. Pakshin (Ed.), Advances in Control Education (Vol. 9). Resort Automobilist, Russia: IFAC.

Quijano, G., Hernandez, M., Villaverde, S., Thalasso, F., & Muñoz, R. (2010). A step-forward in the characterization and potential applications of solid and liquid oxygen transfer vectors. Applied Microbiology and Biotechnology, 85, 543-551.

Roberts, P. D. (1979). AN ALGORITHM FOR STEADY-STATE SYSTEM OPTIMIZATION AND PARAMETER-ESTIMATION. International Journal of Systems Science, 10, 719-734.

Roberts, P. D. (1995). Coping with model-reality differences in industrial process optimisation -- A review of integrated system optimisation and parameter estimation (). Computers in Industry, 26, 281-290.

Roberts, P. D. (2000). Broyden derivative approximation in ISOPE optimising and optimal control algorithms. In 11th IFAC workshop on control applications of optimisation (Vol. 1, pp. 283 - 288). Saint-Petersburg, Russia: IFAC.

Roberts, P. D., & Becerra, V. M. (1999). Model predictive control using dynamic integrated system optimisation and parameter estimation (DISOPE). In Model Predictive Control: Techniques and Applications - Day 1 (Ref. No. 1999/095), IEE Two-Day Workshop on (pp. 8/1-8/4).

Rodger, E. (2010). Dual Modifier Adaptation Methodology For the On-line Optimization of Uncertain Processes. Mc:Master University, Hamilton, Ontario, Canada.

Shobrys, D. E., & White, D. C. (2002). Planning, scheduling and control systems: why cannot they work together. Computers & Chemical Engineering, 26, 149-160.

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

Skogestad, S. (2000b). Self-optimizing control: the missing link between steady-state optimization and control. Computers & Chemical Engineering, 24, 569-575.

Srinivasan, B. (2007). Real-time optimization of dynamic systems using multiple units. International Journal of Robust and Nonlinear Control, 17, 1183-1193.

Tatjewski, P. (2002). Iterative Optimizing Set-Point Control - The Basic Principle Redesigned. In 15th Triennial World Congress, Barcelona, Spain: IFAC.

Treybal, R. E. (1980). Mass-transfer operations (3d ed.). New York: McGraw-Hill.

Valleur, M., & Grue, J. L. (2004). Optimize short-term refinery scheduling. Hydrocarbon Processing, 46-53.

Walters, F. H., Parker, L. R., Morgan, S. L., & Stanley, D. N. (1991). Sequential simplex optimization: a technique for improving quality and productivity in research, development, and manufacturing. Boca Raton, Fla.: CRC Press.

Wang, J. (2011). What is the outlook for advanced control engineering? . Hydrocarbon Processing, 46-53.

Williams, T. J., & Otto, R. E. (1960). A generalized chemical processing model for the investigation of computer control. AIEE Trans, 79, 458–473.

Woodward, L., Perrier, M., & Srinivasan, B. (2009). Improved performance in the multi-unit optimization method with non-identical units. Journal of Process Control, 19, 205-215.

Würth, L., Hannemann, R., & Marquardt, W. (2009). Neighboring-extremal updates for nonlinear model-predictive control and dynamic real-time optimization. Journal of Process Control, 19, 1277-1288.

Würth, L., Hannemann, R., & Marquardt, W. (2011). A two-layer architecture for economically optimal process control and operation. Journal of Process Control, 21, 311-321.

Young, R. (2006). Petroleum refining process control and real-time optimization. IEEE Control Systems Magazine, 26, 73-83.

Zhang, H., & Roberts, P. D. (1990). On-line steady-state optimisation of nonlinear constrained processes with slow dynamics. Transactions of the Institute of Measurement and Control, 12, 251-261.

Zhang, Y., & Forbes, J. F. (2006). Performance Analysis of Perturbation-Based Methods for Real-Time Optimization. The Canadian Journal of Chemical Engineering, 84, 209-218.

Zhang, Y., & Fraser Forbes, J. (2000). Extended design cost: a performance criterion for real-time optimization systems. Computers & Chemical Engineering, 24, 1829-1841.

Zhao, B., Li, Y., Tong, H., Zhuo, Y., Zhang, L., Shi, J., & Chen, C. (2005). Study on the reaction rate of sulfite oxidation with cobalt ion catalyst. Chemical Engineering Science, 60, 863-868.

4 FINAL CONCLUSIONS AND OPEN ISSUES

4.1 Final Conclusions

This work presents a study of how to handle the uncertainty present in the process optimization from the point of view of the stochastic nature of process variables and the modeling mismatch that can be experimented in the model-based optimization implemented in the RTO layer.

In chapter 2 we have presented the application of stochastic optimization for a hydrodesulfuration unit. With the aim to overcome the unknown and random behavior of the quality of the raw materials and the products to be treated, we have tested the implementation of two methodologies: chance constrained optimization and two-stage programming.

We have proposed an optimal policy to be applied in open loop that guarantees a given degree of feasibility in the process constraints, in spite of the uncertain behavior of the random variables, solving a dynamic optimization problem in the continuous domain of time. Regarding the implementation the conclusions are:

- In the two-stage optimization, a solution method has been proposed for the scenario aggregation methodology that allows obtaining a unique trajectory of the first-stage decision variables using the single shooting approach.

- The discrete optimal trajectories obtained with the scenario approach in the second stage of the two-stage optimization, can be efficiently generalized for the original distribution probability by means of an interpolation step tested with Montecarlo simulations. Using the interpolations and the Montecarlo simulations it is necessary to redefine the estimated value of perfect information, since the application of the decision variables of the second stage depends on the interpolation method chosen.

- The probability of fulfillment of the inequality constraints in the chance constrained optimization can be calculated with the inverse mapping technique, estimating the bounds of the integrals solving a parameter estimation problem with the single shooting approach.

- Both, two-stage and chance constrained optimization have been successfully applied in the hydrodesulfuration unit in terms of optimality and feasibility.

- Comparing the outcomes obtained with the two methods, it seems to be more efficient the use of the two-stage optimization. However, this requires an intermediate future step where the random variables must be estimated, unlike the chance constrained optimization where no additional information is required.

- Due to the high computational times observed, we have proposed the open loop execution based on an interpolation stage. Nevertheless, the implementation in a real application requires improving this time.

In chapter 3 we have studied the modifier-adaptation methodology as an efficient way to deal with the unavoidable modeling mismatch in model-based optimization, oriented to reach the optimum of a process. Also it has been discussed and proposed some improvements to overcome the implementation issues of this methodology in terms of: avoiding violations in the process constraints, skipping the estimation of the process gradient and applying the modifiers in dynamic optimization.

In this chapter we have proposed and tested methodologies that increase the field of application of this approach, making it feasible, robust and easier to apply in real conditions. From the specific objectives for this part of the thesis the conclusions are:

- We have studied and tested the generalities of the modifier adaptation methodology comparing them with the classic two-step approach, noting that the systematic correction of an erroneous model, and in particular its curvature, allows finding the real optimum of a process.

- Using previous analysis from literature and a test example, we have identified the main challenges of this methodology: the presence of infeasible points in its evolution, the problems related with the process gradient estimation step and the possibility to apply it in dynamic optimization.

- We have reduced the occurrence of infeasibilities in the evolution of the methodology, implementing an infeasibility controller that corrects the decision variables given by the RTO layer when the constraints are violated. Also we have proposed a second controller to estimate the process gradient accurately if needed.

- We have detected the optimum of the process without the need to estimate the process gradients, reformulating the modifier-adaptation methodology as a nested optimization problem using a gradient-free algorithm in the upper optimization layer.

- The reformulation as a nested problem allows increasing the application of the modifier-adaptation approach, using different update methods according to the characteristics of the system.

- We have tested the nested approach in several examples, noting that the method detects the optimum of the process in a similar way than the gradient-based approach when the process gradient is well estimated, but without the need to estimate the process derivatives which makes its implementation easier.

- Also, we have increased the field of application of the modifier-adaptation methodology for those kinds of systems where the gradient is not available or is affected with process noise, showing a robust behavior in these conditions. The methodology has also been tested in a runt-to-run batch application with good results

- Regarding the implementation of the modifiers in dynamic optimization, we have proposed some preliminary ideas, correcting the gradients of the process with its natural dynamic response. Also we have applied the same idea to the nested approach for the same kind of problems. This first approach can be viewed as a starting point for future work.

As a final conclusion we can mention that we have studied and proposed methods oriented to implement and improve the treatment of the uncertainties in process optimization, from the point of view of stochastic behavior of the variables and the partial knowledge of the process.

4.2 Open Issues

From the developments proposed in this thesis, it is necessary to comment what we think that it is necessary to improve and keep researching.

For the stochastic optimization, it is mandatory closing the loop of the two-stage and the chance constrained implementation, which only can be done if the computational times are reduced considerably. To do this, we plan to continue the study of reduction techniques that use previous values of the random variables to decrease its expected range in the next optimization. In the same way, other numerical environments can be used to reduce the computation times. Also, it is necessary testing these methodologies in a wider range of applications.

Concerning the applications for the modifier-adaptation method, we propose the study of other type of controllers to manage the infeasibilities together with the influence on the performance of this intermediate layer with respect to the pairing between the controlled inequality and the manipulated set-points.

In the nested modifier-adaptation it is necessary to extend the test with real world applications.

Finally, in the dynamic implementation of the modifier-adaptation and the nested approach for a receding horizon problem, it is compulsory removing the condition of waiting until the end prediction horizon and continue the research in order to extend the application to this important type of systems.

4.3 Conclusiones Finales

En este trabajo se ha presentado un estudio sobre el manejo de las incertidumbres en la optimización de procesos desde el punto de vista del comportamiento aleatorio de las variables, así como de los errores de modelado que pueden encontrarse en la capa de optimización en tiempo real.

En el capítulo 2 se ha presentado la aplicación de técnicas de optimización estocástica. Con el objetivo de manejar el comportamiento aleatorio de la calidad de las materias primas y de los productos a ser transformados, se ha probado la implementación de dos metodologías: optimización de dos etapas y optimización probabilística.

Se ha propuesto una política de operación óptima a ser aplicada en lazo abierto, que garantiza la factibilidad en la operación a pesar del comportamiento aleatorio de las variables de proceso, mediante la resolución de un problema de optimización dinámica en el dominio continuo del tiempo. Con respecto a la implementación, las conclusiones son:

- Para la aplicación de la optimización de dos etapas, la técnica de agregación de escenarios implementada con el método de single shooting ha permitido obtener una única trayectoria para la primera etapa.

- Las soluciones óptimas obtenidas para la segunda etapa utilizando la discretización de la función de probabilidad original, han sido generalizadas eficientemente mediante la propuesta de interpolación. Es importante destacar que esta generalización implica redefinir el concepto de valor estimado de información perfecta, puesto que la aplicación de las variables de decisión en la segunda etapa depende del método de interpolación escogido.

- Para la optimización probabilística, la probabilidad de las restricciones de desigualdad se ha calculado mediante el método de mapeo inverso, estimando los límites de integración en las variables aleatorias resolviendo un problema de estimación de parámetros mediante el método de optimización secuencial.

- Tanto la optimización de dos etapas como la de restricciones probabilísticas han sido aplicadas de manera satisfactoria en un ejemplo de una unidad de hidrodesulfuración, en términos de optimalidad y factibilidad.

- Comparando los resultados obtenidos con ambos métodos, parece ser más eficiente el uso de la optimización de dos etapas en este sistema, sin embargo este requiere un paso intermedio de estimación de las variables inciertas (fuera de línea), a diferencia de la optimización probabilística que no necesita información adicional.

- Debido a los tiempos computacionales observados, se ha propuesto la aplicación en lazo abierto utilizando el método de interpolación. Sin embargo, la implementación final requiere mejorar los tiempos de resolución.

En el capítulo 3 se ha estudiado el método de adaptación de modificadores como una técnica que permite detectar el óptimo de un proceso utilizando modelos erróneos, derivados a partir de un conocimiento parcial del sistema real. A partir de este estudio, se han identificado los desafíos que presenta este método, para los cuales se han propuesto algunas mejoras en términos de: evitar violaciones en las restricciones del proceso, proponer una alternativa que evite la necesidad de estimar el gradiente del proceso y probar la aplicación de los modificadores en optimización dinámica.

En este capítulo se han propuesto y probado métodos que incrementan el campo de aplicación de esta técnica, haciendo su implementación más robusta, factible y sencilla de aplicar para condiciones más cercanas a la realidad. Para los objetivos específicos de esta parte de la tesis, las siguientes conclusiones se pueden listar:

- Se ha estudiado la técnica de modificación de adaptadores, comparándola con el método clásico de RTO basado en dos etapas para un ejemplo de simulación. Los resultados indican que la corrección de un modelo (y en especial su curvatura), permite detectar el óptimo de un proceso.

- Mediante el análisis previo de la literatura y con la ayuda de un ejemplo, se han identificado algunos desafíos que requieren estudio en el método de adaptación de modificadores: infactibilidades en la evolución hacia el óptimo, problemas relacionados con la estimación del gradiente del proceso y posibilidad de aplicación en optimización dinámica.

- Se ha reducido la ocurrencia de infactibilidades en la convergencia del método, mediante la implementación de un controlador de infactibilidades que se activa cuando las restricciones son violadas, modificando el valor dado por la capa de RTO. Adicionalmente se ha implementado un segundo controlador (dual) que mantiene la excitación del sistema.

- Se ha implementado un método que permite detectar el óptimo del proceso sin estimar los gradientes del mismo, mediante la reformulación del método original como un problema de optimización anidado, que utiliza un algoritmo de optimización de búsqueda directa en la capa superior.

- La reformulación como una optimización anidada permite aumentar el campo de aplicación del método original, puesto que es posible escoger el método de actualización dependiendo de las características del sistema.

- Se ha probado el método de optimización anidada en varios ejemplos, mostrando como el método propuesto converge hacia el óptimo del proceso (o una región cercana a él) de la misma forma como lo haría el método basado en gradientes si es que las derivadas del proceso estuviesen bien estimadas. Lo anterior descartando la información de las derivadas del proceso, lo que adicionalmente simplifica su implementación.

- Adicionalmente, se ha ampliado la aplicación del método de adaptación de modificadores para aquellos sistemas en los cuales el gradiente no está disponible o está contaminado con ruido del proceso, lo que implica un comportamiento más robusto del método en condiciones reales.

- Referente a la implementación de los modificadores en la optimización dinámica, se han propuesto algunas ideas preliminares sobre la corrección de los gradientes del proceso utilizando la respuesta natural del sistema. Adicionalmente se ha implementado el método anidado para el mismo tipo de problemas. Este es un primer intento que puede servir como base en futuros desarrollos.

Como conclusión final, se puede indicar que se han estudiado y propuesto métodos que permiten el manejo de las incertidumbres y mejoran su tratamiento en la optimización de procesos, tanto desde el punto de vista

del comportamiento aleatorio de las variables de proceso, como de los errores de modelado derivados del conocimiento parcial del sistema real.

4.4 Trabajo Futuro

A partir de lo presentado en esta tesis, se cree necesario comentar los temas que se abren y que deben seguir explorándose.

Desde el punto de vista de la optimización estocástica, es imprescindible aplicar la optimización de dos etapas y la probabilística cerrando el lazo de control, lo cual sólo puede hacerse si es que se reducen los tiempos computacionales de manera considerable. Para realizar esto, se planea continuar la investigación con el estudio de técnicas que permitan reducir el rango esperado en las variables aleatorias en la optimización siguiente, a partir de las medidas disponibles del proceso. En la misma dirección, también se compararán las posibles mejoras obtenidas al probar otros programas computacionales. Adicionalmente se propone la implementación de estos métodos en otros ejemplos de optimización de procesos.

Para el caso de las aplicaciones propuestas en el campo de la optimización en tiempo real, se propone el estudio del desempeño de otro tipo de controladores en el corrector de infactibilidades así como de la influencia que tiene el emparejamiento de las desigualdades controladas con respecto a las consignas manipuladas.

En el método de optimización anidada, se propone continuar con las pruebas utilizando sistemas reales.

Por último, referente a la implementación dinámica para un problema de horizonte móvil, es completamente necesario buscar formas de eliminar la restricción de esperar hasta el final del horizonte de predicción y así continuar con la investigación referente a la aplicación de estos métodos en sistemas dinámicos.

CONTRIBUTIONS AND PUBLICATIONS

The list of contributions developed during the realization of this thesis is the following:

Journal Papers

- Navia, D., Sarabia, D., Gutierrez, G., Cubillos, F., & De Prada Moraga, C. (2012). A Comparison Between Two Methods of Stochastic Optimization for a Dynamic Hydrogen Consuming Plant. Computers & Chemical Engineering, Submitted on 31 Jan 2012, manuscript number: 7029 (Accepted with Corrections).

- Navia, D., Gutierrez, G., & De Prada Moraga, C. (2012). Real-Time Optimization Reformulated as a Nested Optimization Problem. Industrial & Engineering Chemistry Research, Manuscript ID: ie-2012-01176u. Submitted on 06 May 2012(Accepted with Corrections).

Congress Presentations

- Navia, D., Gutiérrez, G. & De Prada Moraga, C. Nested Modifier-Adaptation for Real Time Optimization in the Otto Williams Reactor. ECC 13 (Submitted).

- Navia, D., Marti, R., Sarabia, D., Gutierrez, G., & De Prada Moraga, C. (2012). Handling Infeasibilities in Dual Modifier-Adaptation Methodology for Real-Time Optimization. In V. Kariwala, L. Samavedham & R. D. Braatz (Eds.), 8th IFAC International Symposium on Advanced Control of Chemical Processes (Vol. 8). Furama Riverfront, Singapore: IFAC.

- Navia, D., Marti, R., Sarabia, D., & De Prada Moraga, C. (2012). Laboratory Plant of a Gas Distribution Network. In P. Pakshin (Ed.),

Advances in Control Education (Vol. 9). Resort Automobilist, Russia: IFAC.

- Navia, D., Gutierrez, G., & De Prada Moraga, C. (2012). Optimización en Tiempo Real de un Reactor Químico Usando Adaptación de Modificadores. In XXXIII Jornadas Españolas de Automática. Vigo, Spain: CEA.

- Navia, D., Sarabia, D., Gutierrez, G., & De Prada Moraga, C. (2010). Improving Scenario Stochastic Optimization for a Dynamic Hydrogen Consumption Plant. In XXXI Jornadas Españolas de Automática. Jaén, Spain: CEA.

- Navia, D., Sarabia, D., Gutierrez, G., & De Prada Moraga, C. (2010). A comparison Between Two Ways of Stochastic Optimization for a Dynamic Hydrogen Consuming Plant. In 3rd International Conference on Continuous Optimization (ICCOPT). Santiago, Chile.

- Navia, D., Gutierrez, G., & De Prada Moraga, C. (2009). Programación Estocástica en una Red de Distribución de Hidrógeno. In XXX Jornadas Españolas de Automática. Valladolid, Spain: CEA.

Other Publications

- Navia, D., Sarabia, D., Marti, R., Gómez, E., De Prada, C., Lucia, S., & Sonntag, C. (2011). Preliminary report on the state of the art and research needs in hierarchical coordination in energy systems and processes. In: Technical report of Highly-complex and networked control systems (HYCON2) European Project.

Other Publications unrelated with the Thesis

- Martí, R., Navia, D., Sarabia, D., & De Prada, C. (2012). Shared Resources Management by Price Coordination. In B. Ian David Lockhart & F. Michael (Eds.), Computer Aided Chemical Engineering (Vol. Volume 30, pp. 902-906): Elsevier.

- Marti, R., Navia, D., Sarabia, D., & De Prada Moraga, C. (2011). Control Predictivo Descentralizado de una Red de Oxígeno. In XXXII Jornadas Españolas de Automática. Sevilla, Spain: CEA.

- De Prada Moraga, C., Gutierrez, G., & Navia, D. (2011). Integrated process and control design of a chemical reactor. In 8th European Congress of Chemical Engineering. Berlin, Germany.

- Navia, D., Gutierrez, G., Cobas, P., & De Prada Moraga, C. (2010). New Features in Dynamic Optimization with EcosimPro. In X Simposio Internacional de Automatización. la Habana, Cuba.

- Navia, D., De Prada Moraga, C., & Cubillos, F. (2010). Heat Integration and Optimal HEN Design Applied to a Gas Pipeline / Power Generation System. In 19th International Congress of Chemical and Process Engineering CHISA 2010 – PRESS. Praga, Czech Republic

- Navia, D., De Prada Moraga, C., Gutierrez, G., & Cubillos, F. (2010). Synthesis of Multiperiod Heat Exchanger Networks Comprising Heat Exchanger Types. In 20th European Symposium of Computer Aided Chemical Engineering (ESCAPE 20). Naples, Italy.