

## Nested Modifier-Adaptation for RTO in the Otto Williams Reactor

Daniel Navia\*, Gloria Gutiérrez\*\*, César de Prada\*\*

\* Dpt. Ingeniería Química y Ambiental, Universidad Técnica Federico Santa María, Santiago, Chile, (e-mail: daniel.navia@usm.cl).

\*\* Dpt. Systems Engineering and Automatic Control, University of Valladolid, c/ Real de Burgos s/n, 47011, (e-mail: prada@autom.uva.es).

---

**Abstract:** This paper deals with the problem of uncertainty management in real time optimization (RTO). It proposes a new architecture in the modifier-adaptation methodology, reformulating the algorithm as a nested optimization problem with two layers. Using this approach, it is possible to find a point that satisfies the KKT conditions of a process using an inaccurate model, but unlike the original modifier method, with no need to estimate the experimental gradients of the process. The proposed method has been tested in the Otto Williams Reactor considering structural mismatches and perfect and noisy measurements. The results are compared with the previous modifier adaptation methodology using dual control optimization showing that the method finds a KKT point of the process with the advantage that no experimental gradient information is required and with less sensitivity to process noise.

*Keywords:* Process optimization, Real Time Optimization, modifier-adaptation, direct search, uncertainty

---

### 1. INTRODUCTION

Because of the complexity of industrial processes, finding the optimal operation point is not a trivial task due to the inherent difficulties of the process itself, as well as the uncertainties and disturbances that continuously modify the operating conditions or modelling errors. In highly automated plants, optimal operation is typically managed by a decision hierarchy with several levels including planning, real-time optimization (RTO) and process control. At the RTO level, medium-term decisions are made on a time scale of hours to a few days considering economic objectives.

RTO emerged in the late 1970's as a two-stage algorithm: i) parameter estimation and ii) economic optimization. In the first step, the uncertainties are taken into account updating the parameters  $\alpha$  of a nonlinear model, solving a parameter estimation problem (1), where  $u \in \mathbb{R}^{N_u}$  is the actual value of the manipulated variables and  $y \in \mathbb{R}^{N_y}$  represents the model outputs,  $f \in \mathbb{R}^{N_x}$  is the model of the process and  $g \in \mathbb{R}^{N_g}$  are the inequality constraints of the system. The superscript “ $\bar{\cdot}$ ” indicates in the entire document that the variable is measured from the process.

$$\begin{aligned} \min_{\alpha} \Phi_{id} &:= (y - \bar{y})^T R (y - \bar{y}) \\ \text{s.t.} &: f(x, y, u, \alpha) = 0, \quad g(x, u, \alpha) \leq 0 \end{aligned} \quad (1)$$

After the model update, an economic optimization is solved (2), obtaining a new set of decision variables. In (2),  $\phi \in \mathbb{R}$  represents an economic objective function. The solution of (2) is then applied to the process in an iterative scheme.

$$\begin{aligned} \min_u \Phi_{eco} &:= \phi(x, u, \alpha) \\ \text{s.t.} &: f(x, y, u, \alpha) = 0, \quad g(x, u, \alpha) \leq 0 \end{aligned} \quad (2)$$

Due to the uncertainty that affects the process, there is an unavoidable difference between the model used in the optimization and the process (model mismatch). This point, plus the type of interaction that takes place in the two-stage implementation, implies that the classical approach of RTO will not necessarily converge to a point that satisfies the KKT conditions of the process, noting that the optimum of the process cannot be reached using this procedure. With this in mind, Roberts solved the integrated system optimization and parameter estimation problem, adding an additional modifier to the economic optimization that takes into account the difference between the gradient of the cost computed with the model and the real one of the process (Roberts, 1979). Later, Tatjewski showed that the convergence to a KKT point of the process does not depend on the parameter estimation problem, but only on the equality between the output of the process and the model in each RTO iteration (Tatjewski, 2002). Following the same idea, Gao and Engell proposed an extension to handle process-dependent constraints, being necessary the estimation of the experimental gradients of the constraints too (Gao and Engell, 2005). Recently, Chachuat and co-workers have generalized this methodology in the *Modifier-adaptation method* (Chachuat et al., 2009)

Even though previous methods can find a point that satisfies the necessary conditions of optimality (NCO) of the process, it is mandatory to estimate the real gradients of the system, which is not a trivial task. To avoid this step, the modifier-adaptation has been reformulated in this work as a nested optimization procedure. This allows converging to a KKT point of the process in a similar way than the original method, but with no need to estimate the process gradient.

The paper is organized as follows. Section two presents a summary of the modifier adaptation method. Section three

shows the reformulation proposed. Section four describes the application of the methodology proposed in the Otto Williams Reactor comparing the results with the former procedure. At last, section five gives some concluding remarks.

## 2. MODIFIER ADAPTATION METHODOLOGY

The modifier adaptation problem, in order to guarantee convergence to a point that satisfies the KKT conditions of the process, changes problem (2) in to (3).

$$\begin{aligned} \min_{u \in U} \Phi_{\text{mod}} &:= \phi(x, u, \alpha) + \lambda_k^T u \\ \text{s.t.} & \\ f(x, y, u, \alpha) &= 0 \\ G_{\text{mod}} &:= g(x, u, \alpha) + \gamma_k^T (u - u_{k-1}) + \varepsilon_k \leq 0 \end{aligned} \quad (3)$$

Being  $\Phi_{\text{mod}} \in \mathbb{R}$  and  $G_{\text{mod}} \in \mathbb{R}^{N_g}$  the modified objective function and inequality constraints of the model.

The modifiers  $\lambda$ ,  $\gamma$  and  $\varepsilon$  directly correct the problem formulation (2) modifying the KKT conditions of the model, so that the NCO conditions of (3) match with the ones from the process in an iterative implementation. They are given by the errors in the cost function gradient, in the constraints gradients and the constraints respectively between the process and the model. Typically, a first order filter is applied to these modifiers for implementation (Marchetti et al., 2009). Equation (4) shows its definition, that uses information from the  $k-1^{\text{th}}$  RTO iteration of the algorithm:  $\lambda_k$  and  $\gamma_k$  are in charge of correcting the gradient of the objective function and the inequality constraints, while  $\varepsilon_k$  attempts to modify the value of the process dependent constraints in the  $k^{\text{th}}$  RTO iteration. Being computed with values from the previous iteration, they are considered constants when solving (3).

$$\begin{aligned} \lambda_k &:= (I - K^\lambda) \lambda_{k-1} + K^\lambda (\nabla_u \bar{\phi} - \nabla_u \phi(x_{k-1}, u_{k-1}, \alpha)) \\ \gamma_k &:= (I - K^\gamma) \gamma_{k-1} + K^\gamma (\nabla_u \bar{G} - \nabla_u G(x_{k-1}, u_{k-1}, \alpha)) \\ \varepsilon_k &:= (I - K^\varepsilon) \varepsilon_{k-1} + K^\varepsilon (\bar{\varepsilon} - \varepsilon(x_{k-1}, u_{k-1}, \alpha)) \end{aligned} \quad (4)$$

The modifiers depend both on the gap between process and model variables and the first order constant  $K$ , that tries to smooth the path to achieve the optimum.

### 2.2 Implementation

Fig. 1 shows the diagram of the implementation of the modifier adaptation methodology. It can be seen that for every iteration of the RTO layer, it is necessary to estimate the modifiers of the cost function and the constraints. The bias  $\varepsilon$  is not difficult to obtain, but the estimation of the gradient of the cost and constraints measured on the process is not always available and is the key issue of this methodology. Mansour and Ellis (Mansour and Ellis, 2003) present a compilation of methods for estimating these process gradients. In this work, the dual control optimization will be used to compare with the nested methodology proposed since it has the advantage that fewer perturbations are required to

estimate the process derivatives with respect to other methods (Gao and Engell, 2005).

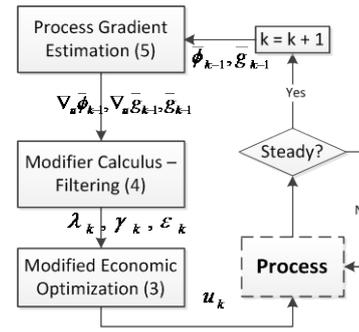


Fig. 1. Implementation of the Modifier Adaptation Methodology.

In dual methodology it is assumed that there are as many past values of the manipulated variables as the number of outputs. Then, by using the definition of directional derivative, the gradient of each measured output can be estimated as in (5) (Brdys and Tatjewski, 2005).

$$\begin{aligned} \nabla_u \bar{y}^j_{k-1} &\approx (S^{k-1})^{-1} \Delta \bar{y}^j_{k-1} \\ S^{k-1} &:= [s^{k-1,1} \quad \dots \quad s^{k-1,Nu}]^T \\ s^{k-1,i} &:= u_{k-1-i} - u_{k-1}, i = 1 \dots Nu \\ \Delta \bar{y}^j_{k-1} &:= [\bar{y}^j_{k-2} - \bar{y}^j_{k-1} \quad \dots \quad \bar{y}^j_{k-Nu-1} - \bar{y}^j_{k-1}]^T \end{aligned} \quad (5)$$

Due to the fact that this technique involves the inversion of the matrix  $S$ , the value of its condition number ( $\kappa$ ) is crucial in order to ensure appropriate estimations of the gradient. Therefore, it is necessary to add the additional non-convex constraint (6) to the general formulation of the modifier adaptation methodology (3), decreasing the original feasible region. The modified problem augmented with the constraint from (6) is the so called ‘‘Dual’’ problem, being  $\delta^{LO}$  a lower bound for the inverse of  $\kappa$ .

$$\delta(S^k) \geq \delta^{LO}, \quad d(\cdot) := \kappa(\cdot)^{-1} \quad (6)$$

## 3. PROPOSED NESTED METHODOLOGY

The iterative implementation of the modifier-adaptation methodology from Fig. 1 can be summarized as follows: for a given value of the modifiers, obtained with any of the methods available to calculate the process gradients, an optimization problem is solved using (3) in order to calculate and apply the next operating point to the process. Once the process has reached the new steady state, repeat this procedure until no further changes in the decision variables are observed, which implies, according the modifier adaptation assumptions, that a KKT point of the process has been found (Chachuat et al., 2009).

In principle, any policy for updating the modifiers could be used, provided that improve the process objective function and respect the constraints of the process. 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 optimization layer that uses the modifiers as decision variables to be applied over the inner modified optimization from (3) 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 the process gradients

This is the basic idea of the nested-modifier adaptation methodology (Fig 2) presented in this section and summarized in the following algorithm:

*Step 0:* Set  $k=0$  and start the algorithm with an initial guess of the modifiers  $\lambda_k$ ,  $\gamma_k$  and  $\varepsilon_k$ . Solve (3) and apply  $u_{k+1}$  into the process.

*Step 1:* Once the process steady state is reached, measure the value  $\bar{\phi}_k$  and  $\bar{g}_k$ . Compute  $\varepsilon_k = \bar{g}_k - g_k$

*Step 2:* Update the modifiers  $\lambda_k$  and  $\gamma_k$  by using the value of  $\bar{\phi}_k$  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:* Filter the modifiers given by the upper optimization layer using (4).

*Step 4:* Calculate the decision variable  $u_{k+1}$  solving the inner modified optimization problem from equation (3), using  $\lambda_k$  and  $\gamma_k$  and the value of  $\varepsilon_k$ . Set  $k=k+1$ .

*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  $\lambda$  and  $\gamma$  as the decision variables, and the cost function of the process as the objective function for the upper optimization layer. The value of  $\varepsilon$ , 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.

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 process derivatives in order to look for the real optimum.

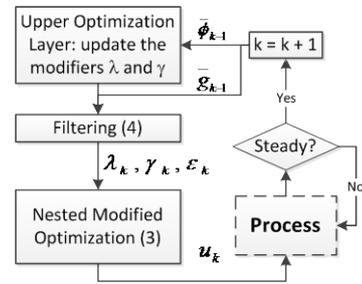


Fig. 2. Implementation of the Nested Modifier Adaptation Methodology.

The Nelder-Mead (NM) algorithm has been chosen in the upper layer, since it is particularly parsimonious in function evaluations per iteration. This property is very important considering that each function evaluation implies changing the operation point of the real process (Walters et al., 1991).

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. Also, another advantage of the proposed method is the fact that one of the most sensitive parameters to tune in modifier-adaptation methodology is neglected: the size of the perturbation to estimate the gradient, translated in the dual methodology as the specification of  $\delta^L$ , a number that is not easy to select a priori, implying simpler way to apply the modifier method compared with the original one.

### 3.1 Equivalence with the Original Modifier-Adaptation

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  $\varepsilon$ . 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. However, a summary of the equivalence with previous method can be commented here.

Starting from the KKT conditions of (3), it is clear that, under stationary assumption, the definition of  $\varepsilon$  allows to ensure primal feasibility of the process. On the other hand, dual feasibility is given by the solution of the nested problem. Also, complementary slackness for the inequality constraints of the process is given by the definition of  $\varepsilon$  and the dual feasibility.

Regarding the optimum of the measured objective function, it is possible to manipulate the complementary slackness condition and the gradient of the Lagrangean function of (3) from its KKT conditions to obtain:

$$(\nabla_u \phi - \lambda^T) \nabla_{\lambda} u = 0, \quad (\nabla_u \phi - \lambda^T) \nabla_{\gamma} u = 0 \quad (6)$$

Equation (6), holds for any value of  $\lambda$  and  $\gamma$ . In particular, if we replace these values with the definition of the modifiers (equation (4), setting  $K = 0$ ) and apply the chain rule, we get:

$$\nabla_{\lambda} \bar{\phi} = 0, \quad \nabla_{\gamma} \bar{\phi} = 0 \quad (7)$$

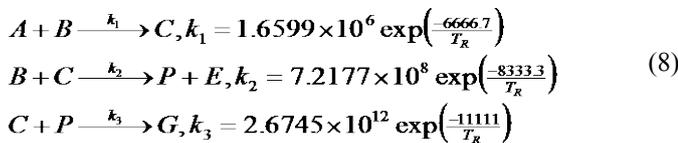
This means that the application of the modifier adaptation policy leads to a (local) unconstrained optimum of the process economic cost function with respect to the modifiers  $\lambda$  and  $\gamma$ , which is also the point given by the upper optimization layer of the nested approach after convergence, as it implements an unconstrained optimization of the process economic cost function having the modifiers  $\lambda$  and  $\gamma$  as decision variables.

#### 4. APPLICATION IN THE OTTO WILLIAMS REACTOR

In order to test the proposed algorithm, we have implemented the modifier-adaptation methodology in the Otto Williams Reactor (Williams and Otto, 1960). Two scenarios were considered in the RTO evolution: with and without noise in the process measurements. The implementation was compared with dual modifier-adaptation methodology for different values of  $\delta^O$ .

##### 4.1 Process Description

The Otto Williams Reactor is a CSTR with two sources of pure raw material  $A$  and  $B$  (with flows  $F_A$  and  $F_B$  respectively), being  $F_A$  a known disturbance and  $F_B$  a decision variable. The other manipulated variable in the system is the reactor temperature  $T_R$ , which can be modified using the heating/cooling system. Inside the vessel, three chemical reactions take place in parallel (8), forming in total 6 compounds: four products  $C$ ,  $E$ ,  $G$  and  $P$ , along with the unused reactants. This mixture leaves the reactor from the bottom of the vessel in a unique stream ( $F_R$ ). Fig. 3 presents a diagram of the example where  $X_i$  is the mass fraction of the “ $i$ ” compound inside the reactor.



The objective of the process optimization is to find a combination of the manipulated variables that achieves the maximal profit in steady state of the reactor operation, expressed as a function of the flow rate of the compounds (9). Here  $P_i$  and  $C_i$  are prices of the products and costs of the raw materials respectively.

$$\phi_{\text{Otto}} := F_R (X_P P_P - X_E P_E) - F_A X_A C_A + F_B X_B C_B \quad (7)$$

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 and Marlin, 1996), defines a model of the Otto Williams example to be used in the model-based optimization, neglecting the existence of the product  $C$  and supposing only two parallel reactions inside the reactor (8).

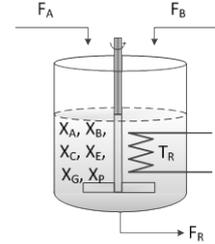
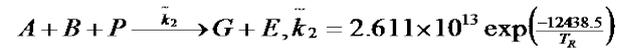
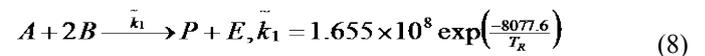


Fig. 3. Diagram of the Otto Williams Reactor.



We can test the Nested Modifier Adaptation method using a model based on (8) for computing the optimal operation point of the reactor and applying the results to a simulated reactor based on (6). The optimization should find the decision variables inside a feasible space that maximize the profit (7), subject to a wrong model with only five compounds and two chemical reactions, as in (9).

$$\begin{aligned} \min_{F_B, T_R} & \phi_{\text{Otto}} \\ \text{s.t.} & \end{aligned} \quad (9)$$

**Model with mismatch**

$$F_B^{Lo} \leq F_B \leq F_B^{Up}, \quad T_R^{Lo} \leq T_R \leq T_R^{Up}$$

The parameters, disturbances and bounds utilized in the model-based optimization and in the simulation of the process, are presented in Table 1.

**Table 1. Value of the Parameters and Bounds**

Parameter	Value	Units	Bound	Value	Units
$C_A$	76.23	€/kg	$F_B^{Lo}$	3	kg/s
$C_B$	114.34	€/kg	$F_B^{Up}$	6	kg/s
$F_A$	1.8275	kg/s	$T_R^{Lo}$	70	°C
$P_P$	1143.38	€/kg	$T_R^{Up}$	100	°C
$P_E$	25.92	€/kg			

#### 4.2 Results

##### Noise Free Scenario

The Nested and dual modifier adaptation methodologies were tested starting from the optimum of the model,

corresponding to point (A) in Fig 4. The starting values for the modifiers were obtained from finite differences.

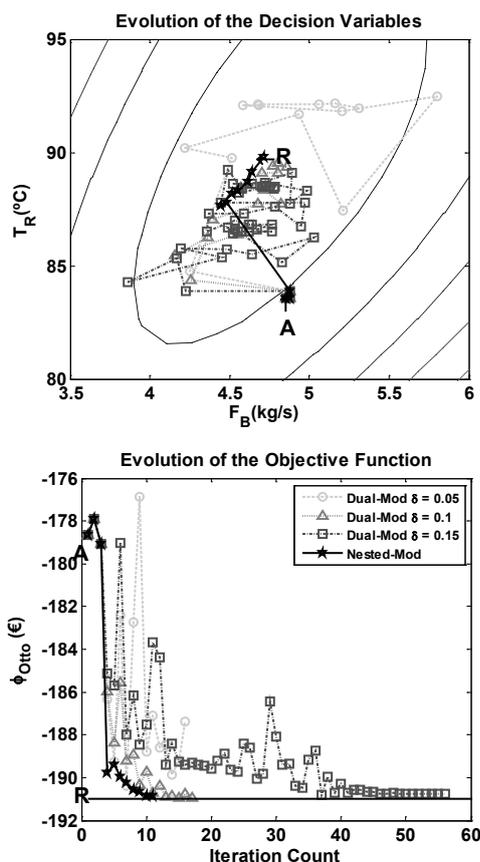


Fig. 4. Evolution of the decision variables and the objective function Dual and Nested-modifier adaptation for noise-free scenario

Fig.4 shows the evolution of the decision variables and the contours of the economic cost function. One can see how the nested modifier- adaptation (in black) converges to the real optimum of the process (R), in a similar way than the dual methodology (grey triangles and squares) does, from the starting point. The convergence to the real optimum is quite expectable after the analysis presented in section III where it was commented that, under convergence assumption of the upper layer, the results of the nested method would be similar to the real optimum.

Even though the proposed methodology presents a similar behaviour than the previous one, it can be noted that there is no need to take in to account the grade of excitation of the process to estimate the experimental gradient because is no longer needed. In the dual methodology, this is equivalent to choose  $\delta^{LO}$ , which is very sensitive to the evolution to the real optimum of the process as can be noted from Fig.4. In fact, if this value is equal to 0.05 the dual methodology does not converge to the desired point. As this lower bound increases, the gradient-based algorithm reaches the real optimum of the process, but increasing the number of the RTO iterations, which can be seen more clearly in the evolution of the objective function.

### Noisy Scenario

To test the performance of the proposed methodology under more realistic conditions, a band-limited white noise error was added in the measurements of the mass fractions. The maximum amplitude of the band was a 10% of the total range of the mass fractions. With this error, the measurement of the compound “i” from the process can be represented as  $X_i \pm 0.05$  (an expected resolution for a composition probe). In experimental setups, it is quite expectable the presence of noise in the measurements, therefore it is also common trying to face this problem considering more than one measurement in time. This step was also simulated filtering the random noise using the average of the last ten measurements obtained each 20s to compute the cost function of the process.

The evolution of the nested methodology, as well as the dual one with different values of  $\delta^{LO}$  is presented in Fig. 5.

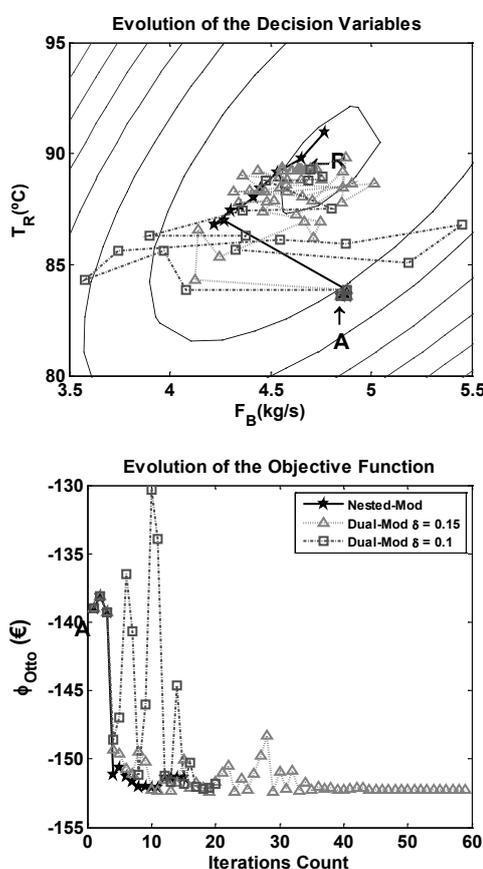


Fig. 5. Evolution of the decision variables and the objective function Dual and Nested-modifier adaptation for noise-free scenario

Analysing the path of the nested methodology, it can be noted that there is a wrong direction in the estimation of the modifiers at the beginning of the iterations due to the inaccuracy in the measurements. Nevertheless, the NM algorithm corrects the search direction updating the decision variables converging in to a point close to the process optimum. Regarding the number of iterations that the method needs, it can be said that the presence of noise increases this number; however, the inaccurate measurements do not affect

the evolution of the gradient-free based method, which is quite similar to the previous case. Comparing the evolution of the decision variables with the dual method, it can be distinguished how the gradient-based methodology seems to be more sensitive with the noise in the measurement, from the zigzag behaviour observed in the path described with dual optimization. This situation was critical for a value of  $\delta^{LO} = 0.05$  that is not presented in Fig. 5 because it diverged completely. Previous description can also be observed in the evolution of the objective function, where important jumps are detected in the gradient-based iterations.

#### Discussion of the Results

The evolution of the nested methodology under the two scenarios tested, shows that it converges close to the optimum of the process in a similar way than the dual methodology, but with no need to estimate the gradient of the process.

Neglecting the gradient information implies that there is no need to take care about the excitation of the process, which is an additional advantage regarding the implementation of the methodology. In the case of dual optimization, the evolution of the method seems to be very sensitive to this issue, which is reflected in the variations of the outcomes observed with the value of  $\delta^{LO}$ . For small values of this parameter the method is not able to reach the optimum of the process since there is not enough excitation to estimate accurately the process derivatives. As this value increases, the algorithm is able to reach the optimum of the process due to the fact that the gradient is identified in a better way; however, the feasible region of the original problem is reduced, producing a path with significant changes in the direction and increasing the number of RTO iterations. This tradeoff is reflected in a better way if we consider process noise in the measurements, since the information of the curvature can be very inaccurate if the gradient is estimated with not enough energy. Because of this, we can also note that the proposed algorithm seems to be more robust due to fact that the modifiers have been updated with a method that is specially recommended for these cases (Walters et al., 1991).

The Otto Williams Reactor is a small size test bed, nevertheless it gives an indication about the scalability of the proposed methodology for a large-scale problem. It is clear that an increase in the decision variables grows the number of the vertex of the simplex in the NM algorithm and therefore, in a linear way, the evaluations in the objective function that are needed before the algorithm can start correcting the modifiers. In the same way as it happens in the computation of the gradients in the dual methodology.

## 5. CONCLUSIONS

The modifier-adaptation methodology has been reformulated as a nested optimization, allowing the use of gradient free algorithms to avoid the estimation of the experimental gradients, expanding the application of this methodology for cases where the experimental gradients are expensive to obtain.

Neglecting the gradient estimation also facilitate the implementation of methods based in modifiers, since there is no need to take into account the grade of excitation of the process. This situation is critical when there is noise in the process measurements where the methodology proposed seems to be more robust.

## ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement n°257462 HYCON2 Network of excellence. The financial support of the project “Control y Optimización de Redes de Distribución, MICINN DPI 2009-12805”, is also greatly appreciated

## REFERENCES

- Brdys, M. A. & Tatjewski, P. 2005. *Iterative algorithms for multilayer optimizing control*, London, Imperial College Press.
- Chachuat, B., Srinivasan, B. & Bonvin, D. (2009). Adaptation strategies for real-time optimization. *Computers & Chemical Engineering*, 33, 1557-1567.
- 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.
- Gao, W. & Engell, S. (2005). Iterative set-point optimization of batch chromatography. *Computers & Chemical Engineering*, 29, 1401-1409.
- 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., Chachuat, B. & Bonvin, D. (2009). Modifier-Adaptation Methodology for Real-Time Optimization. *Industrial & Engineering Chemistry Research*, 48, 6022-6033.
- Roberts, P. D. (1979). An Algorithm for Steady-State System Optimization and Parameter-Estimation. *International Journal of Systems Science*, 10, 719-734.
- Tatjewski, P. (2002). Iterative Optimizing Set-Point Control - The Basic Principle Redesigned. *15th Triennial World Congress, Barcelona, Spain*. IFAC.
- 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.
- Williams, T. J. & Otto, R. E. (1960). A generalized chemical processing model for the investigation of computer control. *AIEE Trans*, 79, 458-473.