Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

An Application of Modifier Adaptation with Quadratic Approximation on a Pilot Scale Plant in Industrial Environment ⋆

A.R. Gottu Mukkula [*] **S. Kern** [**] **M. Salge** [***]
M. Holtkamp [***] **S. Guhl** [**] **C. Fleicher** [***] **K. Meyer** [**]
M.P. Remelhe [****] **M. Maiwald** [**] **S. Engell** [*]

[*] *Process Dynamics and Operations Group, Technische Universität Dortmund, Dortmund, Germany*
(email: anweshreddy.gottumukkula@tu-dortmund.de)
[**] *Bundesanstalt für Materialforschung und -prüfung, Germany*
[***] *INVITE GmbH, Leverkusen, Germany*
[****] *Bayer AG, Leverkusen, Germany*

Abstract:
The goal of this work is to identify the optimal operating input for a lithiation reaction that is performed in a highly innovative pilot scale continuous flow chemical plant in an industrial environment, taking into account the process and safety constraints. The main challenge is to identify the optimum operation in the absence of information about the reaction mechanism and the reaction kinetics. We employ an iterative real-time optimization scheme called modifier adaptation with quadratic approximation (MAWQA) to identify the plant optimum in the presence of plant-model mismatch and measurement noise. A novel NMR PAT-sensor is used to measure the concentration of the reactants and of the product at the reactor outlet. The experiment results demonstrate the capabilities of the iterative optimization using the MAWQA algorithm in driving a complex real plant to an economically optimal operating point in the presence of plant-model mismatch and of process and measurement uncertainties.

Keywords: Iterative real-time optimization, Modifier adaptation, Plant-model mismatch, Reactor control, PAT-sensor, NMR.

1. INTRODUCTION

Active pharmaceutical ingredients (APIs) have been produced in multi-purpose batch and semi-batch reactors since long (Mleczko and Zhao, 2014; Porta et al., 2016). Recently, flow technologies have gained popularity within the pharmaceutical industry for the production of APIs. In order to survive the increasing competition, it is necessary to operate these processes optimally with respect to product yield and production cost. This stimulates the development of completely automated processes with increased efficiency.

Identifying the process optimum is possible by solving a model based optimization problem using a mathematical model which captures the behavior of the underlying process accurately. However, it is expensive or sometimes not possible to build high accuracy models. It is therefore in the interest of the process industries to employ optimization schemes which can identify economically optimal operating conditions of the process which satisfy the process and safety constraints, using an imperfect model with structural and/or parametric mismatch.

⋆ The research leading to these results has received funding from the European Unions Horizon 2020 research and innovation program under grant agreement number 636942 "CONSENS–Integrated Control and Sensing".

Iterative real-time optimization techniques have gained popularity in recent years to identify the optimal operating condition of a process despite the presence of plant-model mismatch. Initially, a two-step approach to handle parametric plant-model mismatch was proposed in Jang et al. (1987). In the two-step approach, the model parameters are updated using the plant measurements and the model based optimization problem with updated parameters is solved and the procedure is iterated. In Roberts (1979), the integrated system optimization and parameter estimation (ISOPE) approach was proposed where in addition to updating the model parameters, the objective function of the model based optimization problem is modified by adding bias and gradient correction terms which are also iteratively updated. The ISOPE scheme can handle both structural and parametric plant-model mismatch and the inputs obtained from the ISOPE converge iteratively to the KKT point of the plant. To avoid the estimation of the model parameters in each iteration, Tatjewski (2002) proposed the redesigned-ISOPE scheme. It was shown in Tatjewski (2002) that the uncertain parameters do not have to be updated in each iteration (or not at all) to achieve convergence to the KKT point of the plant. The redesigned-ISOPE scheme was further extended in Gao and Engell (2005) to the handling of process-dependent constraints and called iterative gradient-modification op-

timization (IGMO). They also refined the computation of the plant gradients which are needed in the iterative optimization. This approach was further analyzed in Marchetti et al. (2009) and the term modifier adaptation (MA) was proposed.

The ISOPE, IGMO and MA schemes require the knowledge of the gradients of the cost function and of the constraints with respect to the inputs of the real plant which are usually hard to obtain (Roberts, 2000). Finite differences are used to compute the process gradients in the ISOPE and the redesigned-ISOPE schemes. In presence of measurement noise, the gradient approximation using finite differences is prone to error. Additionally, it has a disadvantage of requiring additional plant perturbations around each input (Roberts, 2000). In Gao and Engell (2005), Broyden's formula from Roberts (2000) is used to approximate the plant gradients from the past moves, avoiding additional plant perturbations unless the data matrix becomes nearly singular, then optimized new test points are evaluated. In Gao et al. (2016), modifier adaptation with quadratic approximation (MAWQA) has been proposed as a combination of IGMO, quadratic approximation (QA) and elements from derivative free optimization (DFO) (Conn et al., 2009). It was demonstrated in Gao et al. (2016) that the QA handles the measurement noise well and is capable to decrease the gradient approximation error when compared to the above mentioned methods. Recently, Matias and Jäschke (2019) proposed to use radial-basis function network to compute the plant gradients in modifier adaptation.

Although there exists a large amount of literature on iterative real-time optimization methods, they are still not very widely used in the industry. According to Darby et al. (2011), the most successful applications of iterative real-time optimization are found in ethylene plants. In Hernandez et al. (2018), an experimental implementation of MAWQA at a miniplant for hydroformylation has been reported. Recently, de Avila Ferreira et al. (2019) used iterative real-time optimization to optimize a solid-oxide fuel-cell system. In this work, the optimum operating conditions of a lithiation reaction that takes place in an intensified continuous reactor is identified using the iterative real-time optimization scheme MAWQA and a novel NMR PAT sensor in an industrial environment, the INVITE test facility. The challenge here is to identify the process optimum in spite of the lack of knowledge about the reaction mechanism and kinetics.

The remainder of the paper is organized as follows: In Section 2, the iterative real-time optimization scheme MAWQA is presented. The lithiation reaction scheme, the process description and the models, both the nominal and the plant model used in the iterative optimization are presented in Section 3. Simulation results and experiment results are also presented and discussed in Section 3. Finally, in Section 4, the paper is summarized.

2. MODIFIER ADAPTATION WITH QUADRATIC APPROXIMATION

Consider a process for which a steady state mathematical model (1b) has been built. The mapping function $\boldsymbol{F}_m : \mathbb{R}^{n_u} \to \mathbb{R}^{n_y}$ maps the input variables represented by the n_u-dimensional vector $\boldsymbol{u}$ to the n_y-dimensional vector of measured variables $\hat{\boldsymbol{y}}$. Let $J_m : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$ be a function of input and measured variables that we want to minimize and $\boldsymbol{G}_m : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_c}$ be a vector of n_c inequality constraints. We assume that the functions J_m and G_m are at least twice differentiable with respect to $\boldsymbol{u}$. Optimal inputs of the process between bounds $\boldsymbol{u}^L$ and $\boldsymbol{u}^U$ can be obtained by solving the optimization problem (1). The identified optimum $\boldsymbol{u}_m^*$ using the a priori known process model (nominal model) however may differ considerably from the optimum $\boldsymbol{u}_p^*$ of the real process as $\boldsymbol{F}_m$ may not describe the real process (plant) accurately.

$$\boldsymbol{u}_m^* = \arg \min_{\boldsymbol{u} \in [\boldsymbol{u}^L, \boldsymbol{u}^U]} J_m(\hat{\boldsymbol{y}}, \boldsymbol{u}) \tag{1a}$$

$$\text{s.t. } \hat{\boldsymbol{y}} = \boldsymbol{F}_m(\boldsymbol{u}), \tag{1b}$$

$$\boldsymbol{G}_m(\hat{\boldsymbol{y}}, \boldsymbol{u}) \leq \boldsymbol{0}. \tag{1c}$$

MAWQA is an iterative gradient-modification optimization scheme that makes use of elements of derivative free optimization (DFO), in particular local quadratic approximations (Gao et al., 2016). In MAWQA the plant-model mismatch is handled by modifying and iteratively updating the objective and the constraint functions of the model-based optimization problem in (1) as in Gao and Engell (2005); Marchetti et al. (2009). The objective function is modified by adding a gradient correction term and the constraint function is modified by adding bias and gradient correction terms according to the following equations:

$$\begin{aligned} J_m^{ad,k} := {} & J_m(\hat{\boldsymbol{y}}, \boldsymbol{u}) + \\ & (\nabla J_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k) - \nabla J_m(\hat{\boldsymbol{y}}^k, \boldsymbol{u}^k))^T (\boldsymbol{u} - \boldsymbol{u}^k), \end{aligned} \tag{2a}$$

$$\begin{aligned} \boldsymbol{G}_m^{ad,k} := {} & \boldsymbol{G}_m(\hat{\boldsymbol{y}}, \boldsymbol{u}) + (\boldsymbol{G}_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k) - \boldsymbol{G}_m(\hat{\boldsymbol{y}}^k, \boldsymbol{u}^k)) + \\ & (\nabla \boldsymbol{G}_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k) - \nabla \boldsymbol{G}_m(\hat{\boldsymbol{y}}^k, \boldsymbol{u}^k))^T (\boldsymbol{u} - \boldsymbol{u}^k), \end{aligned} \tag{2b}$$

and the following optimization problem is solved iteratively:

$$\hat{\boldsymbol{u}}^{k+1} = \arg \min_{\boldsymbol{u} \in [\boldsymbol{u}^L, \boldsymbol{u}^U]} J_m^{ad,k} \tag{3a}$$

$$\text{s.t. } \hat{\boldsymbol{y}} = \boldsymbol{F}_m(\boldsymbol{u}), \tag{3b}$$

$$\boldsymbol{G}_m^{ad,k} \leq \boldsymbol{0}. \tag{3c}$$

The optimization problem (3) when solved in the k^{th} iteration computes the input for the $k+1^{\text{th}}$ iteration ($\boldsymbol{u}^{k+1}$). $\nabla J_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k)$, $\nabla J_m(\hat{\boldsymbol{y}}^k, \boldsymbol{u}^k)$ in (2) represent the gradient of the plant objective function $J_p(\boldsymbol{y}_p, \boldsymbol{u})$ and the nominal model objective function $J_m(\hat{\boldsymbol{y}}, \boldsymbol{u})$ with respect to the process inputs obtained at the k^{th} iteration. Similarly $\nabla G_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k)$, $\nabla G_m(\hat{\boldsymbol{y}}^k, \boldsymbol{u}^k)$ represent the gradient of the plant constraint function $G_p(\boldsymbol{y}_p, \boldsymbol{u})$ and the constraint functions of the nominal model $G_m(\hat{\boldsymbol{y}}, \boldsymbol{u})$ with respect to the process inputs at the k^{th} iteration.

In MAWQA, $\nabla J_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k)$ and $\nabla G_p(\boldsymbol{y}_p^k, \boldsymbol{u}^k)$ are approximated using a surrogate quadratic approximation model (Q). To fit Q, a minimum of $\frac{(n_u+1)(n_u+2)}{2}$ number of data points including the input point $\boldsymbol{u}^k$ at which the gradient is approximated are required. From the set of all evaluated plant inputs up to the k^{th} iteration ($\mathbb{U}^k$), a subset of input data points $\mathcal{U}^k$ are selected. The values of J_p and $\boldsymbol{G}_p$ for all $\mathcal{U}^k$ are used to approximate the plant objective function J_p and each of the constraint functions $\boldsymbol{G}_p$ using quadratic functions. The gradients are obtained by analytically dif-

ferentiating the fitted quadratic functions and evaluating them at $\boldsymbol{u}^k$. A general quadratic function can be expressed as $Q(\boldsymbol{p},\boldsymbol{u}) = \sum_{i=1}^{n_u}\sum_{j=1}^{i} a_{i,j}u_iu_j + \sum_{i=1}^{n_u} b_iu_i + c$, where $\boldsymbol{p} := \{a_{1,1},\ldots,a_{n_u,n_u},b_1,\ldots,b_{n_u},c\}$ is the vector of parameters of the quadratic function. The set $\mathcal{U}^k$ is identified by screening all the available data points in $\mathbb{U}^k$. In general, it is attempted to select $\mathcal{U}^k$ such that it consists of well distributed distant data points $\mathcal{U}^k_{dist}$ which act as anchor points and neighboring points $\mathcal{U}^k_{nb}$ which lie in the vicinity of $\boldsymbol{u}^k$. The inverse of the condition number of $\boldsymbol{s}^k$ ($\kappa^{-1}(\boldsymbol{s}^k)$) in (4) is used to assess the quality of the distribution of the data points (Gao et al., 2016). In (4), $[\boldsymbol{u}^k]$ and $[\mathcal{U}^k_{dist}]$ are matrix representations of $\boldsymbol{u}^k$ and of the set $\mathcal{U}^k_{dist}$. If $\kappa^{-1}(\boldsymbol{s}^k)$ is less than a desired value (δ), additional plant perturbations are performed and added to $\mathcal{U}^k$ to improve the distribution of the points. This condition is referred to as the conditionality condition.

$$\boldsymbol{s}^k = [\boldsymbol{u}^k]^{n_u\times 1} \otimes 1^{\text{cardinality}(\mathcal{U}^k_{dist})\times 1} - [\mathcal{U}^k_{dist}] \tag{4}$$

As the fitted quadratic functions are only local approximations of J_p and G_p, they are valid only in the vicinity of $\boldsymbol{u}^k$. In MAWQA, the process input $\hat{\boldsymbol{u}}^{k+1}$ obtained from the optimization problem (3) is restricted to lie inside a confidence ellipsoid (trust-region) by adding the following constraint to the modified optimization problem (3):

$$(\boldsymbol{u}-\boldsymbol{u}^k)^T \text{cov}(\mathcal{U}^k)(\boldsymbol{u}-\boldsymbol{u}^k) \leq \gamma^2, \tag{5}$$

where γ is a tuning parameter which scales the size of the trust region defined by the confidence ellipsoid (Gao et al., 2016).

In addition to QA, there are also other elements of derivative free optimization (DFO) Conn et al. (2009) included in MAWQA. It includes the criticality-check, the quality-check and a switch to an optimization based on the quadratic approximation model. The criticality-check is used to ensure that the anchor points $\mathcal{U}^k_{dist}$ that are used in fitting Q do not lie far from $\boldsymbol{u}^k$. The algorithm for the criticality-check is shown in Algorithm 1. In the quality-check, the quality of $J^{ad,k}_m$ and $\boldsymbol{G}^{ad,k}_m$ are compared with the fitted quadratic models J^k_Q and $\boldsymbol{G}^k_Q$. The minimum of ρ^k_m, ρ^k_Q in (6) determines the best among the adapted and quadratic models (Gao et al., 2016).

$$\rho^k_m := \max\left\{\left|1-\frac{J^{ad,k}_m - J^{ad,k-1}_m}{J^k_p - J^{k-1}_p}\right|, \left|1-\frac{G^{ad,k}_{m,1} - G^{ad,k-1}_{m,1}}{G^k_{p,1} - G^{k-1}_{p,1}}\right|, \ldots, \left|1-\frac{G^{ad,k}_{m,n_c} - G^{ad,k-1}_{m,n_c}}{G^k_{p,n_c} - G^{k-1}_{p,n_c}}\right|\right\}, \tag{6a}$$

$$\rho^k_Q := \max\left\{\left|1-\frac{J^{k}_Q - J^{k-1}_Q}{J^k_p - J^{k-1}_p}\right|, \left|1-\frac{G^{k}_{Q,1} - G^{k-1}_{Q,1}}{G^k_{p,1} - G^{k-1}_{p,1}}\right|, \ldots, \left|1-\frac{G^{k}_{Q,n_c} - G^{k-1}_{Q,n_c}}{G^k_{p,n_c} - G^{k-1}_{p,n_c}}\right|\right\}, \tag{6b}$$

From the quality-check, if the quality of the quadratic model is better than the adapted model, i.e. $\rho^k_Q < \rho^k_m$, instead of solving the optimization problem in (3) using the modified objective and constraint functions, the following optimization problem based on the fitted quadratic approximation model is solved:

$$\boldsymbol{u}^{k+1} = \arg\min_{\boldsymbol{u}\in[\boldsymbol{u}^L,\boldsymbol{u}^U]} Q(\boldsymbol{p}^{k,J},\boldsymbol{u}) \tag{7a}$$

$$\text{s.t. } \boldsymbol{Q}(\boldsymbol{p}^{k,\boldsymbol{G}},\boldsymbol{u}) \leq \boldsymbol{0}, \tag{7b}$$

$$(\boldsymbol{u}-\boldsymbol{u}^k)^T \text{cov}(\mathcal{U}^k)(\boldsymbol{u}-\boldsymbol{u}^k) \leq \gamma^2. \tag{7c}$$

Algorithm 1 Criticality-check

Require: $\triangle\boldsymbol{u}, \boldsymbol{u}^{k+1}, \boldsymbol{u}^k, \mathcal{U}^k_{dist}$
  if $||\boldsymbol{u}^{k+1}-\boldsymbol{u}^k||_2 < \triangle\boldsymbol{u}$ **then**
    Solve
    $$\boldsymbol{u}^* = \arg\max_{\boldsymbol{u}_i\in\mathcal{U}^k_{dist}} ||\boldsymbol{u}_i - \boldsymbol{u}^k||_2$$
    while $||\boldsymbol{u}^*-\boldsymbol{u}^k||_2 > 2\triangle\boldsymbol{u}$ **do**
      $\boldsymbol{u}^* := \frac{\boldsymbol{u}^*+\boldsymbol{u}^k}{2}$
      $\boldsymbol{u}^{k+1} := \boldsymbol{u}^*$
    end while
  end if

3. LITHIATION PROCESS

3.1 Process Description

A lithiation reaction is performed in a containerized reactor module that was developed within the F^3-factory project (Bieringer et al., 2013) at the INVITE facility in Leverkusen, Germany. The important modules used in the lithiation process, i.e. the coiled tubular reactor, an online NMR sensor, a NIR flow cell and a product filter are shown in Figure 1. The reactants Aniline (A), 1-Fluoro-2-nitrobenzene (E) and Lithium bis(trimethylsilyl)amide (B) are fed into the reactor through a mixer to produce Lithium 2-Nitrodiphenylamine (H). The reaction scheme is:

$$\text{A} + \text{B} \longrightarrow \text{C} + \text{D}, \qquad \text{C} + \text{E} \longrightarrow \text{F} + \text{G}, \tag{8a}$$

$$\text{C} + \text{F} \longrightarrow \text{H} + \text{A}, \qquad \text{F} + \text{B} \longrightarrow \text{H} + \text{D}, \tag{8b}$$

$$\text{H} + \text{E} \longrightarrow \text{I} + \text{G}. \tag{8c}$$

where C, D, F and I represent Lithium phenylazanide, Hexamethyldisilazane, 2-Nitrodiphenylamine and 4-Nitro-N-(4-nitrophenyl)-N-phenylaniline. Tetrahydrofuran is used as a solvent in this reaction. The reaction is performed in a coiled tube which does not have mixing elements and is cooled by the ambient air. As the reactor does not have a cooling jacket, for safety reasons the throughput of the reactants is limited by the requirement that the temperature at the reactor outlet is less than 54 °C.

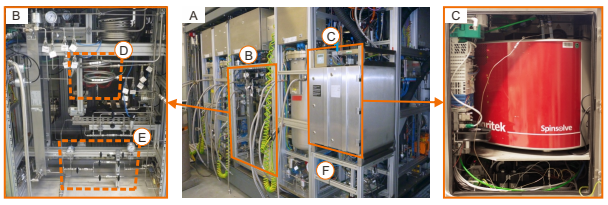

Fig. 1. Photograph of the modular pilot plant (A) for the continuous production of Li-NDPA. (B) is a close-up of the tubular reactor (D, Inner diameter = 12.4 mm) and the filter section (E). (C) is a photograph of the integrated compact NMR spectrometer (43 MHz) with ATEX certified pressurized housing for online concentration measurements. (F) indicates the location of the NIR flow cell.

During the reaction, the byproduct Lithium fluoride (G) precipitates along the length of the reactor, forming a fouling layer. Fouling in tubular reactors causes local constrictions of the tube, thus leading to varying flow velocities and increasing pressure drops. In the worst case, a complete blockage of the tubular reactor happens (Schoenitz et al., 2015). The reaction mixture from the outlet of the reactor is passed through a filter to remove the solid particles of G that are produced during the reaction. The filtered mixture is then passed to a storage tank. A sample of the filtered reaction mixture is continuously passed to the online NMR and to an NIR flow cell to measure the concentrations of the reactants, of the byproducts and of the product in the reaction mixture at the reactor outlet. We refer to Kern et al. (2019) for further information about the online NMR device used in this study.

3.2 Plant model

The following assumptions are made in modelling the reaction in the coiled tubular reactor: (i) the effect of coiling is negligible, (ii) the reactor is an ideal plug flow reactor, (iii) the specific heat capacity and the densities are time invariant, (iv) relative reaction rate constants were taken as provided by INVITE based on their observations and are time invariant, (v) the 5th elementary reaction (H + E $\longrightarrow$ I + G) does not take place and (vi) the concentration and temperature gradients in the radial direction are negligible.

Based on the above assumptions, the following partial differential equations (PDE) model using general mass and energy balance equations for tubular reactors results:

$$\frac{\partial C_i}{\partial t} = -\frac{\partial(v_z C_i)}{\partial z} + D_{cz}\frac{\partial^2 C_i}{\partial z^2} + r_i \quad \forall i \in \{A,B,C,D,E,F,G,H\}, \tag{9a}$$

$$\frac{\partial T_R}{\partial t} = -\frac{\partial(v_z T_R)}{\partial z} + \frac{1}{\rho c_p}(D_{tz}\frac{\partial^2 T_R}{\partial z^2} + KA(T_R - T_E) - \triangle H_1 k_1 C_A C_B - \triangle H_2 k_2 C_C C_E - \triangle H_3 k_3 C_C C_F - \triangle H_4 k_4 C_B C_F), \tag{9b}$$

where C_i, r_i denote the concentration and the reaction rate of the i^{th} reactant, v_z represents the velocity of the reaction mixture in the axial direction, D_{cz}, D_{tz} are the concentration and thermal dispersion coefficients. The algebraic equations for the reaction rates are:

$$r_A = -k_1 C_A C_B + k_3 C_C C_F, \quad r_E = -k_2 C_C C_E, \tag{10a}$$
$$r_B = -k_1 C_A C_B + k_4 C_B C_F, \quad r_G = k_2 C_C C_E, \tag{10b}$$
$$r_C = k_1 C_A C_B - k_2 C_C C_E - r_A, \tag{10c}$$
$$r_D = k_1 C_A C_B + k_4 C_F C_B, \tag{10d}$$
$$r_F = k_2 C_C C_E - k_3 C_C C_F - k_4 C_B C_F, \tag{10e}$$
$$r_H = k_3 C_C C_F + k_4 C_F C_B, \tag{10f}$$

where the reaction rates $k_i \forall i = 1,\ldots,4$ are given by

$$k_i = k_{i0}\exp(\frac{-E}{RT_R}). \tag{11}$$

The steady state model for the lithiation reaction is obtained by substituting $\frac{\partial C_i}{\partial t} = 0$ and $\frac{\partial T_R}{\partial t} = 0$ in the PDE model (9). This substitution leads to a second order differential equation in the longitudinal coordinate z. As precise information is lacking, the dispersion coefficients of the concentrations and of the temperature are set to zero. The resulting simplified model is

$$\frac{\partial C_i}{\partial z} = \frac{r_i}{v_z} \ \forall i \in \{A,B,C,D,E,F,G,H\}, \tag{12a}$$

$$\frac{\partial T_R}{\partial z} = \frac{1}{v_z \rho c_p}\big(KA(T_R - T_E) - \triangle H_1 k_1 C_A C_B - \triangle H_2 k_2 C_C C_E - \triangle H_3 k_3 C_C C_F - \triangle H_4 k_4 C_B C_F\big), \tag{12b}$$

and the model parameter values are listed in Table 1. The steady state optimum of the plant model results as $\boldsymbol{u}_p^* := [4.21, 8, 6.65]$ kg/h for the flow rates of reactants A, B and E.

Table 1. Parameter values used in the plant model

Parameter	Value	Parameter	Value
$\triangle H_1,\ldots,\triangle H_4$	-6 J/mol	T_{in}	283.15 K
$k_{10},\ldots,k_{40}$	17.8 s^{-1}	E	25×10^3 J/mol
KA	0.001 W/K	$\triangle H$	-19 J/mol
T_E	293.15 K	c_p	0.123 J/(mol K)
ρ	900 kg/m^3	R	8.314 J/(mol K)

3.3 Nominal model

The nominal steady-state model approximates the five step reaction mechanism in (8) by assuming a single reaction

$$\text{A} + 2\text{B} + \text{E} \longrightarrow 2\text{D} + \text{G} + \text{H}. \tag{13}$$

The algebraic equations for the reaction rates are

$$\overline{r}_A = -kC_A C_B C_E, \quad \overline{r}_B = -2kC_A C_B C_E, \tag{14a}$$
$$\overline{r}_D = 2kC_A C_B C_E, \quad \overline{r}_E = -kC_A C_B C_E, \tag{14b}$$
$$\overline{r}_G = kC_A C_B C_E, \quad \overline{r}_H = kC_A C_B C_E, \tag{14c}$$

where $k = 17.8\exp(\frac{E}{RT_R})$. The steady-state reactor model is

$$\frac{\partial C_i}{\partial z} = \frac{\overline{r}_i}{v_z} \ \forall i \in \{A,B,D,E,G,H\}, \tag{15a}$$

$$\frac{\partial T_R}{\partial z} = \frac{1}{v_z \rho c_p}\big(KA(T_R - T_E) - \triangle H k C_A C_B C_E\big). \tag{15b}$$

This nominal model has a different structure compared to the real process leading to structural plant-model mismatch. The steady state optimum of the nominal model is $\boldsymbol{u}_m^* := [4.32, 8, 6.7]$ kg/h for the flows of the reactants A, B and E.

3.4 Communication

In order to run the MAWQA scheme at the real plant autonomously, it was necessary to establish a reliable communication between MAWQA and the Siemens PCS7 control system of the plant and between MAWQA and the SQL database of the online NMR. A schematic of the communication scheme used for the experimental implementation of the MAWQA scheme is shown in Figure 2. The PCS7 control system has an OPC-DA server installed. The temperature of the reaction mixture at reactor outlet is recorded in the OPC-DA server from which this information is read by the MAWQA scheme. The SQL database of the online NMR records the concentration of the reactants,

Table 2. Description of the markers used in Figures 3,4. Input markers in Red and Black refer to A (Aniline); Blue and Magenta refer to B (Li-HMDS); Green and Cyan refer to E (FNB).

Marker	Description	Marker	Description
[■, ■, ■]	Successful-iteration input from model-based optimization	[□, □, □]	Perturbation input
[★, ★, ★]	Successful-iteration input from optimization using quadratic model	[✳, ✳, ✳]	Optimal input for the plant
[☆, ☆, ☆]	Explorative-iteration input from optimization using quadratic model	– –○– –	Evolution of the profit function
[▶, ▶, ▶]	Input due to criticality, which improved the value of profit function		
[▷, ▷, ▷]	Input due to criticality, which did not improve the value of profit function		

by-products and the product. This information is read by the MAWQA scheme from the SQL database. The input calculated by the iterative optimization scheme is written onto the SQL database of the online NMR which writes it onto the OPC-DA server of the PCS7 control system. The Siemens PCS7 control system finally applies this input to the plant.

3.5 Problem formulation

Our goal is to identify the optimal inlet flows for the lithiation process during continuous operation. Additionally the optimal inputs have to satisfy the process and safety constraints on the feed flow rates of the reactants and on the temperature of the reaction mixture at the reactor outlet. The optimum is defined by a profit function which depends on the cost of the feeds and the price of the valuable product at the reactor outlet. The inputs from the iterative RTO algorithm are passed to the local mass flow controllers via the communication system described above which establish the set-points in the dosing units. The optimization problem for identifying the process optimum is formulated as

$$\max_{\boldsymbol{u}} \quad \frac{\bar{w}_H C_H M_H}{\rho} \sum_{i=1}^{3} u_i - \bar{w}_A u_1 - \bar{w}_B u_2 - \bar{w}_E u_3 \quad (16a)$$

$$\text{s.t. Plant model (12)}, \quad (16b)$$

$$T_R - 327.15 \leq 0, \quad (16c)$$

$$\boldsymbol{u}^L \leq \boldsymbol{u} \leq \boldsymbol{u}^U, \quad (16d)$$

where M_H is the molecular weight of component H (the target product), $\boldsymbol{u} := \{u_1, u_2, u_3\}$ represents the feed flow rates of components A, B, E and $(\bar{w}_A, \bar{w}_B, \bar{w}_E, \bar{w}_H)$ are the weights used in the profit function. The weights reflect the cost of the reactants and of the product. The weight percentage of A, B and E in the feed is 10.61, 20.0 and 10.18. Additionally, the limits for the feed flow rates of A, B and E are [3-8] kg/h.

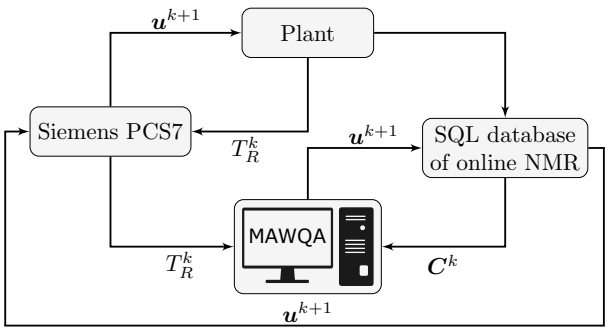

Fig. 2. Illustration of the setup used to communicate with the plant via the Siemens PCS7 system and via the SQL database of the online NMR.

3.6 Simulation results

A simulation result with tuning parameters $\gamma = 1$, $\triangle \boldsymbol{u} = 0.1$ and $\delta = 0.1$ of the MAWQA algorithm is shown in Figure 3. The description of the markers used in the figure is provided in Table 2. In order to take into account the process noise, the input flow rates and the concentration measurement in the plant profit function are corrupted by Gaussian noise with a standard deviation of 0.05. Before the first iteration in Figure 3, after performing plant perturbations according to finite differences, 9 additional plant perturbations are performed to benefit early from the quadratic approximation. These additional perturbation inputs are chosen such that they are well distributed over the operating region. The MAWQA scheme already identifies a near optimum operating input $\boldsymbol{u}^3$ in the second iteration. Although the optimum of the plant and the nominal model are close, the optimization problem (7) using the quadratic approximation model is chosen by the MAWQA over the modifier adaptation problem (2) using the nominal model, as the approximated QA model represents the plant measurements better than the nominal model. Therefore, the optimization problem in (7) is solved in every iteration starting from the 1st iteration. After identifying a near optimum input $\boldsymbol{u}^3$, the MAWQA scheme performed several exploratory moves and moves to satisfy the criticality conditions before it converged to the known plant optimum.

3.7 Experimental result

For the experiment, the screening algorithm for choosing $\mathcal{U}^k$ data points from $\mathbb{U}^k$ was suppressed. This is done to use all the available data points for fitting a quadratic function, to avoid oscillating input set-points from the MAWQA scheme which can be caused by flipping of the quadratic function. Recently in Gottu Mukkula and Engell (2020), a guaranteed model adequacy scheme for MAWQA was proposed to avoid flipping of the quadratic function in MAWQA, thereby avoiding oscillating input set-points from the MAWQA scheme.

The input moves computed by the MAWQA scheme and the evolution of the profit function are shown in Figure 4. The online concentration measurements from the online NMR, the offline concentration measurements from the NIR sensor and the feed flow rates are shown in Figure 5. The MAWQA scheme was initialized with a flow rate of 3.58 kg/h for A, B and E. In the beginning, plant perturbations were realized to approximate the plant gradients using finite differences. Additionally, 9 additional plant evaluations were performed to ensure the availability of well distributed data points for quadratic approximation. From the 1st iteration onwards, quadratic approximation

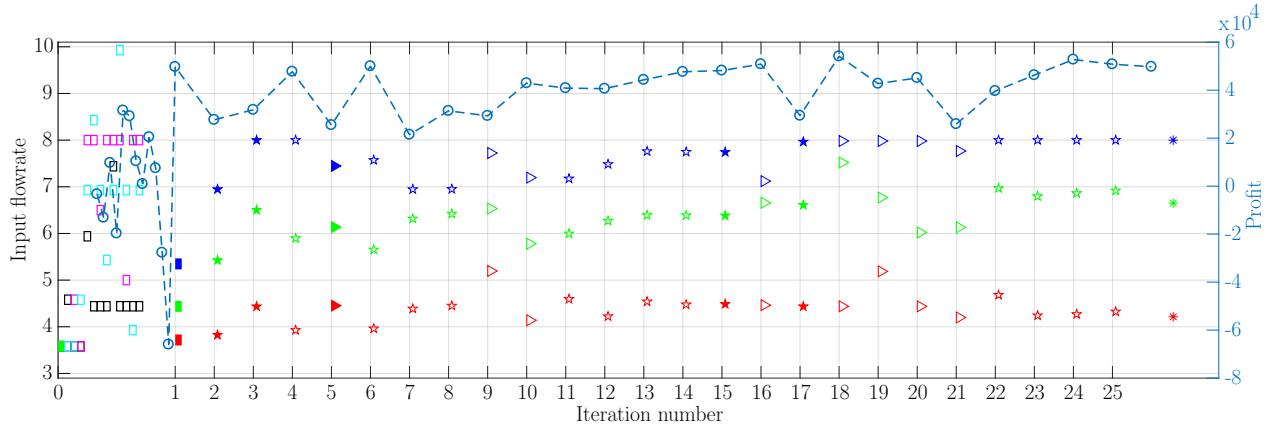

Fig. 3. Simulation result: Evolution of the input flow rates of reactants A, B and E computed using MAWQA and the resulting profit function using simulated measurements corrupted by Gaussian noise. The description of the markers used is provided in Table 2.

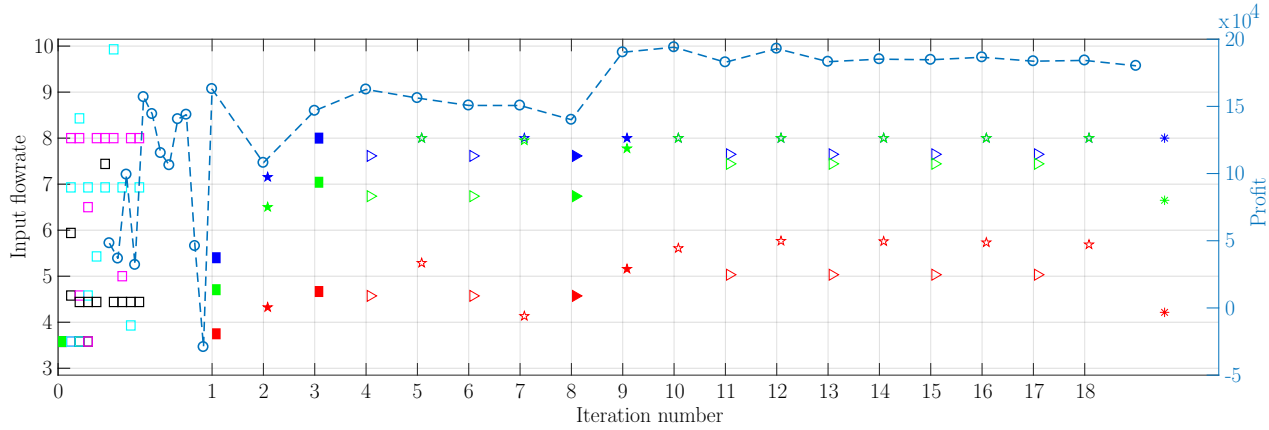

Fig. 4. Experimental result: Evolution of the input flow rates of reactants A, B and E optimized by MAWQA and the obtained profit function using measurements from the online NMR sensor. The description of the markers used can be found in Table 2.

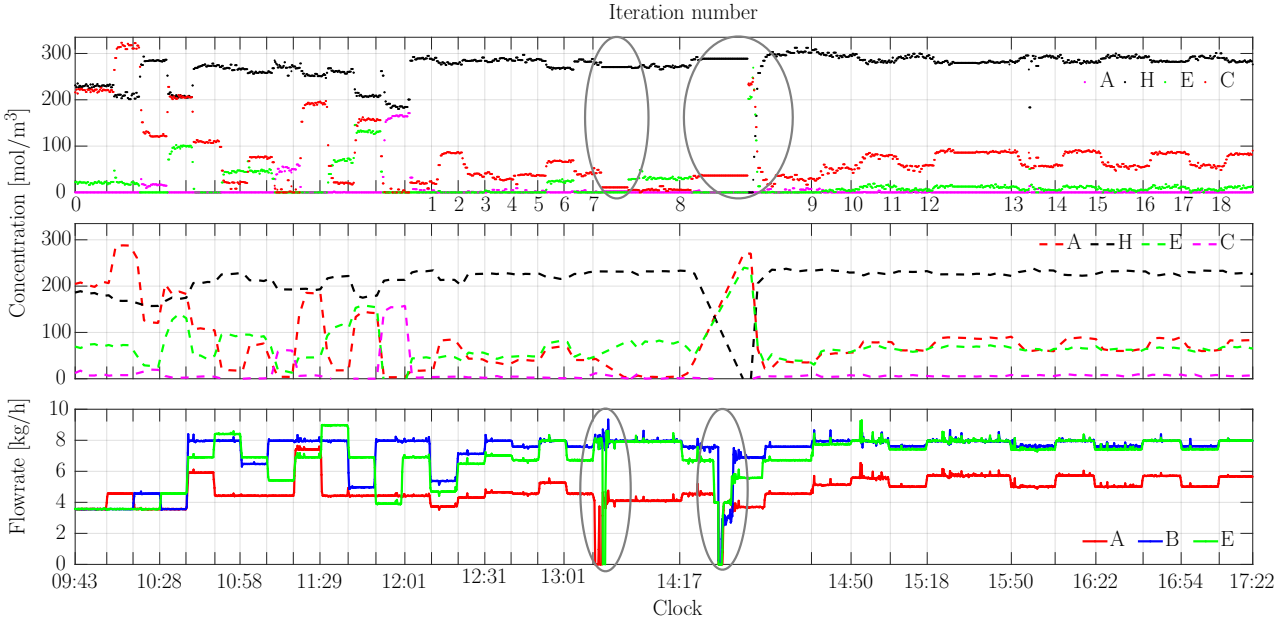

Fig. 5. Top figure: Concentration measurements from the online NMR; Middle figure: Concentration measurements from the NIR; Bottom figure: Input flow rates of reactants A, B and E as computed by MAWQA. Technical glitches that occurred during the experiment are highlighted by ellipses.

was used for the approximation of the plant gradient due to the availability of the minimum required number of data points. The 3rd iteration input is already close to the theoretical plant optimum. During the 7th MAWQA iteration, an unplanned event took place, the feed container was changed and the weight percentage of Li-HMDS in the feed was increased to 20.6. This information was not made available to the control system. Nonetheless, as the controlled process was not yet stationary, it was reflected in the plant gradients that were computed from the quadratic approximation and the control scheme responded by making moves to improve the plant performance. The improvement in the plant profit over the MAWQA iterations is shown in Figure 4. The input moves between the 3rd to the 7th and from the 9th to the 18th iteration are caused due to the tuning of the algorithm, specifically by not employing the screening algorithm and the criticality-check. Due to switching off the screening, there always are data points that lie farther away from the current point by more than $2\triangle \boldsymbol{u}$ which triggered additional moves via the criticality-check. This has the positive effect that the algorithm reacts quickly to changes in the plant but leads to slightly oscillating inputs which however have almost no influence on the profit. Nonetheless, the online optimization algorithm showed a very good performance, driving the plant to the optimal operation in the presence of significant structural and parametric plant-model mismatch. As usually in industrial plants, such probing moves of the controlled inputs are unwanted, extensions are needed that suppress further moves but restart the algorithm when plant model mismatch is detected. First proposals in this direction can be found in Ye et al. (2018); Gottu Mukkula et al. (2019).

4. CONCLUSION

In this paper, we reported the development of an online real-time optimization solution for a highly innovative containerized continuous production process in a pilot plant. MAWQA, a novel iterative optimization scheme that uses a plant model as well as measured data, was chosen as a suitable algorithm for the case study. The simulation results show that the MAWQA scheme is capable to drive the plant to its optimum despite significant plant-model mismatch in few iterations. It then was validated experimentally that the combination of online NMR measurements of the concentrations at the plant outlet with the iterative optimization algorithm MAWQA can drive the plant to an optimal operation despite significant deviations between the model used in the optimization algorithm and the behaviour of the real plant. It was even possible to exploit an unknown change in the operating conditions to improve the efficiency of the operation of the plant. In our future work, we will focus on developing a standardized approach for choosing the tuning parameters in MAWQA scheme.

REFERENCES

Bieringer, T., Buchholz, S., and Kockmann, N. (2013). Future production concepts in the chemical industry: Modular – small-scale – continuous. *Chemical Engineering & Technology*, 36(6), 900–910.

Conn, A.R., Scheinberg, K., and Vicente, L.N. (2009). *Introduction to derivative-free optimization*. SIAM.

Darby, M.L., Nikolaou, M., Jones, J., and Nicholson, D. (2011). Rto: An overview and assessment of current practice. *Journal of Process Control*, 21(6), 874–884.

de Avila Ferreira, T., Wuillemin, Z., et al. (2019). Real-time optimization of an experimental solid-oxide fuel-cell system. *Journal of Power Sources*, 429, 168–179.

Gao, W. and Engell, S. (2005). Iterative set-point optimization of batch chromatography. *Computers & Chemical Engineering*, 29(6), 1401–1409.

Gao, W., Wenzel, S., and Engell, S. (2016). A reliable modifier-adaptation strategy for real-time optimization. *Computers & Chemical Engineering*, 91, 318–328.

Gottu Mukkula, A.R., Ahmad, A., and Engell, S. (2019). Start-up and shut-down conditions for iterative real-time optimization methods. In *2019 6th Indian Control Conference (ICC)*, 158–163.

Gottu Mukkula, A.R. and Engell, S. (2020). Guaranteed model adequacy for modifier adaptation with quadratic approximation. In *2020 19th European Control Conference (ECC)*, 1037–1042.

Hernandez, R., Dreimann, J., Vorholt, A., et al. (2018). Iterative real-time optimization scheme for optimal operation of chemical processes under uncertainty: Proof of concept in a miniplant. *Industrial & Engineering Chemistry Research*, 57(26), 8750–8770.

Jang, S.S., Joseph, B., and Mukai, H. (1987). On-line optimization of constrained multivariable chemical processes. *AIChE Journal*, 33(1), 26–35.

Kern, S., Wander, L., Meyer, K., et al. (2019). Flexible automation with compact nmr spectroscopy for continuous production of pharmaceuticals. *Analytical and Bioanalytical Chemistry*, 411(14), 3037–3046.

Marchetti, A., Chachuat, B., and Bonvin, D. (2009). Modifier-adaptation methodology for real-time optimization. *Industrial & engineering chemistry research*, 48(13), 6022–6033.

Matias, J. and Jäschke, J. (2019). Using a neural network for estimating plant gradients in real-time optimization with modifier adaptation. *IFAC-PapersOnLine*, 52(1), 808–813.

Mleczko, L. and Zhao, D. (2014). Technology for continuous production of fine chemicals. 1181, 403–440.

Porta, R., Benaglia, M., and Puglisi, A. (2016). Flow chemistry: Recent developments in the synthesis of pharmaceutical products. *Organic Process Research & Development*, 20(1), 2–25.

Roberts, P. (1979). An algorithm for steady-state system optimization and parameter estimation. *International Journal of Systems Science*, 10(7), 719–734.

Roberts, P. (2000). Broyden derivative approximation in isope optimising and optimal control algorithms. *IFAC Proceedings Volumes*, 33(16), 293–298.

Schoenitz, M., Grundemann, L., Augustin, W., and Scholl, S. (2015). Fouling in microstructured devices: a review. *Chemical communications*, 51 39, 8213–28.

Tatjewski, P. (2002). Iterative optimizing set-point control – the basic principle redesigned. *IFAC Proceedings Volumes*, 35(1), 49–54.

Ye, L., Shen, F., Ge, Z., and Song, Z. (2018). On an aspect of implementing real-time optimization: Establishing the suspending and activating conditions incorporating process monitoring. *IFAC-PapersOnLine*, 51(18), 79–84.