

Article

# A Study of Explorative Moves during Modifier Adaptation with Quadratic Approximation

Weihua Gao \*, Reinaldo Hernández and Sebastian Engell

Biochemical and Chemical Engineering Department, TU Dortmund, Emil-Figge-Str. 70, 44221 Dortmund, Germany; reinaldo.hernandez@bci.tu-dortmund.de (R.H.); sebastian.engell@bci.tu-dortmund.de (S.E.)

\* Correspondence: weihua.gao@bci.tu-dortmund.de; Tel.: +49-231-755-5131

Academic Editor: Dominique Bonvin

Received: 31 October 2016; Accepted: 22 November 2016; Published: 26 November 2016

**Abstract:** Modifier adaptation with quadratic approximation (in short MAWQA) can adapt the operating condition of a process to its economic optimum by combining the use of a theoretical process model and of the collected data during process operation. The efficiency of the MAWQA algorithm can be attributed to a well-designed mechanism which ensures the improvement of the economic performance by taking necessary explorative moves. This paper gives a detailed study of the mechanism of performing explorative moves during modifier adaptation with quadratic approximation. The necessity of the explorative moves is theoretically analyzed. Simulation results for the optimization of a hydroformylation process are used to illustrate the efficiency of the MAWQA algorithm over the finite difference based modifier adaptation algorithm.

**Keywords:** real-time optimization; modifier adaptation; quadratic approximation

## 1. Introduction

In the process industries, performing model-based optimization to obtain economic operations usually implies the need of handling the problem of plant-model mismatch. An optimum that is calculated using a theoretical model seldom represents the plant optimum. As a result, real-time optimization (RTO) is attracting considerable industrial interest. RTO is a model based upper-level optimization system that is operated iteratively in closed loop and provides set-points to the lower-level regulatory control system in order to maintain the process operation as close as possible to the economic optimum. RTO schemes usually estimate the process states and some model parameters or disturbances from the measured data but employ a fixed process model which leads to problems if the model does not represent the plant accurately.

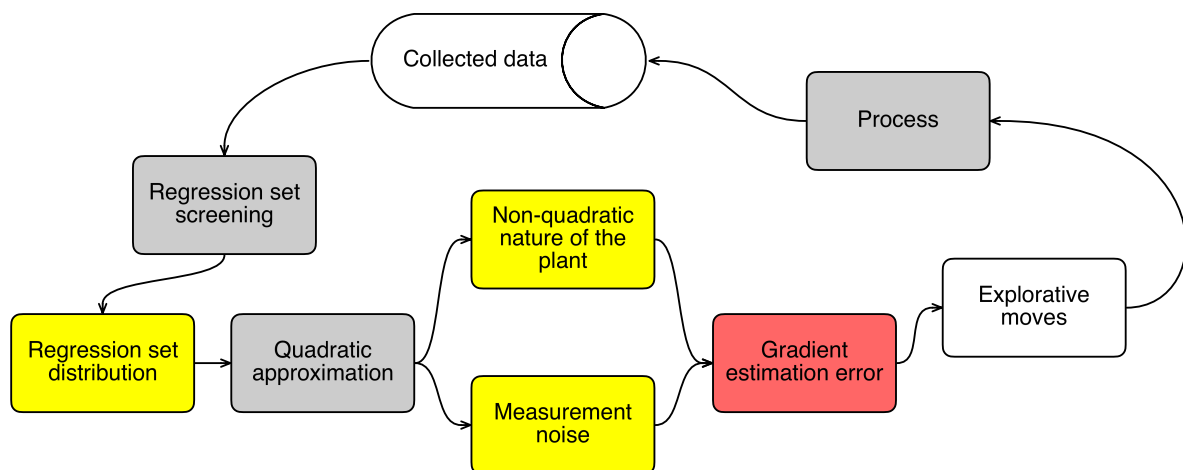
Several schemes have been proposed towards how to combine the use of theoretical models and of the collected data during process operation, in particular the model adaptation or two-step scheme [1]. It handles plant-model mismatch in a sequential manner via an identification step followed by an optimization step. Measurements are used to estimate the uncertain model parameters, and the updated model is used to compute the decision variables via model-based optimization. The model adaptation approach is expected to work well when the plant-model mismatch is only of parametric nature, and the operating conditions lead to sufficient excitation for the estimation of the plant outputs. In practice, however, both parametric and structural mismatch are typically present and, furthermore, the excitation provided by the previously visited operating points is often not sufficient to accurately identify the model parameters.

For an RTO scheme to converge to the plant optimum, it is necessary that the gradients of the objective as well as the values and gradients of the constraints of the optimization problem match those of the plant. Schemes that directly adapt the model-based optimization problem by using the

update terms (called modifiers) which are computed from the collected data have been proposed [2–4]. The modifier-adaptation schemes can handle considerable plant-model mismatch by applying bias- and gradient-corrections to the objective and to the constraint functions. One of the major challenges in practice, as shown in [5], is the estimation of the plant gradients with respect to the decision variables from noisy measurement data.

Gao et al. [6] combined the idea of modifier adaptation with the quadratic approximation approach that is used in derivative-free optimization and proposed the modifier adaptation with quadratic approximation (in short MAWQA) algorithm. Quadratic approximations of the objective function and of the constraint functions are constructed based on the screened data which are collected during the process operation. The plant gradients are computed from the quadratic approximations and are used to adapt the objective and the constraint functions of the model-based optimization problem. Simulation studies for the optimization of a reactor benchmark problem with noisy data showed that by performing some explorative moves the true optimum can be reliably obtained. However, neither the generation of the explorative moves nor their necessity for the convergence of set-point to the optimum was theoretically studied. Due to the fact that the estimation of the gradients using the quadratic approximation approach requires more data than those that are required by using a finite difference approach, the efficiency of the MAWQA algorithm, in terms of the number of plant evaluations to obtain the optimum, has been questioned, in particular for the case of several decision variables. In addition, in practice, it is crucial for plant operators to be confident with the necessity of taking the explorative moves which may lead to a deterioration of plant performance.

This paper reports a detailed study of the explorative moves during modifier adaptation with quadratic approximation. It starts with how the explorative moves are generated and then the factors that influence the generation of these moves are presented. The causality between the factors and the explorative moves is depicted in Figure 1, where the blocks with a yellow background represent the factors. The use of a screening algorithm to optimize the regression set for quadratic approximations is shown to ensure that an explorative move is only performed when the past collected data cannot provide accurate gradient estimates. Simulation results for the optimization of a hydroformylation process with four optimization variables are used to illustrate the efficiency of the MAWQA algorithm, which takes necessary explorative moves, over the finite difference based modifier adaptation algorithm.



**Figure 1.** Causality between the explorative moves and the influencing factors.

## 2. Modifier Adaptation with Quadratic Approximation

Let  $J_m(\mathbf{u})$  and  $\mathbf{C}_m(\mathbf{u})$  represent the objective and the vector of constraint functions of a static model-based optimization problem, assumed to be twice differentiable with respect to the vector of decision variables  $\mathbf{u} \in \mathbb{R}^{n_u}$

$$\begin{aligned} \min_{\mathbf{u}} \quad & J_m(\mathbf{u}) \\ \text{s.t.} \quad & \mathbf{C}_m(\mathbf{u}) \leq \mathbf{0}. \end{aligned} \quad (1)$$

At each iteration of the modifier adaptation algorithm, bias- and gradient-corrections of the optimization problem are applied as

$$\begin{aligned} \min_{\mathbf{u}} \quad & J_m(\mathbf{u}) + \left( \nabla J_p^{(k)} - \nabla J_m^{(k)} \right)^T \left( \mathbf{u} - \mathbf{u}^{(k)} \right) \\ \text{s.t.} \quad & \mathbf{C}_m(\mathbf{u}) + \mathbf{C}_p^{(k)} - \mathbf{C}_m^{(k)} + \left( \nabla \mathbf{C}_p^{(k)} - \nabla \mathbf{C}_m^{(k)} \right)^T \left( \mathbf{u} - \mathbf{u}^{(k)} \right) \leq \mathbf{0}. \end{aligned} \quad (2)$$

The symbols are explained in Table 1.  $\nabla J_p^{(k)}$  and  $\nabla \mathbf{C}_p^{(k)}$  are usually approximated by the finite difference approach

$$\nabla J_p^{(k)} \approx \begin{bmatrix} u_1^{(k)} - u_1^{(k-1)} & \dots & u_{n_u}^{(k)} - u_{n_u}^{(k-1)} \\ \vdots & \vdots & \vdots \\ u_1^{(k)} - u_1^{(k-n_u)} & \dots & u_{n_u}^{(k)} - u_{n_u}^{(k-n_u)} \end{bmatrix}^{-1} \begin{bmatrix} J_p^{(k)} - J_p^{(k-1)} \\ \vdots \\ J_p^{(k)} - J_p^{(k-n_u)} \end{bmatrix}, \quad (3)$$

where  $n_u$  is the number of dimensions of  $\mathbf{u}$ ,  $J_p^{(k-i)}$ ,  $i = 0, \dots, n_u$ , are the plant objectives at set-points  $\mathbf{u}^{(k-i)}$ ,  $i = 0, \dots, n_u$ , and  $\nabla \mathbf{C}_p^{(k)}$  is approximated similarly. The accuracy of the finite difference approximations is influenced by both the step-sizes between the set-points and the presence of measurement noise. In order to acquire accurate gradient estimations, small step-sizes are preferred. However, the use of small step-sizes leads to a high sensitivity of the gradient estimates to measurement noise.

**Table 1.** Symbols used in the modifier adaptation formulation.

Symbol	Description
$k$	Index of iteration
$\mathbf{u}^{(k)}$	Current set-point
$\nabla J_p^{(k)}$	Gradient vector of the plant objective function at $\mathbf{u}^{(k)}$
$\nabla J_m^{(k)}$	Gradient vector of the model-predicted objective function at $\mathbf{u}^{(k)}$
$\mathbf{C}_p^{(k)}$	Vector of the plant constraint values at $\mathbf{u}^{(k)}$
$\mathbf{C}_m^{(k)}$	Vector of the model-predicted constraint values at $\mathbf{u}^{(k)}$
$\nabla \mathbf{C}_p^{(k)}$	Gradient matrix of the plant constraint functions at $\mathbf{u}^{(k)}$
$\nabla \mathbf{C}_m^{(k)}$	Gradient matrix of the model-predicted constraint functions at $\mathbf{u}^{(k)}$

In the MAWQA algorithm, the gradients are computed analytically from quadratic approximations of the objective function and of the constraint functions that are regressed based on a screened set (represented by  $\mathcal{U}^{(k)}$  at the  $k$ th iteration) of all the collected data (represented by  $\mathbb{U}$ ). The screened set consists of near and distant points:  $\mathcal{U}^{(k)} = \mathcal{U}_n \cup \mathcal{U}_d$ , where  $\mathcal{U}_n = \{\mathbf{u} : \|\mathbf{u} - \mathbf{u}^{(k)}\| < \Delta \mathbf{u}; \text{ and } \mathbf{u} \in \mathbb{U}\}$ , and  $\mathcal{U}_d$  is determined by

$$\begin{aligned}
 \min_{\mathcal{U}_d} & \frac{\sum_{\mathbf{u} \in \mathcal{U}_d} \|\mathbf{u} - \mathbf{u}^{(k)}\|}{\theta(\mathcal{U}_d)} \\
 \text{s.t.} & \text{size}(\mathcal{U}_d) \geq C_2^{n_u+2} - 1 \\
 & \mathcal{U}_d \subset \mathbb{U} \setminus \mathcal{U}_n,
 \end{aligned} \tag{4}$$

where  $\Delta \mathbf{u}$  is sufficiently large so that  $\mathcal{U}_d$  guarantees robust quadratic approximations with noisy data,  $\theta(\mathcal{U}_d)$  is the minimal angle between all possible vectors that are defined by  $\mathbf{u} - \mathbf{u}^{(k)}$ , and  $C_2^{n_u+2} = (n_u + 2)(n_u + 1)/2$  is the number of data required to uniquely determine the quadratic approximations.

In the MAWQA algorithm, the regression set  $\mathcal{U}^{(k)}$  is also used to define a constrained search space  $\mathcal{B}^{(k)}$  for the next set-point move

$$\mathcal{B}^{(k)} : (\mathbf{u} - \mathbf{u}^{(k)})^T M^{-1} (\mathbf{u} - \mathbf{u}^{(k)}) \leq \gamma^2, \tag{5}$$

where  $M = cov(\mathcal{U}^{(k)})$  is the covariance matrix of the selected points (inputs) and  $\gamma$  is a scaling parameter.  $\mathcal{B}^{(k)}$  is a  $n_u$ -axial ellipsoid centered at  $\mathbf{u}^{(k)}$ . The axes of the ellipsoid are thus aligned with the eigenvectors of the covariance matrix. The semi-axis lengths of the ellipsoid are related to the eigenvalues of the covariance matrix by the scaling parameter  $\gamma$ . The adapted optimization (2) is augmented by the search space constraint as

$$\begin{aligned}
 \min_{\mathbf{u}} & J_{ad}^{(k)}(\mathbf{u}) \\
 \text{s.t.} & \mathbf{C}_{ad}^{(k)}(\mathbf{u}) \leq \mathbf{0} \\
 & \mathbf{u} \in \mathcal{B}^{(k)},
 \end{aligned} \tag{6}$$

where  $J_{ad}^{(k)}(\mathbf{u})$  and  $\mathbf{C}_{ad}^{(k)}(\mathbf{u})$  represent the adapted objective and constraint functions in (2).

In the application of the modifier adaptation with quadratic approximation, it can happen that the nominal model is inadequate for the modifier-adaptation approach and that it is better to only use the quadratic approximations to compute the next plant move. In order to ensure the convergence, it is necessary to monitor the performance of the adapted optimization and possibly to switch between model-based and data-based optimizations. In each iteration of the MAWQA algorithm, a quality index of the adapted optimization  $\rho_m^{(k)}$  is calculated and compared with the quality index of the quadratic approximation  $\rho_\phi^{(k)}$ , where

$$\rho_m^{(k)} = \max \left\{ \left| 1 - \frac{J_{ad}^{(k)} - J_{ad}^{(k-1)}}{J_p^{(k)} - J_p^{(k-1)}} \right|, \left| 1 - \frac{C_{ad,1}^{(k)} - C_{ad,1}^{(k-1)}}{C_{p,1}^{(k)} - C_{p,1}^{(k-1)}} \right|, \dots, \left| 1 - \frac{C_{ad,n_c}^{(k)} - C_{ad,n_c}^{(k-1)}}{C_{p,n_c}^{(k)} - C_{p,n_c}^{(k-1)}} \right| \right\} \tag{7}$$

and

$$\rho_\phi^{(k)} = \max \left\{ \left| 1 - \frac{J_\phi^{(k)} - J_\phi^{(k-1)}}{J_p^{(k)} - J_p^{(k-1)}} \right|, \left| 1 - \frac{C_{\phi,1}^{(k)} - C_{\phi,1}^{(k-1)}}{C_{p,1}^{(k)} - C_{p,1}^{(k-1)}} \right|, \dots, \left| 1 - \frac{C_{\phi,n_c}^{(k)} - C_{\phi,n_c}^{(k-1)}}{C_{p,n_c}^{(k)} - C_{p,n_c}^{(k-1)}} \right| \right\} \tag{8}$$

with  $J_\phi^{(k)}$  and  $\mathbf{C}_\phi^{(k)}$  are the quadratic approximations of the objective and the constraint functions. If  $\rho_m^{(k)} \leq \rho_\phi^{(k)}$ , the predictions of the adapted model-based optimization are more accurate than that of the quadratic approximations and (6) is performed to determine the next set-point. Otherwise, an optimization based on the quadratic approximations is done

$$\begin{aligned}
& \min_{\mathbf{u}} J_{\phi}^{(k)}(\mathbf{u}) \\
& \text{s.t.} \quad \mathbf{C}_{\phi}^{(k)}(\mathbf{u}) \leq 0 \\
& \quad \mathbf{u} \in \mathcal{B}^{(k)}.
\end{aligned} \tag{9}$$

The MAWQA algorithm is given as follows:

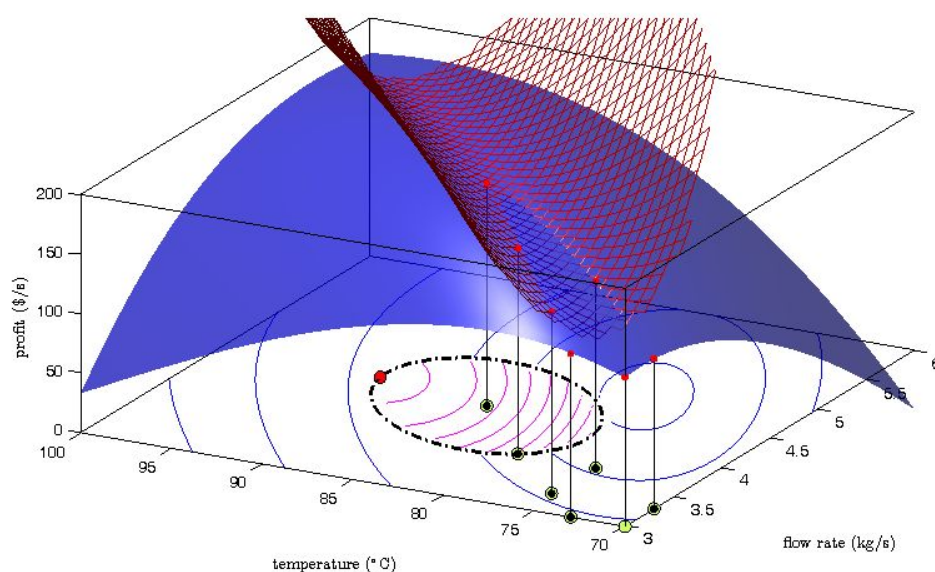
- Step 1. Choose an initial set-point  $\mathbf{u}^{(0)}$  and probe the plant at  $\mathbf{u}^{(0)}$  and  $\mathbf{u}^{(0)} + h\mathbf{e}_i$ , where  $h$  is a suitable step size and  $\mathbf{e}_i \in \mathbb{R}^{n_u}$  ( $i = 1, \dots, n_u$ ) are mutually orthogonal unit vectors. Use the finite difference approach to calculate the gradients at  $\mathbf{u}^{(0)}$  and run the IGMO approach [3] until  $k \geq C_2^{n_u+2}$  set-points have been generated. Run the screening algorithm to define the regression set  $\mathcal{U}^{(k)}$ . Initialize  $\rho_m^{(k)} = 0$  and  $\rho_{\phi}^{(k)} = 0$ .
- Step 2. Calculate the quadratic functions  $J_{\phi}^{(k)}$  and  $\mathbf{C}_{\phi}^{(k)}$  based on  $\mathcal{U}^{(k)}$ . Determine the search space  $\mathcal{B}^{(k)}$  by (5).
- Step 3. Compute the gradients from the quadratic functions. Adapt the model-based optimization problem and determine the optimal set-point  $\hat{\mathbf{u}}^{(k)}$  as follows:
  - (a) If  $\rho_m^{(k)} \leq \rho_{\phi}^{(k)}$ , run the adapted model-based optimization (6).
  - (b) Else perform the data-based optimization (9).
- Step 4. If  $\|\hat{\mathbf{u}}^{(k)} - \mathbf{u}^{(k)}\| < \Delta\mathbf{u}$  and there exists one point  $\mathbf{u}^{(j)} \in \mathcal{U}^{(k)}$  such that  $\|\mathbf{u}^{(j)} - \mathbf{u}^{(k)}\| > 2\Delta\mathbf{u}$ , set  $\hat{\mathbf{u}}^{(k)} = (\mathbf{u}^{(j)} + \mathbf{u}^{(k)}) / 2$ .
- Step 5. Evaluate the plant at  $\hat{\mathbf{u}}^{(k)}$  to acquire  $J_p(\hat{\mathbf{u}}^{(k)})$  and  $\mathbf{C}_p(\hat{\mathbf{u}}^{(k)})$ . Prepare the next step as follows
  - (a) If  $\hat{J}_p^{(k)} < J_p^{(k)}$ , where  $\hat{J}_p^{(k)} = J_p(\hat{\mathbf{u}}^{(k)})$ , this is a performance-improvement move. Define  $\mathbf{u}^{(k+1)} = \hat{\mathbf{u}}^{(k)}$  and run the screening algorithm to define the next regression set  $\mathcal{U}^{(k+1)}$ . Update the quality indices  $\rho_m^{(k+1)}$  and  $\rho_{\phi}^{(k+1)}$ . Increase  $k$  by one and go to Step 2.
  - (b) If  $\hat{J}_p^{(k)} \geq J_p^{(k)}$ , this is an explorative move. Run the screening algorithm to update the regression set for  $\mathbf{u}^{(k)}$ . Go to Step 2.

Note that the index of iteration of the MAWQA algorithm is increased by one only when a performance-improvement move is performed. Several explorative moves may be required at each iteration. The number of plant evaluations is the sum of the numbers of both kinds of moves. The next section studies why the explorative moves are required and how they contribute to the improvement of the performance on a longer horizon.

### 3. Analysis of the Explorative Moves

In the MAWQA algorithm, the quadratic approximations of the objective and the constraint functions are started once  $C_2^{n_u+2}$  data have been collected. It can happen that the distribution of the set-points is not “well-poised” [7] to ensure that the gradients are accurately estimated via the quadratic approximations, especially when the initial set-point is far away from the optimum and the following set-point moves are all along some search direction. Interpolation-based derivative-free optimization algorithms rely on a model-improvement step that generates additional set-point moves to ensure the well-poisedness of the interpolation set. Although the MAWQA algorithm was designed without an explicit model-improvement step, the generation of explorative moves can be considered as an implicit step to improve the poisedness of the regression set for the quadratic approximations. This section gives a theoretical analysis of the explorative moves. We start with some observations from the simulation results in [6] and relate the explorative moves to the estimation error of the gradients. The factors that influence the accuracy of the estimated gradients are analyzed. It is shown that the screening of the regression set leads to very pertinent explorative moves which, on the one hand, are sufficient to improve the accuracy of the gradient estimations, and, on the other hand, are less expensive than the model-improvement step in the derivative-free optimization algorithms.

The generation of the explorative moves is presented in Figure 2 where one MAWQA iteration for the optimization of the steady-state profit of the Williams-Otto reactor with respect to the flow rate and the reaction temperature [6] is illustrated. Here the blue surface represents the real profit mapping, and the mesh represents the quadratic approximation which was computed based on the regression set (●: set-point moves, ●: measured profit values). The bottom part shows the contours of the profit as predicted by the uncorrected model (blue lines), the constrained search space (dash-dot line), and the contours of the modifier-adapted profit (inside, magenta lines). Comparing the surface plot and the mesh plot, we can see that the gradient along the direction of the last set-point move is estimated well. However, a large error can be observed in the perpendicular direction. The gradient error propagates to the modifier-adapted contours and therefore, the next set-point move (●) points to the direction where the gradient is badly estimated. Despite the fact that the move may not improve the objective function, the data collection in that direction can later help to improve the gradient estimation.



**Figure 2.** Illustration of one MAWQA iteration with noisy data. Surface plot: real profit mapping, mesh plot: quadratic approximation, ●: regression set-point, ●: not chosen set-point, ●: measured profit, ●: next set-point move, blue contours: model-predicted profit, magenta contours: modifier-adapted profit, dash-dot line: constrained search space.

The example illustrates how the gradient error in a specific direction may lead to an explorative move along the same direction. In the MAWQA algorithm, the gradients are determined by evaluating  $\nabla J_\phi$  and  $\nabla C_\phi$  at  $\mathbf{u}^{(k)}$ . In order to be able to quantify the gradient error, we assume that the screened set  $\mathcal{U}^{(k)}$  is of size  $C_2^{n_u+2}$  and the quadratic approximations are interpolated based on  $\mathcal{U}^{(k)}$ . In the application of the MAWQA algorithm, this assumption is valid when the current set-point is far away from the plant optimum. Recall the screening algorithm,  $\mathcal{U}^{(k)}$  consists of the near set  $\mathcal{U}_n$  and the distant set  $\mathcal{U}_d$ . From (4), we can conclude that the distant set  $\mathcal{U}_d$  is always of size  $C_2^{n_u+2} - 1$ . Step 4 of the MAWQA algorithm ensures that the near set  $\mathcal{U}_n$  only consists of  $\mathbf{u}^{(k)}$  until the optimized next move is such that  $\|\hat{\mathbf{u}}^{(k)} - \mathbf{u}^{(k)}\| < \Delta \mathbf{u}$  and there are no points  $\mathbf{u}^{(j)} \in \mathcal{U}^{(k)}$  such that  $\|\mathbf{u}^{(j)} - \mathbf{u}^{(k)}\| > 2\Delta \mathbf{u}$ , that is, all the points in  $\mathcal{U}_d$  keep suitable distances away from  $\mathbf{u}^{(k)}$  for good local approximations. The above two conditions imply that  $\|\mathbf{u}^{(k)} - \mathbf{u}^*\| \leq \Delta \mathbf{u}$ , where  $\mathbf{u}^*$  represents the plant optimum. As a result of Step 4, when  $\|\mathbf{u}^{(k)} - \mathbf{u}^*\| > \Delta \mathbf{u}$ ,  $\mathcal{U}_n$  is always of size 1. For simplicity, a shift of coordinates to move  $\mathbf{u}^{(k)}$  to the origin is performed and the points in  $\mathcal{U}^{(k)}$  are reordered as  $\{\mathbf{0}, \mathbf{u}^{d_1}, \dots, \mathbf{u}^{d_n}\}$ , where  $n = C_2^{n_u+2} - 1$ .

Let  $\phi = \{1, u_1, \dots, u_{n_u}, u_1^2, \dots, u_{n_u}^2, \sqrt{2} u_1 u_2, \sqrt{2} u_1 u_3, \dots, \sqrt{2} u_{n_u-1} u_{n_u}\}$  represent a natural basis of the quadratic approximation. Let  $\underline{\alpha}$  represent a column vector of the coefficients of the quadratic approximation. The quadratic approximation of the objective function is formulated as

$$J_\phi(\mathbf{u}) = \alpha_0 + \sum_{i=1}^{n_u} \alpha_i u_i + \sum_{i=1}^{n_u} \alpha_{n_u+i} u_i^2 + \sqrt{2} \sum_{i=1}^{n_u-1} \sum_{j=i+1}^{n_u} \alpha_{2n_u+I(n_u,i,j)} u_i u_j, \tag{10}$$

where  $I(n_u, i, j) = n_u(i - 1) - (i + 1)i/2 + j$ . The coefficients  $\alpha_i, i = 0, \dots, n$  are calculated via the interpolation of the  $n + 1$  data sets  $\{\mathbf{u}, J_p(\mathbf{u})\}$

$$\underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & \varphi_1(\mathbf{u}^{(d_1)}) & \dots & \varphi_n(\mathbf{u}^{(d_1)}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \varphi_1(\mathbf{u}^{(d_n)}) & \dots & \varphi_n(\mathbf{u}^{(d_n)}) \end{bmatrix}}_{\underline{M}(\phi, \mathcal{U}^{(k)})} \underbrace{\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}}_{\underline{\alpha}} = \underbrace{\begin{bmatrix} J_p(\mathbf{0}) \\ J_p(\mathbf{u}^{(d_1)}) \\ \vdots \\ J_p(\mathbf{u}^{(d_n)}) \end{bmatrix}}_{\underline{J}_p(\mathcal{U}^{(k)})} + \underbrace{\begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_n \end{bmatrix}}_{\underline{v}}, \tag{11}$$

where  $\varphi_i(\mathbf{u}), i = 1, \dots, n$ , represent the polynomial bases in  $\phi$ ,  $J_p$  and  $v$  represent the noise-free objective and the measurement noise. Assume  $\underline{M}(\phi, \mathcal{U}^{(k)})$  is nonsingular,

$$\underline{\alpha} = \left(\underline{M}(\phi, \mathcal{U}^{(k)})\right)^{-1} \underline{J}_p(\mathcal{U}^{(k)}) + \left(\underline{M}(\phi, \mathcal{U}^{(k)})\right)^{-1} \underline{v}. \tag{12}$$

The computed gradient vector at the origin via the quadratic approximation is

$$\nabla J_\phi(\mathbf{0}) = (\alpha_1, \dots, \alpha_{n_u})^T = \underline{\alpha}^{J_p} + \underline{\alpha}^v, \tag{13}$$

where  $\underline{\alpha}^{J_p}$  represents the noise-free estimation and  $\underline{\alpha}^v$  represents the influence of the measurement noise. From [8], a bound on the error between the noise-free estimation  $\underline{\alpha}^{J_p}$  and  $\nabla J_p(\mathbf{0})$  can be obtained and simplified to

$$\|\underline{\alpha}^{J_p} - \nabla J_p(\mathbf{0})\| \leq \frac{1}{6} G \Lambda \sum_{i=1}^n \|\mathbf{u}^{(d_i)}\|^3, \tag{14}$$

where  $G$  is an upper bound on the third derivative of  $J_p(\mathbf{u})$ , and  $\Lambda$  is a constant that depends on the distribution of the regression set  $\mathcal{U}^{(k)}$ . Note that the bound in (14) is defined for the error between the plant gradients and the estimated gradients. It is different from the lower and upper bounds on the gradient estimates which were studied by Bunin et al. [5]. To simplify the study of  $\Lambda$ , assume  $\|\mathbf{u}^{(d_i)} - \mathbf{u}^{(k)}\| = \Delta \mathbf{u}$ .  $\Lambda$  is defined as

$$\Lambda \geq \max_{0 \leq i \leq n} \max_{\|\mathbf{u}\| \leq \Delta \mathbf{u}} |\ell_i(\mathbf{u})|, \tag{15}$$

where  $\ell_i(\mathbf{u}), i = 0, \dots, n$ , are the Lagrange polynomial functions that are defined by the matrix determinants

$$\ell_i(\mathbf{u}) = \frac{\det(\underline{M}(\phi, \mathcal{U}_i(\mathbf{u})))}{\det(\underline{M}(\phi, \mathcal{U}^{(k)}))} \tag{16}$$

with the set  $\mathcal{U}_i(\mathbf{u}) = \mathcal{U}^{(k)} \setminus \{\mathbf{u}^{(d_i)}\} \cup \{\mathbf{u}\}$ . The determinant of  $\underline{M}(\phi, \mathcal{U}^{(k)})$  is computed as

$$\det(\underline{M}(\phi, \mathcal{U}^{(k)})) = \begin{vmatrix} \varphi_1(\mathbf{u}^{(d_1)}) & \dots & \varphi_n(\mathbf{u}^{(d_1)}) \\ \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{u}^{(d_n)}) & \dots & \varphi_n(\mathbf{u}^{(d_n)}) \end{vmatrix}. \tag{17}$$

Let  $vol(\phi(\mathcal{U}^{(k)}))$  represent the volume of the  $n$ -dimensional convex hull spanned by the row vectors of the matrix in (17), we have

$$|det(\underline{M}(\phi, \mathcal{U}^{(k)}))| = vol(\phi(\mathcal{U}^{(k)})) n!. \tag{18}$$

Except the vertex at the origin, all the other vertices of the convex hull distribute on a  $n$ -dimensional sphere with radius  $\Delta \mathbf{u} + \Delta \mathbf{u}^2$ .  $vol(\phi(\mathcal{U}^{(k)}))$  reaches its maximal value when the vectors are orthogonal to each other. Let  $\mathbf{v}_i$  and  $\mathbf{v}_j$  represent any two row vectors of the matrix in (17). The angle between them

$$\cos(\theta_{i,j}^v) = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\| \|\mathbf{v}_j\|} = \frac{u_1^{(d_i)} u_1^{(d_j)} + \dots + u_{n_u}^{(d_i)} u_{n_u}^{(d_j)} + \left( u_1^{(d_i)} u_1^{(d_j)} + \dots + u_{n_u}^{(d_i)} u_{n_u}^{(d_j)} \right)^2}{\Delta \mathbf{u} + \Delta \mathbf{u}^2}. \tag{19}$$

Note that the angle  $\theta_{i,j}^v$  is different from the angle  $\theta_{i,j}$  between vectors  $\mathbf{u}^{d_i}$  and  $\mathbf{u}^{d_j}$

$$\cos(\theta_{i,j}) = \frac{u_1^{(d_i)} u_1^{(d_j)} + \dots + u_{n_u}^{(d_i)} u_{n_u}^{(d_j)}}{\Delta \mathbf{u}}. \tag{20}$$

The relationship between  $\theta_{i,j}^v$  and  $\theta_{i,j}$  is illustrated in Figure 3, where the angle between two 3-dimensional unit vectors is changed from 0 to 180 degree and the angle between the corresponding quadratic interpolation vectors increases proportionally when  $\theta \leq 90$  degree and stays in an interval of [90 98] degree from 90 to 180 degree. Recall the screening for the distant set  $\mathcal{U}_d$  via (4), the consideration of the minimal angle at the denominator of the objective function ensures the choosing of the best set in terms of the orthogonality of the matrix in (17) from all the collected data. As a result of (14)–(16) and (18), the lowest bound of the error between  $\underline{\alpha}^{J_p}$  and  $\nabla J_p(\mathbf{0})$  based on the collected data is achieved.

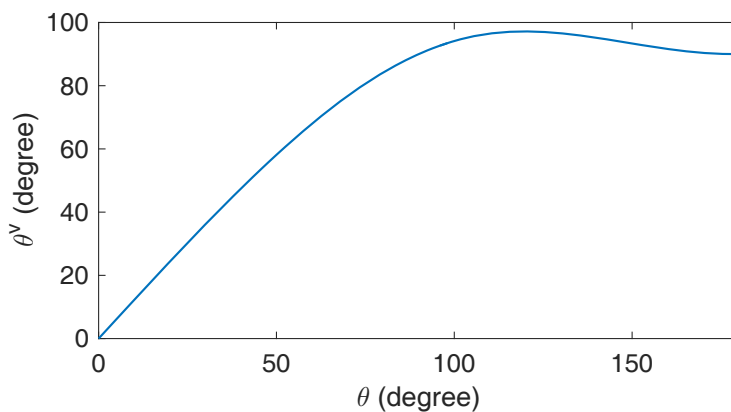


Figure 3. Illustration of the relationship between  $\theta$  and  $\theta^v$ .

The error due to the measurement noise  $\underline{\alpha}^v$  can be calculated by Cramer’s rule

$$\alpha_i^v = \frac{det(\underline{M}_i(\phi, \mathcal{U}^{(k)}, \underline{\mathbf{v}}))}{det(\underline{M}(\phi, \mathcal{U}^{(k)}))} = \frac{\delta_{noise}}{\Delta \mathbf{u}} \frac{det(\underline{M}_i(\phi, \mathcal{U}^{(k)}, \tilde{\mathbf{v}}))}{det(\underline{M}(\phi, \mathcal{U}^{(k)}))}, \tag{21}$$

where  $\underline{M}_i(\phi, \mathcal{U}^{(k)}, \underline{\mathbf{v}})$  is the matrix formed by replacing the  $(i + 1)$ th column of  $\underline{M}(\phi, \mathcal{U}^{(k)})$  by the column vector  $\underline{\mathbf{v}}$ ,  $\delta_{noise}$  is the level of the measurement noise,  $\mathbf{v}_i \in (-\delta_{noise}, +\delta_{noise})$ , and  $\tilde{\mathbf{v}} = \Delta \mathbf{u} \underline{\mathbf{v}} / \delta_{noise}$  represents the scaled vector of noises. In order to reduce  $\underline{\alpha}^v$ , the value of  $\Delta \mathbf{u}$  should be large enough. However, from (14) the error bound will increase accordingly. The optimal tuning of  $\Delta \mathbf{u}$  according to



$\delta_{noise}$  and  $G$  can be a future research direction. For a given distribution of  $\mathcal{U}_d$ , the relative upper bound of  $\alpha_i^y$  is related to the variance of the elements at the  $(i + 1)$ th column. To show that, we start from the angle between the  $(i + 1)$ th column vector and the first column vector

$$\cos(\beta_{i+1,1}) = \frac{\sum_{j=1}^n u_i^{(d_j)}}{\sqrt{n+1} \sqrt{\sum_{j=1}^n \left(u_i^{(d_j)}\right)^2}}, \quad (22)$$

where we use  $\beta$  to differentiate the angle from that formed by the row vectors. The variance of the elements at the  $(i + 1)$ th column is

$$\text{Var}(\{0, u_i^{(d_1)}, \dots, u_i^{(d_n)}\}) = \frac{\sum_{j=1}^n \left(u_i^{(d_j)}\right)^2}{n+1} - \left(\frac{\sum_{j=1}^n u_i^{(d_j)}}{n+1}\right)^2. \quad (23)$$

From (22) and (23), we obtain that

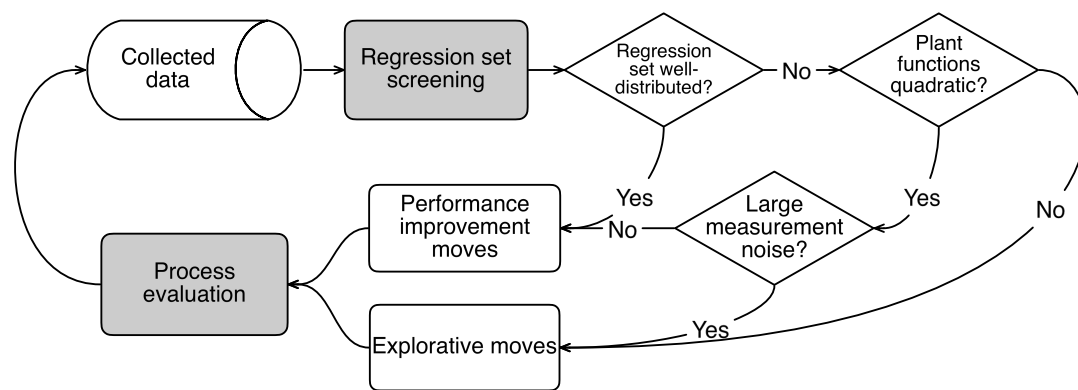
$$\cos(\beta_{i+1,1}) = \frac{\sum_{j=1}^n u_i^{(d_j)}}{(n+1) \sqrt{\text{Var} + \left(\frac{\sum_{j=1}^n u_i^{(d_j)}}{n+1}\right)^2}}. \quad (24)$$

For the same mean value, the orthogonality of the  $(i + 1)$ th column vector to the first column vector can be quantified by the variance of the elements at the  $(i + 1)$ th column. As discussed before, the absolute value of the determinant of  $\underline{M}(\phi, \mathcal{U}^{(k)})$  is influenced by the orthogonality. As a result, to replace a column of elements with small variance leads to a higher upper bound of the error than to replace a column of elements with large variance. Note that this is consistent with the constrained search space which is defined by the covariance matrix of the regression set.

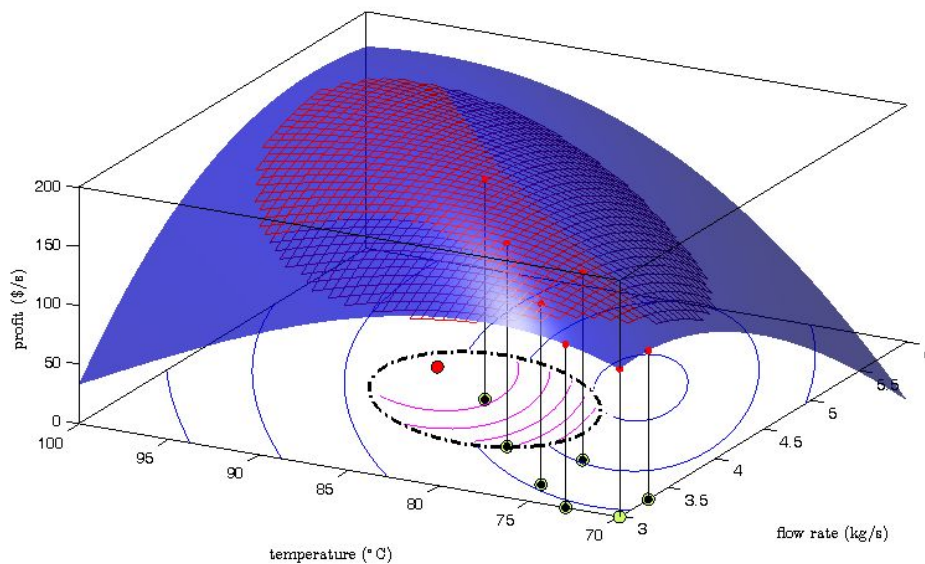
From (14) and (21), three factors determine the gradient estimation errors

- Distribution of the regression set (quantified by  $\Lambda$  and the distance to the current point)
- Non-quadratic nature of the plant (quantified by  $G$ , the upper bound on the third derivative)
- Measurement noise (quantified by  $\delta_{noise}$ ).

Figure 4 depicts how the three factors influence the generation of the explorative moves. Assume the plant functions are approximately quadratic in the region  $\|\mathbf{u} - \mathbf{u}^{(k)}\| \leq \Delta \mathbf{u}$  around the current set-point and the value of  $\Delta \mathbf{u}$  is large enough, a well-distributed regression set normally leads to a performance-improvement move. If the regression set is not well-distributed but the plant functions are approximately quadratic, a performance-improvement move is still possible if the level of the measurement noise  $\delta_{noise}$  is low. This is illustrated in Figure 5, where noise-free data was considered and the MAWQA algorithm did not perform any explorative moves. The explorative moves are only generated when either the plant functions deviate significantly from quadratic functions or the level of the measurement noise  $\delta_{noise}$  is large. In the case shown in Figure 2, the mapping of the plant profit is approximately quadratic but considerable measurement noise was presented. As a result, the unfavourable distribution of the regression set leads to an explorative move.

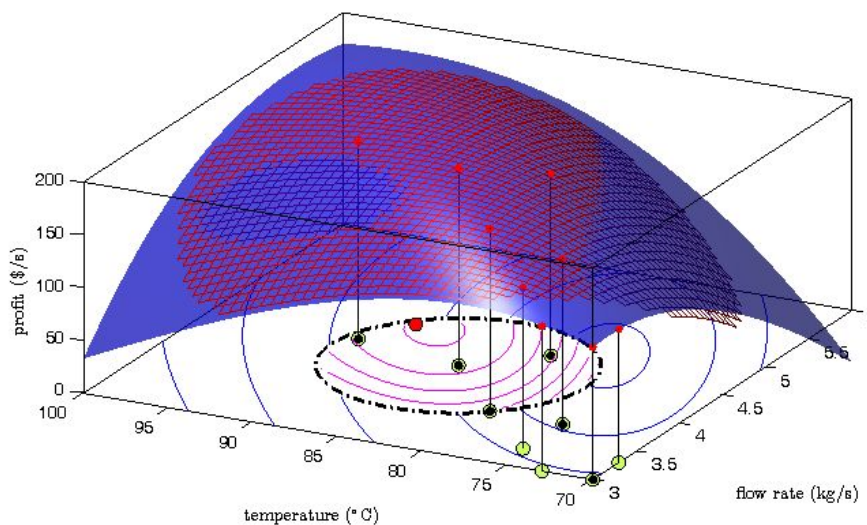


**Figure 4.** Illustration of the generation of explorative moves.



**Figure 5.** Illustration of one MAWQA iteration with noise-free measurements. Surface plot: real profit mapping, mesh plot: quadratic approximation, ●: regression set-point, ○: not chosen set-point, ●: measured profit, ●: next set-point move, blue contours: model-predicted profit, magenta contours: modifier-adapted profit, dash-dot line: constrained search space.

Gao et al. [6] proved that the explorative moves can ensure an improvement of the accuracy of the gradient estimation in the following iterations. As an example, Figure 6 shows the MAWQA iteration after two explorative moves for the optimization of the Williams-Otto reactor with noisy data [6]. The mesh of the quadratic approximation well represents the real mapping (the blue surface) around the current point which locates at the center of the constrained search space (defined by the dash-dot line). The computed gradients based on the quadratic approximation are more accurate than those based on the quadratic approximation in Figure 2. As a result, a performance-improvement move (represented by ●) was obtained.



**Figure 6.** Illustration of the MAWQA iteration after two explorative moves. Surface plot: real profit mapping, mesh plot: quadratic approximation,  $\odot$ : regression set-point,  $\circ$ : not chosen set-point,  $\bullet$ : measured profit,  $\bullet$ : next set-point move, blue contours: model-predicted profit, magenta contours: modifier-adapted profit, dash-dot line: constrained search space.

#### 4. Simulation Studies

The optimization of a hydroformylation process with four optimization variables is used to illustrate the efficiency of the MAWQA algorithm over the finite difference based modifier adaptation algorithm. The continuous hydroformylation of 1-dodecene in a thermomorphic multicomponent solvent (in short TMS) system is considered. This process was developed in the context of the collaborative research centre InPROMPT at the universities of Berlin, Dortmund and Magdeburg and was demonstrated on miniplant scale at TU Dortmund [9]. Figure 7 illustrates the simplified flow diagram of the TMS system together with the reaction network. The TMS system consists of two main sections: the reaction part and the separation part (here a decanter). The feed consists of the substrate 1-dodecene, the apolar solvent *n*-decane, the polar solvent dimethylformamide (in short DMF), and synthesis gas (CO/H<sub>2</sub>). The catalyst system consists of Rh(acac)(CO)<sub>2</sub> and the bidentate phosphite biphephos as ligand. During the reaction step, the system is single phase, thus homogeneous, so that no mass transport limitation occurs. During the separation step a lower temperature than that of the reactor is used and the system exhibits a miscibility gap and separates into two liquid phases, a polar and an apolar phase. The apolar phase contains the organic product which is purified in a down-stream process, while the catalyst mostly remains in the polar phase which is recycled. The main reaction is the catalyzed hydroformylation of the long-chain 1-dodecene to linear *n*-tridecanal. Besides the main reaction, isomerization to iso-dodecene, hydrogenation to *n*-dodecane and formation of the branched iso-aldehyde take place. Hernández and Engell [10] adapted the model developed by Hentschel et al. [11] to the TMS miniplant by considering the material balances of the above components as follows:

$$V_R \frac{dC_i}{dt} = \dot{V}_{in} C_{i,in} - \dot{V}_{out} C_{i,out} + V_R C_{cat} M_{cat} \sum_{l=1}^{n_{react}} \nu_{i,l} r_l \quad (25)$$

$$V_R \frac{dC_j}{dt} = -k_{eff} (C_j - C_j^{eq}) + \dot{V}_{in} C_{j,in} - \dot{V}_{out} C_{j,out} + V_R C_{cat} M_{cat} \sum_{l=1}^{n_{react}} \nu_{j,l} r_l \quad (26)$$

where (25) is applied to the liquid components 1-dodecene, *n*-tridecanal, iso-dodecene, *n*-dodecane, iso-aldehyde, decane, and DMF. (26) is applied to the gas components CO and H<sub>2</sub>. The algebraic equations involved in the model are as follows:

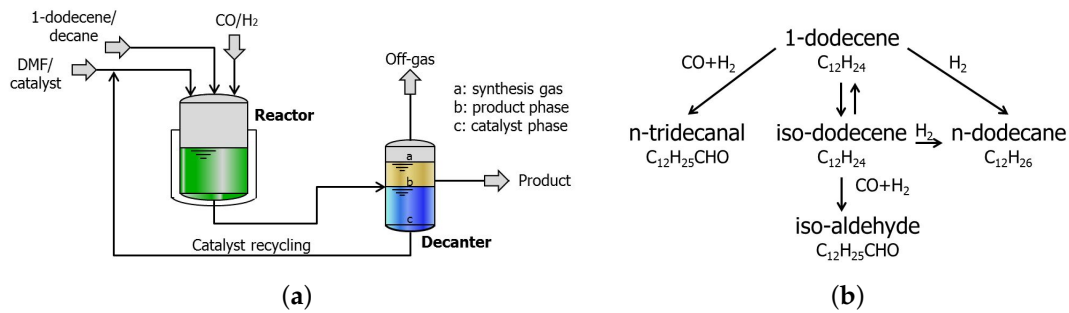
$$C_j^{eq} = \frac{P_j}{H_{j,0} \exp\left(-E_j/RT\right)} \quad (27)$$

$$C_{cat} = \frac{C_{Rh,precursor}}{1 + K_{cat,1}C_{CO} + K_{cat,2}C_{CO}/C_{H_2}} \quad (28)$$

$$K_i = \exp\left(A_{i,0} + \frac{A_{i,1}}{T_{decanter}} + A_{i,2}T_{decanter}\right) \quad (29)$$

$$n_{i,product} = \frac{K_i}{1 + K_i} n_{i,decanter} \quad (30)$$

$$n_{i,catalyst} = \frac{1}{1 + K_i} n_{i,decanter}. \quad (31)$$



**Figure 7.** (a) Thermomorphic multicomponent solvent (TMS) system; (b) Reaction network of the hydroformylation of 1-dodecene. Adapted from Hernández and Engell [10].

Table 2 lists all the symbols used in the model together with their explanations. The optimization problem is formulated as the minimization of the raw material and operating cost per unit of *n*-tridecanal produced

$$\min_{\mathbf{u}} \frac{Pr_{1-dodecene} \cdot F_{1-dodecene} + Pr_{Rh} \cdot F_{Rh} + C_{cooling} + C_{heating}}{F_{tridecanal}}, \quad (32)$$

where  $Pr_{1-dodecene}$  and  $Pr_{Rh}$  represent the prices of 1-dodecene and of the catalyst,  $F_{1-dodecene}$  and  $F_{Rh}$  are the molar flow rates,  $C_{cooling}$  and  $C_{heating}$  are the operating costs of cooling and heating, and  $F_{tridecanal}$  is the molar flow rate of *n*-tridecanal, the vector of the optimization variables  $\mathbf{u}$  consists of the reactor temperature, the catalyst dosage, the pressure and the composition of the synthesis gas. A sensitivity analysis was performed on the different model parameters and it was found that the gas solubility and the equilibrium constants for the catalyst species have the largest influence on the cost function. In our study of the MAWQA approach, the plant-model mismatch is created by decreasing the Henry coefficients  $H_{j,0}$  by 50% and ignoring the influence of CO on the active catalyst ( $K_{cat,2} = 0$ ) in the model that is used by the optimization. Table 3 lists the operating intervals of the optimization variables and compares the real optimum with the model optimum. In order to test the robustness of the approach to noisy data, the real cost data is assumed to be subject to a random error which is normally distributed with standard deviation  $\sigma$ .

**Table 2.** Model parameters and variables.

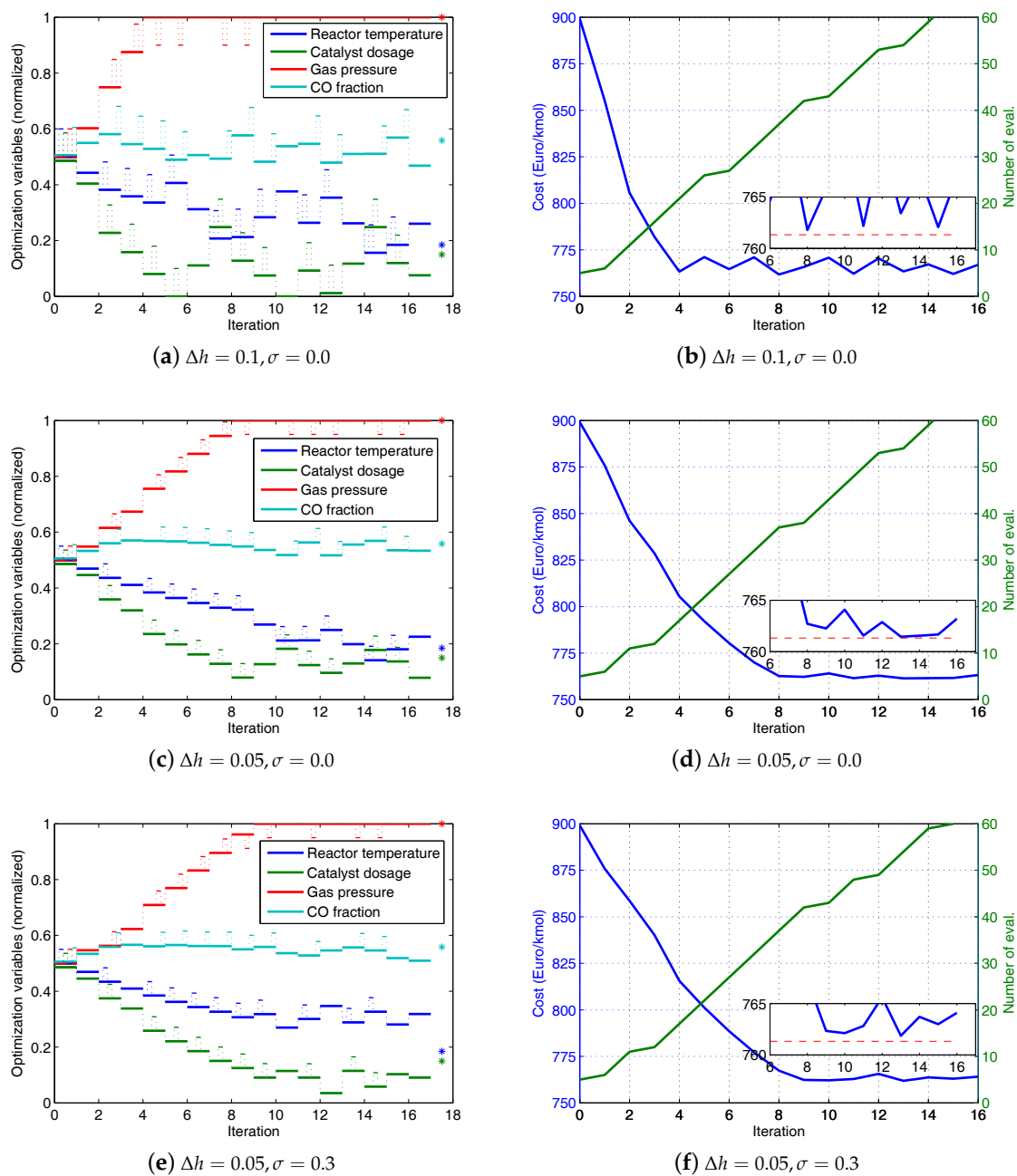
Symbol	Description
$C_i$	Concentration of 1-dodecene, <i>n</i> -tridecanal, iso-dodecene, <i>n</i> -dodecane, iso-aldehyde, decane, and DMF
$C_j$	Concentration of CO and H <sub>2</sub>
$C_j^{eq}$	Equilibrium concentration of CO and H <sub>2</sub> at the G/L interface
$V_R$	Reactor volume
$\dot{V}_{in}$	Inflow rate
$\dot{V}_{out}$	Outflow rate
$C_{cat}$	Concentration of the active catalyst
$M_{cat}$	Molar mass of the catalyst
$v_{i,l}$	Coefficients of the stoichiometric matrix [11]
$r_l$	Reaction rate of the <i>l</i> <sup>th</sup> reaction
$k_{eff}$	Mass transfer coefficient
$P_j$	Partial pressure
$T$	Reaction temperature
$H_{j,0}$	Henry coefficient
$C_{Rh,precursor}$	Concentration of the catalyst precursor
$K_{cat,1\setminus 2}$	Equilibrium constants
$T_{decanter}$	Decanter temperature
$A_{i,0\setminus 1\setminus 2}$	Coefficients regressed from experimental data [12]
$n_{i,product}$	Molar flows of the components in the product stream
$n_{i,catalyst}$	Molar flow of the components in the recycled catalyst stream
$n_{i,decanter}$	Molar flow of the components in the decanter inlet stream

**Table 3.** Operating variables and optimum.

Operating variable	Operating Interval	Initial Set-Point	Real Optimum	Model Optimum
Reactor temperature (°C)	85~105	95.0	88.64	85.10
Catalyst dosage (ppm)	0.25~2.0	1.1	0.51	0.49
Gas pressure (bar)	1.0~3.0	2.0	3.0	3.0
CO fraction	0.0~0.99	0.5	0.55	0.61
Cost (Euro/kmol)		899.04	761.33	818.88

Simulation results of the modifier-adaptation approach using finite-difference approximation of gradient are illustrated in Figure 8. The figures on the left show the evolutions of the normalized optimization variables with respect to the index of RTO iterations. The small pulses, which are superimposed on the evolutions, represent the additional perturbations required for the finite-difference approximations of the gradients. The star symbols at the right end mark the real optima. The figures on the right show the evolutions of the cost and the number of plant evaluations with respect to the index of the RTO iteration. The inset figure zooms in on the cost evolution, and the dashed line marks the real optimum of the cost. Three cases with different combinations of the step size of the perturbation and the noise in the data are considered. In the first case a large step-size,  $\Delta h = 0.1$ , is used and the data is free of noise ( $\sigma = 0.0$ ). From Figure 8a we can see that three of the four optimization variables are still away from their real optimal values after 16 iterations. The cost evolution in Figure 8b shows an oscillating behavior above the cost optimum. This indicates that the step-size is too large to enable an accurate estimation of the gradients.

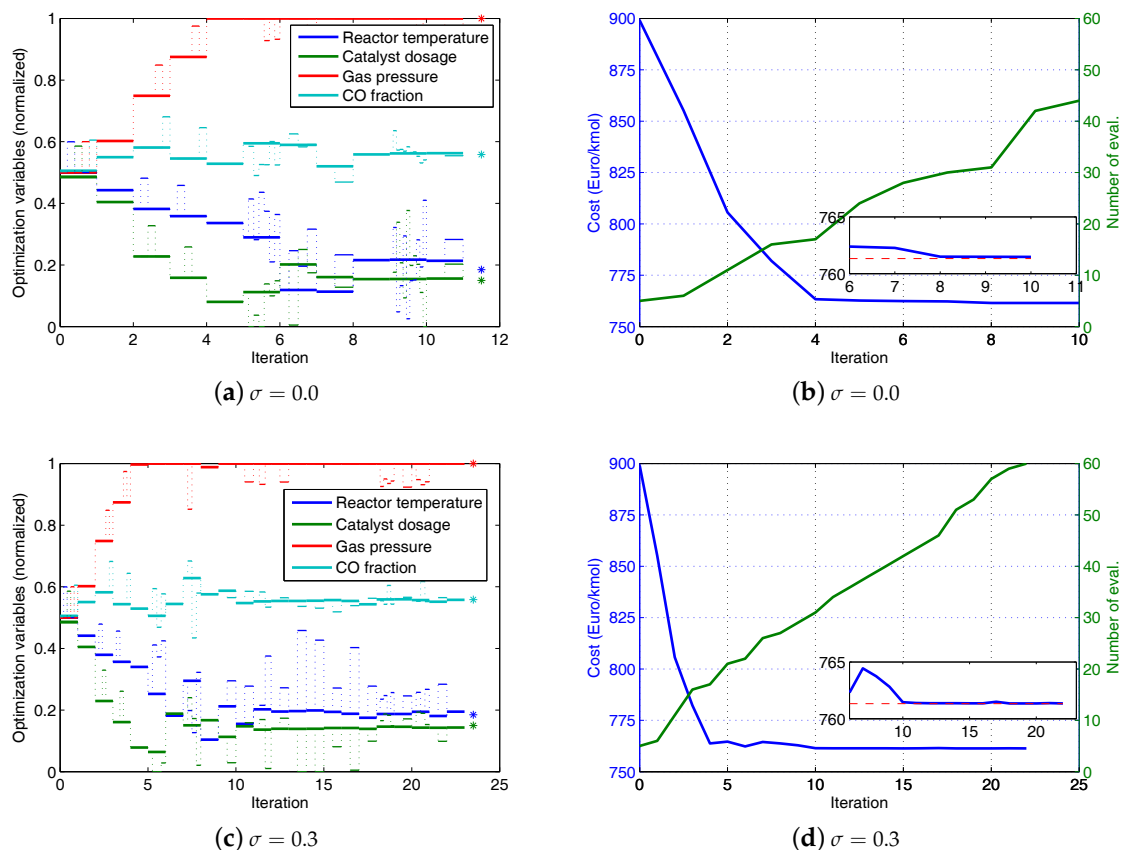
In the second case a reduced step-size ( $\Delta h = 0.05$ ) is tried. From Figure 8c we can see that the optimization variables attain their real optimal values at the 14th iteration. However, numerical errors of the simulation of the plant cause deviations during the following iterations. On the one hand, the use of a small step-size reduces the error of the finite-difference approximation of the gradients. On the other hand, the small step-size leads to a high sensitivity of the gradient approximation to errors. This is illustrated by the third case, in which the data contains a normally distributed error ( $\sigma = 0.3$ ). The optimization variables do not reach the real optima (see Figure 8e,f).



**Figure 8.** Modifier-adaptation optimization of the thermomorphic solvent system using finite-difference approximation of the gradients, the left figures show the evolutions of the normalized optimization variables  $\mathbf{u}$  (the additional set-point perturbations are represented by the small pulses which are superimposed on the set-point evolutions; the star symbols at the end mark the real optima), the right figures show the evolutions of the cost and the number of plant evaluations (the inset figure zooms in on the cost evolution, and the dashed line marks the real optimum).

Simulation results of the MAWQA algorithm are illustrated in Figure 9. The parameters of the MAWQA algorithm are listed in Table 4. The figures on the left show the evolutions of the normalized optimization variables with respect to iteration. The small pulses, which are superimposed on the evolutions, represent the additional plant evaluations, i.e., initial probes and explorative moves. The real optima are marked by the star symbols. The figures on the right show the evolutions of the cost and the number of plant evaluations with respect to the iterations. The inset figure

zooms in on the cost evolution, and the dashed line marks the real optimum of the cost. In the first 4 iterations, the modifier-adaptation approach using finite-difference approximation of gradient is run. Afterwards, enough sampled points (here  $(n_u + 1)(n_u + 2)/2 = 15$ ) are available for the quadratic approximation and the MAWQA approach is run. In the case of noise-free data, the MAWQA approach takes 8 iterations to reach the real optima approximately (see Figure 9a). The total number of plant evaluations is 30, much less than that used in the second case of the finite-difference approximations of gradients which requires 55 plant evaluations to reach a similar accuracy. Note that the additional plant evaluations at the 10th iteration are attributed to the shrinking of the regression region by Step 4 of the MAWQA algorithm. Figure 9c,d show the optimization results in the presence of noise. The MAWQA algorithm takes 10 iterations to reach the real optima.



**Figure 9.** MAWQA optimization of the thermomorphic solvent system, the left figures show the evolutions of the normalized optimization variables (the additional plant evaluations, i.e., initial probes and unsuccessful moves, are represented by the small pulses which are superimposed on the set-point evolutions; the star symbols at the end mark the real optima), the right figures show the evolutions of the cost and the number of plant evaluations (the inset figure zooms in on the cost evolution, and the dashed line marks the real optimum).

**Table 4.** Parameters of MAWQA.

Description	Symbol	Value
Screening parameter	$\Delta \mathbf{u}$	0.1
Search space parameter	$\gamma$	3
Perturbation step size	$\Delta h$	0.1

Finally, the MAWQA algorithm was tested with different values of the parameters. For each parameter, three values are considered and the costs after 30 plant evaluations are compared with

the real optimum. The results are summarized in Table 5. The increase of  $\Delta u$  leads to a decrease of the accuracy of the computed optima. This is due to the fact that the true mapping is only locally quadratic and therefore the use of too distant points can cause large approximation errors. The value of  $\gamma$  influences the rate of convergence since it is directly related to the size of the search space in each iteration. The value of  $\Delta h$  determines the accuracy of the finite-difference approximation of the gradients at the starting stage of the MAWQA approach. In the absence of noise, a small  $\Delta h$  is preferred. However, the search space is also determined by the distribution of the sampled points. A small  $\Delta h$  leads to a more constrained search space and therefore decreases the rate of convergence. The overall influence of  $\Delta h$  on the rate of convergence is a combination of both effects. Note that  $\Delta h$  does not influence the accuracy of the optima if enough plant evaluations are performed.

**Table 5.** Influence of MAWQA parameters.

	$\Delta u$	$\Delta h$	$\gamma$	Cost after 30 Eval.	Deviation from the True Optimum (%)
Initial	0.1	0.1	3	761.5	0.02%
$\uparrow \Delta u$	<b>0.15</b>	0.1	3	762.5	0.15%
$\uparrow\uparrow \Delta u$	<b>0.2</b>	0.1	3	762.9	0.21%
$\uparrow \Delta h$	0.1	<b>0.15</b>	3	761.4	0.01%
$\uparrow\uparrow \Delta h$	0.1	<b>0.2</b>	3	763.3	0.26%
$\downarrow \gamma$	0.1	0.1	<b>2</b>	763.2	0.24%
$\downarrow\downarrow \gamma$	0.1	0.1	<b>1</b>	772.1	1.14%

## 5. Conclusions

This paper focuses on the explorative moves when using the MAWQA algorithm to optimize a plant. The explorative moves are generated as a result of the estimation errors of the plant gradients. Three factors that influence the estimation errors are analyzed: the non-quadratic nature of the plant mappings, the measurement noise, and the distribution of the regression set. The screening algorithm is shown to take the accuracy of the gradient estimation into consideration and to choose the best set from the collected data. This ensures that the MAWQA algorithm takes only necessary explorative moves to improve the gradient estimations, instead of an expensive model-improvement step as in the derivative-free optimization algorithms. The simulation results for the optimization of the hydroformylation process with four decision variables demonstrate the promising performance of the MAWQA algorithm over the finite difference based modifier adaptation algorithm. Further studies will be focused on how to implement a dual control between the performance improvement and the explorative moves.

**Acknowledgments:** The research leading to these results has received funding from the European Commission under grant agreement number 291458 (ERC Advanced Investigator Grant MOBOCON) and by the Deutsche Forschungsgemeinschaft (DFG) via the funding of the Collaborative Research Center “Integrated Chemical processes in Liquid Multiphase Systems-InPROMPT” (TR63). The financial support by the ERC and by DFG is most gratefully acknowledged.

**Author Contributions:** W.G. conceived the theoretical analysis, performed the simulation studies and wrote the paper; R.H. provided the TMS model; R.H. and S.E. corrected the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

RTO	Real-time optimization
MAWQA	Modifier adaptation with quadratic approximation
TMS	Thermomorphic multicomponent solvent system
DMF	Dimethylformamide



## References

1. Chen, C.Y.; Joseph, B. On-line optimization using a two-phase approach: An application study. *Ind. Eng. Chem. Res.* **1987**, *26*, 1924–1930.
2. Roberts, P.D. An algorithm for steady-state system optimization and parameter estimation. *Int. J. Syst. Sci.* **1979**, *10*, 719–734.
3. Gao, W.; Engell, S. Iterative set-point optimization of batch chromatography. *Comput. Chem. Eng.* **2005**, *29*, 1401–1409.
4. Marchetti, A.; Chachuat, B.; Bonvin, D. Modifier-adaptation methodology for real-time optimization. *Ind. Eng. Chem. Res.* **2009**, *48*, 6022–6033.
5. Bunin, G.A.; François, G.; Bonvin, D. From discrete measurements to bounded gradient estimates: A look at some regularising structures. *Ind. Eng. Chem. Res.* **2013**, *52*, 12500–12513.
6. Gao, W.; Wenzel, S.; Engell, S. A reliable modifier-adaptation strategy for real-time optimization. *Comput. Chem. Eng.* **2016**, *91*, 318–328.
7. Conn, A.R.; Scheinberg, K.; Vicente, L.N. *Introduction to Derivative-Free Optimization*; SIAM: Philadelphia, PA, USA, 2009.
8. Ciarlet, P.G.; Raviart, P.A. General lagrange and hermite interpolation in  $R^n$  with applications to finite element methods. *Arch. Ration. Mech. Anal.* **1972**, *46*, 177–199.
9. Zagajewski, M.; Behr, A.; Sasse, P.; Wittman, J. Continuously operated miniplant for the rhodium catalyzed hydroformylation of 1-dodecene in a thermomorphic multicomponent solvent system (TMS). *Chem. Eng. Sci.* **2014**, *115*, 88–94.
10. Hernández, R.; Engell, S. Modelling and iterative real-time optimization of a homogeneously catalyzed hydroformylation process. In Proceedings of the 26th European Symposium on Computer Aided Process Engineering; Portorož, Slovenia, 12–15 June 2016.
11. Hentschel, B.; Kiedorf, G.; Gerlach, M.; Hamel, C.; Seidel-Morgenstern, A.; Freund, H.; Sundmacher, K. Model-based identification and experimental validation of the optimal reaction route for the hydroformylation of 1-Dodecene. *Ind. Eng. Chem. Res.* **2015**, *54*, 1755–1765.
12. Schäfer, E.; Brunsch, Y.; Sadowski, G.; Behr, A. Hydroformylation of 1-dodecene in the thermomorphic solvent system dimethylformamide/decane. Phase behavior-reaction performance-catalyst recycling. *Ind. Eng. Chem. Res.* **2012**, *51*, 10296–10306.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).