Journal of Process Control 54 (2017) 38–46

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

A feasible-side globally convergent modifier-adaptation scheme

Alejandro G. Marchetti [a,*,1], Timm Faulwasser [a,b], Dominique Bonvin [a]

[a] Laboratoire d'Automatique, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
[b] Institute for Applied Informatics, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 25 April 2016
Received in revised form
23 December 2016
Accepted 22 February 2017

Keywords:
Real-time optimization
Static optimization
Optimization under uncertainty
Global convergence
Feasible-side convergence

ABSTRACT

In the context of static real-time optimization (RTO) of uncertain plants, the standard modifier-adaptation scheme consists in adding first-order correction terms to the cost and constraint functions of a model-based optimization problem. If the algorithm converges, the limit is guaranteed to be a KKT point of the plant. This paper presents a general RTO formulation, wherein the cost and constraint functions belong to a certain class of convex upper-bounding functions. It is demonstrated that this RTO formulation enforces feasible-side global convergence to a KKT point of the plant. Based on this result, a novel modifier-adaptation scheme with guaranteed feasible-side global convergence is proposed. In addition to the first-order correction terms, quadratic terms are added in order to convexify and upper bound the cost and constraint functions. The applicability of the approach is demonstrated on a constrained variant of the Williams–Otto reactor for which standard modifier adaptation fails to converge in the presence of plant-model mismatch.

1. Introduction

In many industrial processes, there are strong economic incentives for finding the operating conditions that optimize some performance criterion while satisfying operating constraints. In many cases, the optimization problem relies on first-principles models and can be formulated as a nonlinear program (NLP). As examples, one can mention the steady-state optimization of continuously operating processes [8,11,23] and the optimization of batch and semi-batch processes using parameterized input profiles [28]. In the presence of plant-model mismatch and time-varying disturbances, it is necessary to continuously guide the operation towards the optimum. For this purpose, several model-based real-time optimization (RTO) schemes have been proposed, which iteratively update the model-based NLP problem using some adaptation strategy based on measurements.

The standard strategy used in industry is the two-step approach of parameter estimation followed by re-optimization [9,11]. The parameters of a first-principles model are estimated based on measurements available at the current operating point, and the updated model is used in the optimization problem to compute the next operating point. However, it is well known that, in the presence of structural plant-model mismatch, the two-step approach does not in general converge to the plant optimum [4,12,17]. In response to this deficiency, a modified two-step approach known as *Integrated Systems Optimization and Parameter Estimation* (ISOPE) was proposed by Roberts and co-workers [23,24]. ISOPE incorporates plant-gradient information in a gradient-modification term that is added to the cost function of the optimization problem, such that the Karush–Kuhn–Tucker (KKT) optimality conditions *for the plant* are satisfied upon convergence. The ISOPE algorithm was simplified by Tatjewski [30] by eliminating the parameter estimation problem. Gao and Engell [16] extended the approach of Tatjewski [30] to problems with process-dependent constraints by including first-order correction terms to the constraints in the optimization problem. Finally, Marchetti et al. [17] used the same type of first-order correction terms in the cost *and* constraint functions, and labeled the approach *Modifier Adaptation*, providing a comprehensive analysis of many of the algorithm's properties, such as optimality upon convergence, model-adequacy conditions, and necessary conditions for local asymptotic convergence. Since then, many variants of modifier adaptation have been proposed, such as *dual* modifier adaptation [18,25,19], *directional* modifier adaptation [10], *nested* modifier adaptation [22], and *second-order* modifier adaptation [13].

Bunin et al. [7] made the crucial observation that basically none of the available RTO algorithms can provide practical or even conceptual guarantees for converging to the plant optimum with all

* Corresponding author.
E-mail addresses: marchetti@cifasis-conicet.gov.ar (A.G. Marchetti), timm.faulwasser@kit.edu (T. Faulwasser), dominique.bonvin@epfl.ch (D. Bonvin).
[1] Also with the French-Argentine International Center for Information and Systems Sciences (CIFASIS), CONICET-UNR, S2000EZP Rosario, Argentina.

http://dx.doi.org/10.1016/j.jprocont.2017.02.013
0959-1524/

the iterations being feasible points for the plant, and modifier adaptation is no exception to this assertion. A necessary condition for modifier adaptation to converge to the plant optimum is that the model be adequate, in the sense that it should predict the correct curvature of the cost function in the vicinity of the converged point [17]. François and Bonvin [15] showed that model adequacy is satisfied if the cost and constraint functions of the model used for optimization are selected as convex functions. A general abstract sufficient condition for the convergence of modifier adaptation is given in [13]. However, this condition is difficult to verify and there is no practical way to enforce it. Global convergence conditions have also been proposed in [6] by implementing modifier adaptation as a trust-region algorithm, and exploiting the global convergence results available in trust-region theory. However, the results in [13,6] cannot guarantee *feasible-side* convergence of modifier adaptation. Bunin et al. [7] proposed a set of *sufficient conditions for feasibility and optimality* (SCFO) that, when met by any RTO algorithm, would enforce feasible-side global convergence. The feasible solution computed via RTO is projected onto a cone of feasible descent directions for the cost and constraint functions, and then a line-search step is conducted to improve the cost and remain feasible.

In the mathematical programming literature, a number of sequential convex programming methods have been proposed for solving inequality-constrained nonconvex NLPs [2,21,29]. The idea therein is to replace the nonconvex cost and/or constraints in the optimization problem by convex inner approximations, which results in interior-side monotone convergence to a KKT point. In the present paper, using similar ideas, we present a feasible-side globally convergent RTO formulation, wherein the cost and constraint functions belong to a certain class of convex upper-bounding functions. We propose to construct the required upper-bounding functions by adding quadratic terms to the modified cost and constraint functions used in standard modifier adaptation. The main contribution of the present paper is a modifier-adaptation algorithm guaranteeing global feasible-side convergence to a KKT point of the plant in the presence of plant-model mismatch.

The rest of the paper is organized as follows. Section 2 recalls the RTO problem, presents the necessary conditions of optimality, and introduces the main definitions and assumptions used in this work. An RTO scheme based on using general convex upper-bounding functions is presented and analyzed in Section 3. In particular, a proof of feasible-side global convergence is provided. On the grounds of this result, a modifier-adaptation algorithm with convex upper-bounding functions is proposed in Section 4. The new modifier-adaptation algorithm is applied to an optimization problem that is defined for the Williams–Otto reactor in Section 5. Finally, Section 6 concludes the paper.

2. Preliminaries

2.1. Real-time optimization

The purpose of static real-time optimization (RTO) is to optimize process operation by finding the solution to the following optimization problem

$$\begin{aligned} \min_{\mathbf{u}} \quad & \Phi_p(\mathbf{u}) := \phi(\mathbf{u}, \mathbf{y}_p(\mathbf{u})) \\ \text{s.t.} \quad & G_{p,i}(\mathbf{u}) := g_i(\mathbf{u}, \mathbf{y}_p(\mathbf{u})) \le 0, \quad i = 1, \ldots, n_g, \\ & \mathbf{u} \in \mathcal{U}, \end{aligned} \tag{1}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$ denotes the decision (or input) variables; $\mathbf{y}_p \in \mathbb{R}^{n_y}$ are the measured output variables; $\phi : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$ is the cost function to be minimized; $g_i : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$, $i = 1, \ldots, n_g$, is the set of process-dependent inequality constraints; and $\mathcal{U} = \{\mathbf{u} \in \mathbb{R}^{n_u} : \mathbf{u}^{\mathrm{L}} \le \mathbf{u} \le \mathbf{u}^{\mathrm{U}}\}$. The notation $(\cdot)_p$ is used throughout for variables associated with the plant.

This formulation assumes that ϕ and g_i are known functions of $\mathbf{u}$ and $\mathbf{y}_p$, i.e., they can be directly measured or evaluated from the knowledge of $\mathbf{u}$ and the measurement of $\mathbf{y}_p$. However, the steady-state input-output mapping of the plant $\mathbf{y}_p(\mathbf{u})$ is typically unknown, and only an approximate nonlinear steady-state model is available:

$$\boldsymbol{F}(\mathbf{x}, \mathbf{u}) = \mathbf{0}, \tag{2a}$$

$$\mathbf{y} = \boldsymbol{H}(\mathbf{x}, \mathbf{u}), \tag{2b}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$ are the state variables, and $\mathbf{y} \in \mathbb{R}^{n_y}$ are the output variables predicted by the model. For $\mathbf{u}$ given, the solution to the system (2a) is given by:

$$\mathbf{x} = \boldsymbol{\xi}(\mathbf{u}), \tag{3}$$

where $\boldsymbol{\xi}$ is an operator expressing the steady-state mapping between $\mathbf{u}$ and $\mathbf{x}$. The steady-state input-output mapping predicted by the model can now be expressed as:

$$\mathbf{y}(\mathbf{u}) := \boldsymbol{H}(\boldsymbol{\xi}(\mathbf{u}), \mathbf{u}). \tag{4}$$

The model-based counterpart of Problem (1) is given by the following NLP:

$$\begin{aligned} \min_{\mathbf{u}} \quad & \Phi(\mathbf{u}) := \phi(\mathbf{u}, \mathbf{y}(\mathbf{u})) \\ \text{s.t.} \quad & G_i(\mathbf{u}) := g_i(\mathbf{u}, \mathbf{y}(\mathbf{u})) \le 0, \quad i = 1, \ldots, n_g, \\ & \mathbf{u} \in \mathcal{U}, \end{aligned} \tag{5}$$

In the presence of plant-model mismatch, the solution to Problem (5) does not generally match the solution to Problem (1). In real-time optimization, the solution to Problem (1) is approached by iteratively re-evaluating the operating point applied to the plant. Let $\mathbf{u}_k$ denote the steady-state operating point applied to the plant at the kth RTO iteration. The next optimal RTO solution is obtained by solving the following model-based optimization problem:

$$\begin{aligned} \mathbf{u}^\star_{k+1} = \arg\min_{\mathbf{u}} \quad & \Phi_k(\mathbf{u}) \\ \text{s.t.} \quad & G_{i,k}(\mathbf{u}) \le 0, \quad i = 1, \ldots, n_g, \\ & \mathbf{u} \in \mathcal{U}. \end{aligned} \tag{6}$$

In order to deal with plant-model mismatch, the models used for the cost function Φ_k and the constraint functions $G_{i,k}$, $i = 1, \ldots, n_g$, are typically updated at each RTO iteration k, based on collected measurements. Examples of updating strategies are the computation of new model parameters based on available plant data—i.e. two-step approaches [9,11]—and the computation of first-order correction terms, i.e. modifier-adaptation approaches [17].

For stability reasons, $\mathbf{u}^\star_{k+1}$ is usually filtered before it is applied to the plant:

$$\mathbf{u}_{k+1} = \mathbf{u}_k + K(\mathbf{u}^\star_{k+1} - \mathbf{u}_k), \tag{7}$$

where $K \in (0, 1]$ is the filter gain, with $K = 1$ meaning no filtering. The combination of (6) and (7) constitutes an RTO algorithm.

In general, the design of any RTO algorithm should enforce the following desirable properties:

(i) **Plant optimality**: Despite structural mismatch between (1) and (6), a KKT point of (1) is reached upon convergence of (6)–(7).
(ii) **Plant feasibility**: All RTO iterates $\mathbf{u}_k$ satisfy the constraints of (1).
(iii) **Monotonic cost improvement:** The performance is required to improve between consecutive RTO iterates.

Besides these important basic properties, one would like to have sufficiently fast convergence and sufficient robustness with respect to gradient errors. In the present paper, however, the investigation will be limited to the design of RTO algorithms enforcing (i)–(iii).

2.2. Necessary conditions of optimality

Associated with the constraints $G_{p,i}$ of Problem (1), let us define the sets

$$\mathcal{G}_{p,i} = \{\mathbf{u} \in \mathbb{R}^{n_u} : G_{p,i}(\mathbf{u}) \leq 0\}, \quad i = 1, \ldots, n_g, \tag{8}$$

with which the feasibility set of the plant can be written as

$$\mathcal{F}_p = \left(\bigcap_{i=1}^{n_g} \mathcal{G}_{p,i} \right) \cap \mathcal{U}. \tag{9}$$

In Problem (1), the inequality constraints $G_{p,i}(\mathbf{u}) \leq 0$ and the input box constraints $\mathbf{u} \in \mathcal{U}$ are considered separately because the RTO approach that will be presented will treat them in a different way. However, in order to simplify some of the developments presented in the paper, it is convenient to group all the inequality constraints in the vector $\boldsymbol{H}_p(\mathbf{u})$ as done in the following equivalent optimization problem

$$\begin{aligned} \min_{\mathbf{u}} \quad & \Phi_p(\mathbf{u}) \\ \text{s.t.} \quad & H_{p,j}(\mathbf{u}) \leq 0, \quad j = 1, \ldots, n_h. \end{aligned} \tag{10}$$

Local minima of Problem (1) can be characterized via the necessary conditions of optimality (NCO) [1]. To this end, let us denote the set of active constraints at some point $\mathbf{u}$ by

$$\mathcal{A}(\mathbf{u}) = \left\{ j \in \{1, \ldots, n_h\} \,|\, H_{p,j}(\mathbf{u}) = 0 \right\}. \tag{11}$$

If $\mathbf{u}$ is a (local) minimum of (1), then there exists a scalar ν_0 and a vector $\boldsymbol{\nu} = [\nu_1, \ldots, \nu_{n_h}]^\mathsf{T}$ such that the following Fritz–John conditions hold [1]:

$$\nu_0 \frac{\partial \Phi_p}{\partial \mathbf{u}}(\mathbf{u}) + \sum_{j=1}^{n_h} \nu_j \frac{\partial H_{p,j}}{\partial \mathbf{u}}(\mathbf{u}) = \mathbf{0} \tag{12a}$$

$$H_{p,j}(\mathbf{u}) \leq 0, \quad \forall j \in \{1, \ldots, n_h\} \tag{12b}$$

$$\nu_j H_{p,j}(\mathbf{u}) = 0, \quad \forall j \in \{1, \ldots, n_h\} \tag{12c}$$

$$\nu_0, \nu_j \geq 0, \quad \forall j \in \{1, \ldots, n_h\} \tag{12d}$$

$$(\nu_0, \boldsymbol{\nu}) \neq (0, \mathbf{0}). \tag{12e}$$

The possibility that conditions (12) hold trivially with $\nu_0 = 0$ at some non-optimal solutions is eliminated by introducing a constraint qualification.

Definition 1 *(Linear independence constraint qualification).* The gradients of the active constraints, $\frac{\partial H_{p,j}}{\partial \mathbf{u}}(\mathbf{u})$ for $j \in \mathcal{A}(\mathbf{u})$, are linearly independent.

If a constraint qualification such as linear independence holds at $\mathbf{u}$, then one can set $\nu_0 = 1$, which satisfies (12e) automatically. This way, the Fritz–John conditions (12) reduce to the Karush–Kuhn–Tucker (KKT) conditions [1]. A point $\mathbf{u}$ satisfying these conditions is called a KKT point.

2.3. Plant and model assumptions

Our further developments are based on the technical assumptions introduced next.

Assumption 1 *(Plant properties).* The plant optimization problem (1), or equivalently Problem (10), satisfies the following conditions:

a) For all $\mathbf{u} \in \mathcal{U}$, the plant has no steady-state output multiplicities.
b) Φ_p and $G_{p,i}$, $i = 1, \ldots, n_g$, are twice continuously differentiable functions on $\mathcal{U}$.
c) $\mathcal{F}_p$ is a nonempty compact set.
d) At any boundary point of $\mathcal{F}_p$, the linear independence constraint qualification holds.

Assumption 1 has the following implications:

- For all $\mathbf{u} \in \mathcal{U}$, the steady-state mappings Φ_p and $G_{p,i}$, $i = 1, \ldots, n_g$, are single-valued functions.
- Any point $\bar{\mathbf{u}}_p$ that satisfies the Fritz–John optimality conditions of Problem (1) satisfies the linear independence constraint qualification. In other words, any Fritz–John point is a KKT point.
- At any boundary point of $\mathcal{F}_p$, the cone of interior directions of $\mathcal{F}_p$ is nonempty. That is,

$$\mathcal{C}_0 = \left\{ \mathbf{d} : \frac{\partial H_{p,j}}{\partial \mathbf{u}}(\mathbf{u})\,\mathbf{d} < 0, \quad \text{for } j \in \mathcal{A}(\mathbf{u}) \right\} \neq \emptyset. \tag{13}$$

  The condition $\mathcal{C}_0 \neq \emptyset$ is equivalent to the Mangasarian-Fromovitz constraint qualification [3]. Notice that $\mathbf{d} \in \mathcal{C}_0$ implies that, for some sufficiently small $\lambda > 0$, the inclusion $\mathbf{u} + \lambda \mathbf{d} \in \text{int}(\mathcal{F}_p)$ holds.
- The gradients of the active constraints do not vanish, that is, $\frac{\partial G_{p,i}}{\partial \mathbf{u}}(\mathbf{u}) \neq \mathbf{0}$, for all $i \in \mathcal{A}(\mathbf{u})$.

Assumption 2 *(Model properties).* The model satisfies the following conditions:

a) For all $\mathbf{u} \in \mathcal{U}$, the steady-state nonlinear model equations (2a) have a unique solution.
b) Φ and G_i, $i = 1, \ldots, n_g$, are twice continuously differentiable functions on $\mathcal{U}$.

Assumption 3 *(Perfect gradient estimates).* The constrained values and the cost and constraint gradients of the plant are perfectly known at each RTO iteration.

2.4. Standard modifier-adaptation scheme

Unlike the two-step approach of repeated parameter estimation and optimization, modifier-adaptation schemes do not rely on updating model parameters. Instead, these schemes are based on iteratively modifying the cost and constraint functions using correction terms. Standard modifier-adaptation uses measurements to update first-order correction terms that are added to the model cost and constraint functions. At the kth RTO iteration, the next optimal RTO inputs are computed by solving the following *modified* optimization problem [17]:

$$\mathbf{u}_{k+1}^\star = \underset{\mathbf{u}}{\text{argmin}} \quad \Phi_m(\mathbf{u}) := \Phi(\mathbf{u}) + \varepsilon_k^\Phi + (\boldsymbol{\lambda}_k^\Phi)^\mathsf{T}(\mathbf{u} - \mathbf{u}_k) \tag{14a}$$

s.t.

$$G_{m,i}(\mathbf{u}) := G_i(\mathbf{u}) + \varepsilon_k^{G_i} + (\boldsymbol{\lambda}_k^{G_i})^\mathsf{T}(\mathbf{u} - \mathbf{u}_k) \leq 0, \quad i = 1, \ldots, n_g, \tag{14b}$$

$$\mathbf{u} \in \mathcal{U}, \tag{14c}$$

with

$$\varepsilon_k^\Phi = \Phi_p(\mathbf{u}_k) - \Phi(\mathbf{u}_k), \tag{15a}$$

$$\varepsilon_k^{G_i} = G_{p,i}(\mathbf{u}_k) - G_i(\mathbf{u}_k), \tag{15b}$$

$$(\boldsymbol{\lambda}_k^\Phi)^\mathsf{T} = \frac{\partial \Phi_p}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial \Phi}{\partial \mathbf{u}}(\mathbf{u}_k), \tag{15c}$$

$$(\boldsymbol{\lambda}_k^{G_i})^{\mathsf{T}} = \frac{\partial G_{p,i}}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial G_i}{\partial \mathbf{u}}(\mathbf{u}_k), \tag{15d}$$

where the scalars ε_k^{Φ} and $\varepsilon_k^{G_i}$, $i=1, \ldots, n_g$, are the zero-order modifiers, and the vectors $\boldsymbol{\lambda}_k^{\Phi}$, and $\boldsymbol{\lambda}_k^{G_i}$, $i=1, \ldots, n_g$, are the first-order modifiers. The optimal RTO input $\mathbf{u}_{k+1}^{\star}$ is filtered with a gain $K \in (0, 1]$:

$$\mathbf{u}_{k+1} = \mathbf{u}_k + K(\mathbf{u}_{k+1}^{\star} - \mathbf{u}_k). \tag{16}$$

Alternatively, it is possible to filter the modifiers (15a)–(15d), as proposed in [17]. The appeal of modifier adaptation lies in its ability to reach a KKT point of the plant upon convergence. For the sake of completness, we restate a result presented in [17].

Lemma 1 (MA convergence → KKT matching [17]). *Let Assumptions 1–3 hold. Then, at any fixed point $\bar{\mathbf{u}}$, the KKT conditions of the RTO problem (14) match those of the plant problem (1). Furthermore, if the modifier-adaptation algorithm (14)–(16) converges, it converges to a KKT point of the plant.*

3. RTO using convex upper bounds

Next, we propose a general formulation of the RTO scheme (6) that enforces feasible-side global convergence to a KKT point of the plant. Our approach is based on a convex inner approximation of the feasible set and a convex upper bound on the cost.

Definition 2 *(Convex upper-bounding function).* Let the function $f : \mathcal{U} \to \mathbb{R}$ be continuously differentiable. Then, any differentiable convex function $f^{\mathrm{U}} : \mathcal{U} \times \mathcal{U} \to \mathbb{R}$ such that, for all $\mathbf{u}, \mathbf{u}_k \in \mathcal{U}$,

$$f^{\mathrm{U}}(\mathbf{u}_k, \mathbf{u}_k) = f(\mathbf{u}_k), \tag{17a}$$

$$\frac{\partial f^{\mathrm{U}}}{\partial \mathbf{u}}(\mathbf{u}_k, \mathbf{u}_k) = \frac{\partial f}{\partial \mathbf{u}}(\mathbf{u}_k), \tag{17b}$$

$$f^{\mathrm{U}}(\mathbf{u}_k, \mathbf{u}) \geq f(\mathbf{u}). \tag{17c}$$

is said to be a convex upper-bounding function of f. If, additionally, the function f^{U} is strictly convex, then it is called a strictly convex upper-bounding function of f.

Note that the upper-bounding function f^{U} is always a function of two arguments: the point $\mathbf{u}_k$, at which the bound f^{U} is required to match f up to first order (17a)–(17b), and the input $\mathbf{u}$. Subsequently, we require convex upper bounding of the constraints and strict convex upper bounding of the cost in order to enforce uniqueness of optimal solutions to the RTO optimization problem.

Assumption 4 *(Convex upper bounds).* At each RTO iteration k, one can compute a strictly convex upper-bounding function of the plant cost Φ_p and convex upper-bounding functions of the plant constraints $G_{p,i}$, $i=1, \ldots, n_g$.

Henceforth, for the sake of readability, we surpress the explicit dependence of the upper-bounding functions on $\mathbf{u}_k$. Instead, we write $\Phi_k^{\mathrm{U}}(\mathbf{u})$ to denote the strictly convex upper-bounding function of the plant cost, and $G_{i,k}^{\mathrm{U}}(\mathbf{u})$ to denote the convex upper-bounding function of the ith constraint at the iterate $\mathbf{u}_k$.

Remark 1 *(Nominal RTO analysis).* Assumption 4 implies that the gradients of Problem (1) are exactly known, cf. (17b). Since the plant gradients cannot be measured directly in real applications, they must be estimated, for example using finite differences. In other words, Assumption 4 implies that this paper focuses on nominal feasibility and nominal global convergence guarantees, that is, we consider the ideal scenario wherein Assumption 3 holds. The consideration of gradient uncertainty is beyond the scope of this paper (note that an RTO scheme that is robust to uncertainty in the cost gradient has been proposed recently [26]).

Remark 2 *(Upper bounding the cost and all constraints).* The properties of the convex upper-bounding functions Φ_k^{U} and $G_{i,k}^{\mathrm{U}}$, $i=1, \ldots, n_g$, are the same as those required by the general inner approximation algorithm presented by Marks and Wright [21]. Since the aim in that paper was to locate KKT solutions to nonconvex programs, only the nonconvex functions in the optimization problem were replaced by convex upper-bounding functions. The situation is different here since the cost and constraint functions of the plant are typically unknown in RTO applications. Hence, even if some of these unknown functions are convex, it is still necessary to replace them in the RTO problem by <u>known</u> convex upper-bounding functions.

3.1. RTO formulation

To solve RTO Problem (1), we propose to replace the iterative solution to (6) by the iterative solution to the following convex approximation:

$$\begin{aligned} \mathbf{u}_{k+1} = \arg\min_{\mathbf{u}} \quad & \Phi_k^{\mathrm{U}}(\mathbf{u}) \\ \text{s.t.} \quad & G_{i,k}^{\mathrm{U}}(\mathbf{u}) \leq 0, \quad i = 1, \ldots, n_g, \\ & \mathbf{u} \in \mathcal{U}. \end{aligned} \tag{18}$$

The feasible sets associated with the upper-bounding constraint functions $G_{i,k}^{\mathrm{U}}$ are

$$\mathcal{G}_{i,k} = \{\mathbf{u} \in \mathbb{R}^{n_u} : G_{i,k}^{\mathrm{U}}(\mathbf{u}) \leq 0\}, \quad i = 1, \ldots, n_g. \tag{19}$$

Hence, the feasible set for the kth RTO iteration can be written as

$$\mathcal{F}_k = \left(\bigcap_{i=1}^{n_g} \mathcal{G}_{i,k}\right) \cap \mathcal{U}. \tag{20}$$

Notice that the solution to (18) can be applied directly to the plant without the need to include a filter. Hence, we shall also refer to (18) as an RTO algorithm.

The following three lemmas present useful properties of this algorithm.

Lemma 2 (Plant feasibility). *Let Assumption 4 hold. Then, at each RTO iteration k, the solution $\mathbf{u}_{k+1}$ to Problem (18) is feasible for the plant. If, additionally, Assumption 1 holds, then the feasible set $\mathcal{F}_k$ has a non-empty interior.*

Proof. The proof follows directly by noting that, from (17c), enforcing $G_{i,k}^{\mathrm{U}}(\mathbf{u})$ to be non-positive enforces $G_{p,i}(\mathbf{u}) \leq 0$. Hence, for the ith constraint, the set $\mathcal{G}_{i,k}$ is a convex inner approximation of the plant set $\mathcal{G}_{p,i}$, i.e., $\mathcal{G}_{i,k} \subseteq \mathcal{G}_{p,i}$, and the feasible set $\mathcal{F}_k$ is a convex inner approximation of the feasible set of the plant, i.e., $\mathcal{F}_k \subseteq \mathcal{F}_p$.

By construction, the convex upper-bounding functions $G_{i,k}^{\mathrm{U}}$ match the plant constraints $G_{p,i}$ up to first order. Hence, non-emptiness of the interior cone (13) implies that its inner approximation $\mathcal{F}_k$ also has a non-empty interior. □

Lemma 3 (Uniqueness of solutions). *Let Assumption 4 hold. Then, at each RTO iteration k, Problem (18) has a unique global optimal solution that is also a strong local minimum.*

Proof. It follows from the constraint functions $G_{i,k}^{\mathrm{U}}$ being convex that the zero-level sets $\mathcal{G}_{i,k}$ are convex. Hence, the feasibility set $\mathcal{F}_k$ is also convex as it corresponds to the intersection of convex sets. Furthermore, it follows from the cost function Φ_k^{U} being strictly convex on $\mathcal{F}_k$ that $\mathbf{u}_{k+1}$ is the unique global solution to Problem (18), and also a strong local minimum (see e.g. Theorem 3.4.2 in [1]). □

Lemma 4 (RTO convergence ⇔ KKT matching). *Let Assumptions 1 and 4 hold. Furthermore, let the input $\bar{\mathbf{u}}$ be a fixed point of the RTO algorithm (18) and $\bar{\mathbf{u}}_p$ be a KKT point of the plant problem (1). Then,*

a) *At* $\bar{\mathbf{u}}_p$, *the KKT conditions of the RTO problem* (18) *match those of the plant problem* (1).
b) *The input* $\bar{\mathbf{u}}$ *is a fixed point of the RTO algorithm* (18) *if and only if it is a KKT point of the plant, i.e.,* $\bar{\mathbf{u}} = \bar{\mathbf{u}}_p$.

Proof. Part *a*). $\bar{\mathbf{u}}_p$ is a KKT point of the plant problem. However, with Assumption 4 and the properties (17), $\bar{\mathbf{u}}_p$ is also a KKT point of the RTO problem (18) with the same Lagrange multipliers. In other words, the KKT conditions of the model-based RTO problem (18) match those of the plant problem.

In order to prove the "if" part of *b*), let us assume that, at the *k*th iteration, the RTO problem (18) has produced the input $\mathbf{u}_k = \bar{\mathbf{u}}_p$. The KKT matching of Part *a*) implies that $\bar{\mathbf{u}}_p$ is a KKT point of the RTO problem (18). Since Φ_k^{U} and $G_{i,k}^{\mathrm{U}}$, $i=1, \ldots, n_g$, are convex and differentiable functions, it follows from the first-order KKT sufficient conditions of optimality (see Theorem 4.2.16 in [1]) that $\bar{\mathbf{u}}_p$ is the (unique) global minimum of Problem (18). Hence, $\bar{\mathbf{u}}_p = \mathbf{u}_k = \mathbf{u}_{k+1} = \bar{\mathbf{u}}$, and $\bar{\mathbf{u}}_p$ is a fixed point of the RTO algorithm (18).

To prove the "only if" part of *b*), let us assume that $\bar{\mathbf{u}}$ is a fixed point of the RTO algorithm (18). Hence, $\bar{\mathbf{u}}$ must satisfy the Fritz–John optimality conditions of Problem (18). The properties (17) imply that $\bar{\mathbf{u}}$ is also a Fritz–John point for the plant, and by Assumption 1c, $\bar{\mathbf{u}}$ is a KKT point of the plant, that is, $\bar{\mathbf{u}} = \bar{\mathbf{u}}_p$. □

The preceding technical lemmas prepare the ground for analyzing RTO schemes based on convex upper-bounding functions.

3.2. Feasible-side global convergence

We show next that the RTO algorithm (18) guarantees feasible-side global convergence to a KKT point of the plant.

Theorem 1 (Feasible-side global convergence). *Let Assumptions 1, 3 and 4 hold. Furthermore, let the input* $\bar{\mathbf{u}}$ *be a fixed point of the RTO algorithm* (18) *and* $\bar{\mathbf{u}}_p$ *be a KKT point of the plant problem* (1). *Then, for any feasible initial point* $\mathbf{u}_0 \in \mathcal{F}_p$, *the RTO algorithm* (18) *satisfies the following properties:*

a) *All RTO iterates satisfy the plant constraints.*
b) *The RTO iterates computed according to* (18) *converge to a KKT point of the plant problem* (1).
c) *The plant cost decreases monotonically at each iteration.*

Proof. The proof proceeds over four steps.

Step 1: Plant feasibility expressed in Property *a*) has been proven in Lemma 2.
Step 2: If $\mathbf{u}_k = \bar{\mathbf{u}}_p$, then it follows from Lemma 4 that $\mathbf{u}_k$ is both a KKT point of Problem (18) and a fixed point of the RTO algorithm (18). In other words, the RTO algorithm has converged to a KKT point of the plant.
Step 3: If $\mathbf{u}_k \neq \bar{\mathbf{u}}_p$, then it follows from Lemma 4 that $\mathbf{u}_k$ is not a KKT point of Problem (18). Let us define the following sets:

$$\mathcal{S}_k = \left\{\mathbf{u} \in \mathcal{F}_k : \Phi_k^{\mathrm{U}}(\mathbf{u}) \leq \Phi_k^{\mathrm{U}}(\mathbf{u}_k)\right\} \tag{21a}$$

$$\mathcal{S}_{p,k} = \left\{\mathbf{u} \in \mathcal{F}_p : \Phi_p(\mathbf{u}) \leq \Phi_p(\mathbf{u}_k)\right\} \tag{21b}$$

$$\mathcal{S}'_k = \left\{\mathbf{u} \in \mathcal{F}_k : \Phi_k^{\mathrm{U}}(\mathbf{u}) < \Phi_k^{\mathrm{U}}(\mathbf{u}_k)\right\} \tag{21c}$$

$$\mathcal{S}'_{p,k} = \left\{\mathbf{u} \in \mathcal{F}_p : \Phi_p(\mathbf{u}) < \Phi_p(\mathbf{u}_k)\right\}. \tag{21d}$$

Since $\mathcal{F}_k \subseteq \mathcal{F}_p$ and Φ_k^{U} is an upper-bounding function of Φ_p, it follows that $\mathcal{S}_k \subseteq \mathcal{S}_{p,k}$. Note that $\mathcal{S}_k \subseteq \mathcal{S}_{p,k}$ implies that $\mathcal{S}'_k \subseteq \mathcal{S}'_{p,k}$. It follows that, if $\mathbf{u} \in \mathcal{S}'_k$, there is a strict decrease in the plant cost:

$$\mathbf{u} \in \mathcal{S}'_k \Rightarrow \Phi_p(\mathbf{u}) < \Phi_p(\mathbf{u}_k). \tag{22}$$

Since ϕ_k^{U} is strictly convex and the feasibility set $\mathcal{F}_k$ is convex, if $\mathbf{u}_k$ is not a KKT point, then $\mathcal{S}'_k \neq \emptyset$. Hence, if $\mathbf{u}_k \neq \bar{\mathbf{u}}_p$, the solution to Problem (18) satisfies $\mathbf{u}_{k+1} \in \mathcal{S}'_k$ and therefore

$$\Phi_p(\mathbf{u}_{k+1}) < \Phi_p(\mathbf{u}_k), \quad \text{if} \quad \mathbf{u}_k \neq \bar{\mathbf{u}}_p, \tag{23}$$

which proves Property *c*).

Step 4: Since Φ_p is continuous on the compact set $\mathcal{U}$, then, by Weierstrass' Theorem [3], Φ_p has a minimum and a maximum on $\mathcal{U}$. Hence, if $\mathbf{u}_0 \neq \bar{\mathbf{u}}_p$, the sequence $\{\Phi_p(\mathbf{u}_k)\}$ is strictly monotone decreasing and bounded from below, which proves the existence of a limiting value $\bar{\Phi}_p$. In principle, this does not imply that the sequence $\{\mathbf{u}_k\}$ has a limiting value. In order to show that $\{\mathbf{u}_k\}$ also has a limiting value, we consider two successive points of the sequence $\{\mathbf{u}_k\}$ upon convergence of the sequence $\{\Phi_p(\mathbf{u}_k)\}$, namely, $\bar{\Phi}_p = \Phi_p(\mathbf{u}_{k+1}) = \Phi_p(\mathbf{u}_k)$. Then, $\mathbf{u}_{k+1} \notin \mathcal{S}'_k$, which implies that $\mathcal{S}'_k = \emptyset$ and therefore $\mathcal{S}_k = \mathbf{u}_k = \mathbf{u}_{k+1} = \bar{\mathbf{u}}$. Hence, the algorithm converges to a fixed point $\bar{\mathbf{u}}$ and, by Lemma 4, $\bar{\mathbf{u}}$ is a KKT point of the plant, that is, $\bar{\mathbf{u}} = \bar{\mathbf{u}}_p$. This completes the proof. □

Remark 3 *(Local and global plant optimality).* Theorem 1 shows that the RTO algorithm (18) converges to a KKT point of the plant optimization problem (1) for any initial feasible input $\mathbf{u}_0 \in \mathcal{F}_p$. However, no guarantee can be given that a global optimizer for the plant will be reached. Indeed, the KKT point reached upon convergence may correspond to a local minimum of the plant, or even to a saddle point. Note that Lemma 4 states that any stationary point for the plant is a fixed point of the RTO algorithm. This means that, if the algorithm is started at a local maximum of the plant, it will stay there.

4. Modifier adaptation with feasible-side global convergence

Section 3 indicated that the use of convex upper-bounding functions provides properties that are useful to any RTO algorithm. We show next how these results can be used in the context of modifier adaptation to give rise to a new modifier-adaptation scheme that guarantees feasible-side global convergence as per Theorem 1. For this, the modified cost and constraint functions are augmented with quadratic terms so as to generate an RTO problem with convex upper-bounding functions.

4.1. Modifier adaptation with convex upper bounds

We propose to construct the following modified cost and constraint functions at the *k*th iteration,

$$\Phi_k^{\mathrm{U}}(\mathbf{u}) := \Phi(\mathbf{u}) + \varepsilon_k^{\Phi} + (\boldsymbol{\lambda}_k^{\Phi})^{\mathsf{T}}(\mathbf{u} - \mathbf{u}_k) + \frac{\delta^{\Phi}}{2}(\mathbf{u} - \mathbf{u}_k)^{\mathsf{T}}(\mathbf{u} - \mathbf{u}_k), \tag{24a}$$

$$G_{i,k}^{\mathrm{U}}(\mathbf{u}) := G_i(\mathbf{u}) + \varepsilon_k^{G_i} + (\boldsymbol{\lambda}_k^{G_i})^{\mathsf{T}}(\mathbf{u} - \mathbf{u}_k) + \frac{\delta^{G_i}}{2}(\mathbf{u} - \mathbf{u}_k)^{\mathsf{T}}(\mathbf{u} - \mathbf{u}_k), \quad i = 1, \ldots, n_g, \tag{24b}$$

where the constants δ^{Φ} and δ^{G_i}, $i=1, \ldots, n_g$, are selected such that Φ_k^{U} is a strictly convex upper-bounding function for Φ_p and $G_{i,k}^{\mathrm{U}}$ is a convex upper-bounding function for $G_{p,i}$ as per Definition 2. The next proposition shows how these quadratic upper bounds can be constructed.

Proposition 1 (Construction of upper bounds). *Let Assumptions 1 and 2 hold. Then, there exist non-negative scalars* α^{Φ}, β^{Φ} *and* α^{G_i}, β^{G_i}, *for* $i=1,\ldots,n_g$, *such that the cost and constraint functions* (24) *with*

$$\delta^{\Phi} \geq \max\{\alpha^{\Phi}, \beta^{\Phi}\}, \tag{25a}$$

$$\delta^{G_i} \geq \max\{\alpha^{G_i}, \beta^{G_i}\}, \qquad \text{for} \quad i=1,\ldots,n_g, \tag{25b}$$

are (strictly) convex upper-bounding functions of Φ_p *and* $G_{p,i}$ *in Problem* (1).

Proof. Consider the following choices of α^{Φ} and β^{Φ}

$$\alpha^{\Phi} \geq \max\left\{\epsilon, -\underline{\mu}^{\Phi} + \epsilon\right\},$$

$$\text{with} \quad \underline{\mu}^{\Phi} := \min_{\mathbf{u}\in\mathcal{U}} \min_{\mu} \left\{\mu \in \sigma\left(\frac{\partial^2 \Phi}{\partial \mathbf{u}^2}(\mathbf{u})\right)\right\},$$

$$\beta^{\Phi} \geq \max_{\mathbf{u}\in\mathcal{U}} \max_{\mu} \left\{\mu \in \sigma\left(\frac{\partial^2 \Phi_p}{\partial \mathbf{u}^2}(\mathbf{u}) - \frac{\partial^2 \Phi}{\partial \mathbf{u}^2}(\mathbf{u})\right)\right\},$$

and, for $i=1,\ldots,n_g$, consider the following choices of α^{G_i} and β^{G_i}

$$\alpha^{G_i} \geq \max\left\{0, -\underline{\mu}^{G_i}\right\},$$

$$\text{with} \quad \underline{\mu}^{G_i} := \min_{\mathbf{u}\in\mathcal{U}} \min_{\mu} \left\{\mu \in \sigma\left(\frac{\partial^2 G_i}{\partial \mathbf{u}^2}(\mathbf{u}, \boldsymbol{\theta})\right)\right\},$$

$$\beta^{G_i} \geq \max_{\mathbf{u}\in\mathcal{U}} \max_{\mu} \left\{\mu \in \sigma\left(\frac{\partial^2 G_{p,i}}{\partial \mathbf{u}^2}(\mathbf{u}) - \frac{\partial^2 G_i}{\partial \mathbf{u}^2}(\mathbf{u}, \boldsymbol{\theta})\right)\right\},$$

whereby $\sigma(\cdot)$ denotes the spectrum of a matrix. Without loss of generality, let ϵ be a small positive scalar and consider a choice of δ^{Φ} and δ^{G_i} satisfying (25). Evaluating $\Phi_k^{\mathrm{U}}(\mathbf{u})$ and $G_{i,k}^{\mathrm{U}}(\mathbf{u})$ at $\mathbf{u}_k$ gives

$$\Phi_k^{\mathrm{U}}(\mathbf{u}_k) = \Phi(\mathbf{u}_k) + \varepsilon_k^{\Phi} = \Phi_p(\mathbf{u}_k),$$

$$G_{i,k}^{\mathrm{U}}(\mathbf{u}_k) = G_i(\mathbf{u}_k) + \varepsilon_k^{G_i} = G_{p,i}(\mathbf{u}_k),$$

which satisfies Condition (17a) for the cost and the constraints, respectively. Computing the gradients of $\Phi_k^{\mathrm{U}}(\mathbf{u})$ and $G_{i,k}^{\mathrm{U}}(\mathbf{u})$ leads to

$$\frac{\partial \Phi_k^{\mathrm{U}}}{\partial \mathbf{u}}(\mathbf{u}) = \frac{\partial \Phi}{\partial \mathbf{u}}(\mathbf{u}) + (\boldsymbol{\lambda}_k^{\Phi})^{\mathsf{T}} + \delta^{\Phi}(\mathbf{u} - \mathbf{u}_k)^{\mathsf{T}},$$

$$\frac{\partial G_{i,k}^{\mathrm{U}}}{\partial \mathbf{u}}(\mathbf{u}) = \frac{\partial G_i}{\partial \mathbf{u}}(\mathbf{u}) + (\boldsymbol{\lambda}_k^{G_i})^{\mathsf{T}} + \delta^{G_i}(\mathbf{u} - \mathbf{u}_k)^{\mathsf{T}},$$

which, when evaluated at $\mathbf{u}_k$ gives

$$\frac{\partial \Phi_k^{\mathrm{U}}}{\partial \mathbf{u}}(\mathbf{u}_k) = \frac{\partial \Phi}{\partial \mathbf{u}}(\mathbf{u}_k) + \frac{\partial \Phi_p}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial \Phi}{\partial \mathbf{u}}(\mathbf{u}_k) = \frac{\partial \Phi_p}{\partial \mathbf{u}}(\mathbf{u}_k),$$

$$\frac{\partial G_{i,k}^{\mathrm{U}}}{\partial \mathbf{u}}(\mathbf{u}_k) = \frac{\partial G_i}{\partial \mathbf{u}}(\mathbf{u}_k) + \frac{\partial G_{p,i}}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial G_i}{\partial \mathbf{u}}(\mathbf{u}_k) = \frac{\partial G_{p,i}}{\partial \mathbf{u}}(\mathbf{u}_k),$$

which satisfies Condition (17b) for the cost and the constraints, respectively.

There remains to show that, with the selected choices of δ^{Φ} and δ^{G_i}, Φ_k^{U} is a strictly convex upper-bounding function for Φ_p, and $G_{i,k}^{\mathrm{U}}$ is a convex upper-bounding function for $G_{p,i}$. Note that $\delta^{\Phi} \geq \alpha^{\Phi}$ guarantees that Φ_k^{U} is a strictly convex function, a feature that has also been used in the augmented ISOPE method [5]. Here, the inclusion of ϵ guarantees that Φ_k^{U} is strictly convex also for $\delta^{\Phi} = \alpha^{\Phi}$. Similarly, $\delta^{G_i} \geq \alpha^{G_i}$ guarantees that $G_{i,k}^{\mathrm{U}}$ is a convex function. The proof that $\delta^{G_i} \geq \beta^{G_i}$ guarantees that $G_{i,k}^{\mathrm{U}}$ is an upper-bounding function for $G_{p,i}$ can be found in [20], and the same applies to the cost function. Hence, the upper-bounding functions (24) satisfy the conditions of Definition 2. □

We are now ready to summarize the properties of the proposed modifier-adaptation scheme in the following theorem.

Theorem 2 (Feasible-side global convergence). *Let Assumptions 1-3 hold, and let* δ^{Φ} *and* δ^{G_i} *be chosen according to Proposition 1. Then, for any initially feasible point* $\mathbf{u}_0 \in \mathcal{F}_p$, *the modifier-adaptation scheme using the cost and constraints* (24) *has the following properties:*

a) *All RTO iterates satisfy the plant constraints.*
b) *The RTO iterates converge to a KKT point of the plant.*
c) *The plant cost decreases monotonically at each iteration.*

Proof. The proof follows from observing that the choice of δ^{Φ} and δ^{G_i} according to Proposition 1 implies that the modifier-adaptation scheme using (24) satisfies all the conditions of Theorem 1. □

Remark 4 *(Convexification and feasibility)*. The proposed RTO scheme unifies the convexification of modifier adaptation presented in [15] with the second-order approach presented in [13] and the SCFO approach [7].

Remark 5 *(Adaptation of upper bounds)*. In practice, one usually does not know values of δ^{Φ} and δ^{G_i} that satisfy Proposition 1. At the same time, the Hessian upper bound parameters δ^{Φ} and δ^{G_i} influence the RTO step size significantly. One remedy to this problem could be to apply a trust-region-inspired adaptation of δ^{Φ} and δ^{G_i}. First results in this direction can be found in [27], where such an adaptation is proposed in the context of enforcing plant feasibility via quadratic upper bounds.

5. Case study: Williams–Otto reactor

To illustrate our findings, we consider the Williams–Otto reactor. The reactor is an ideal continuous stirred tank reactor, in which three reactions take place [31]:

$$\begin{aligned} A + B &\longrightarrow C \\ C + B &\longrightarrow P + E \\ C + P &\longrightarrow G\,. \end{aligned}$$

The reactants A and B are fed in the reactor with the mass flowrates F_A and F_B, respectively. The outlet stream has the mass flowrate $F = F_A + F_B$, the reactor temperature is T_R, and the reactor mass holdup is constant. The objective is to maximize the steady-state profit, which is expressed as the difference in price between the products P and E and the reactants:

$$J = 1143.38 X_P F + 25.92 X_E F - 76.23 F_A - 114.34 F_B,$$

where X_i represents the concentration of species i.

Since it is assumed that the reaction scheme is not well understood, the following two-reaction *model* has been proposed [14]:

$$\begin{aligned} A + 2B &\longrightarrow P + E \\ A + B + P &\longrightarrow G\,. \end{aligned}$$

The model equations and parameter values for both the plant and the model can be found in [32]. We shall consider the following optimization problem:

$$\max_{F_A, F_B, T_R} \quad J(\mathbf{u}) \tag{26a}$$

$$\text{s.t.} \quad F_A \in [3, \;\; 4.5]; \qquad F_B \in [6, \;\; 11]; \tag{26b}$$

$$\begin{aligned} &T_R \in [80, \;\; 105], \\ &g = X_G - 0.08 \leq 0, \end{aligned} \tag{26c}$$

Table 1
Plant optimal values.

F_A	F_B	T_R	X_G	Profit
3.887	9.369	91.2	**0.08**	210.33

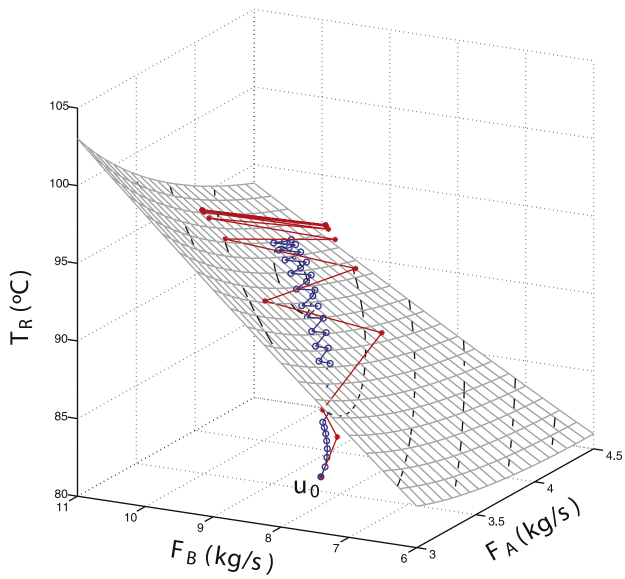

Fig. 1. Standard modifier adaptation applied to the Williams–Otto reactor. Input trajectories for $K = 0.8$ (-*-), and $K = 0.2$ (-o-).

where the decision variables are the flowrates of both reactants and the reactor temperature, $\mathbf{u} = [F_A, F_B, T_R]^\mathsf{T}$. The values of the decision and constrained variables at the plant optimum are given in Table 1, where it is seen that the constraint on the concentration of the undesired product G is active, and the maximum profit is $J = 210.33$.

5.1. Standard modifier adaptation

The standard modifier-adaptation algorithm (14)–(16) is applied first, starting from the initial point $\mathbf{u}_0 = [3.6, 10, 85]^\mathsf{T}$. Here, $\mathbf{u}_0$ was selected as a conservative initial *feasible* point, assuming that the plant is being operated at this point before the implementation of RTO. The first 40 iterations obtained using two different filter gains, $K = 0.8$ and $K = 0.2$, are shown in Fig. 1. This figure also shows the manifold corresponding to the active constraint for the plant ($X_G = 0.08$) in the three-dimensional input space. The constraint is violated above the manifold and respected on and below the manifold. The contour lines of the plant profit are plotted on the manifold. The evolution of the constrained concentration and the profit of the plant are shown in Fig. 2 for the case of $K = 0.2$. It is seen that the RTO iterates start oscillating on the infeasible side of the constraint on X_G, and the algorithm does not converge. The algorithm does not become stable by further decreasing the value of the filter gain. In fact, it turns out that, at the plant optimum, the reduced Hessian of the Lagrangian function predicted by the model is not positive semi-definite. This means that, for this optimization problem, the model based on the two-reaction scheme does not satisfy the model adequacy condition defined in [17].

5.2. Modifier adaptation augmented with quadratic terms

Next, the modifier-adaptation algorithm that includes the additional quadratic terms in the cost and constraint functions is applied. To be able to guarantee feasible-side global convergence,

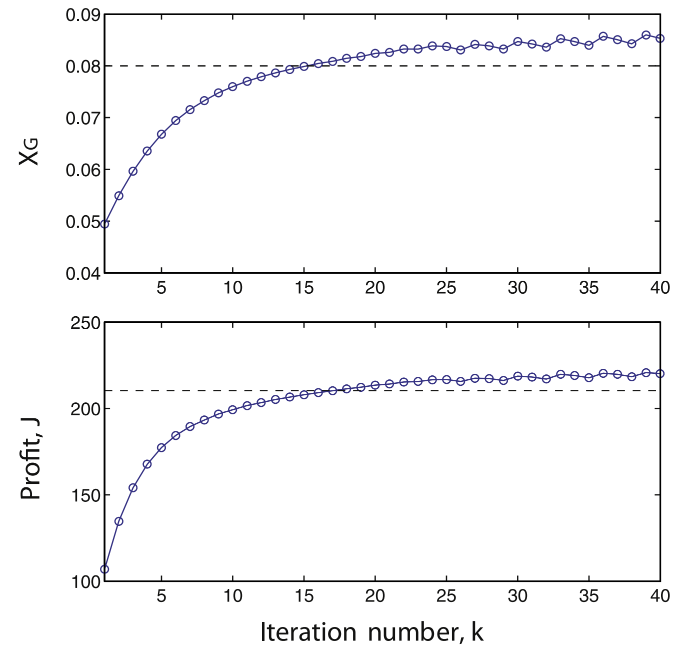

Fig. 2. Standard modifier adaptation applied to the Williams–Otto reactor with $K = 0.2$. Evolution of the constrained concentration X_G and the profit J for the plant. Dashed line: constraint bound (top plot); optimal profit (bottom plot).

Table 2
Estimated parameter values needed in Theorem 2.

$\underline{\mu}^\Phi = -20.12$	$\underline{\mu}^G = -0.0563$
$\underline{\alpha}^\Phi = 20.12$	$\underline{\alpha}^G = 0.0563$
$\beta^\Phi = 74.38$	$\beta^G = 0.0257$

Table 3
Simulation scenarios.

Scenario A:	$\delta^\Phi = 74.38$	$\delta^G = 0.0563$
Scenario B:	$\delta^\Phi = 0$	$\delta^G = 0.01$

we need to find values of the parameters δ^Φ and δ^G that satisfy the conditions in Theorem 2. An approximation of the parameters $\underline{\mu}$, α and β is obtained by evaluating the Hessian matrices $\frac{\partial^2 \Phi}{\partial \mathbf{u}^2}$, $\frac{\partial^2 \Phi_p}{\partial \mathbf{u}^2}$, $\frac{\partial^2 G_i}{\partial \mathbf{u}^2}$ and $\frac{\partial^2 G_{p,i}}{\partial \mathbf{u}^2}$ on a uniform grid of dimension $20 \times 20 \times 20$ in the feasible input space. The resulting parameter values are given in Table 2.

Two simulation scenarios are considered, which differ in the selected values of δ^Φ and δ^G. Scenario A in Table 3 uses the estimated values of δ^Φ and δ^G that satisfy the conditions in Theorem 2, while Scenario B uses much less conservative values. Starting from $\mathbf{u}_0$, the RTO iterates converge to the plant optimum without constraint violations in both scenarios. Fig. 3 shows how the iterates converge to the plant optimum after 150 iterations in Scenario B. Fig. 4 shows the evolution of the constrained concentration X_G and the objective function for the first 40 iterations. In both scenarios, the active constraint is first approached quickly without constraint violations, and then the iterates start moving slowly towards the plant optimum on the active constraint manifold. The values of δ^Φ and δ^G used in Scenario A are conservative, but they *guarantee* feasible-side global convergence starting from any feasible point. On the other hand, the value of $\delta^G = 0.01$ used in Scenario B is sufficient for satisfying the model adequacy condition, and it allows converging to the plant optimum without violating the constraint, but this comes without guarantees.

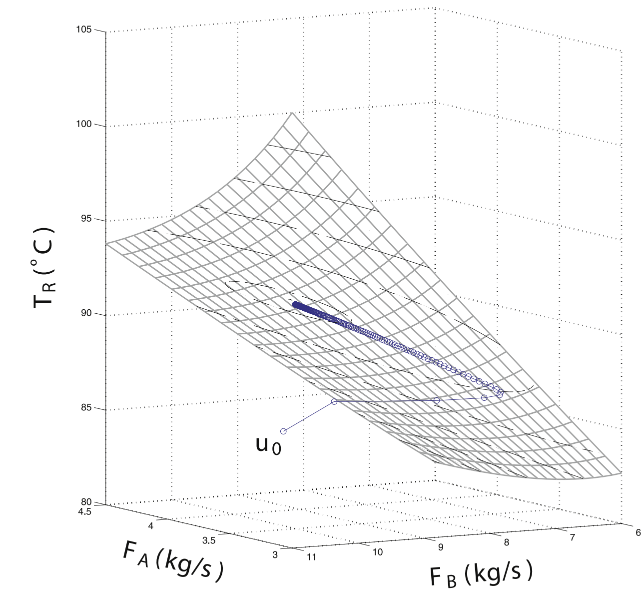

Fig. 3. Modifier adaptation augmented with quadratic terms, applied to the Williams–Otto reactor. Input trajectory for Scenario B.

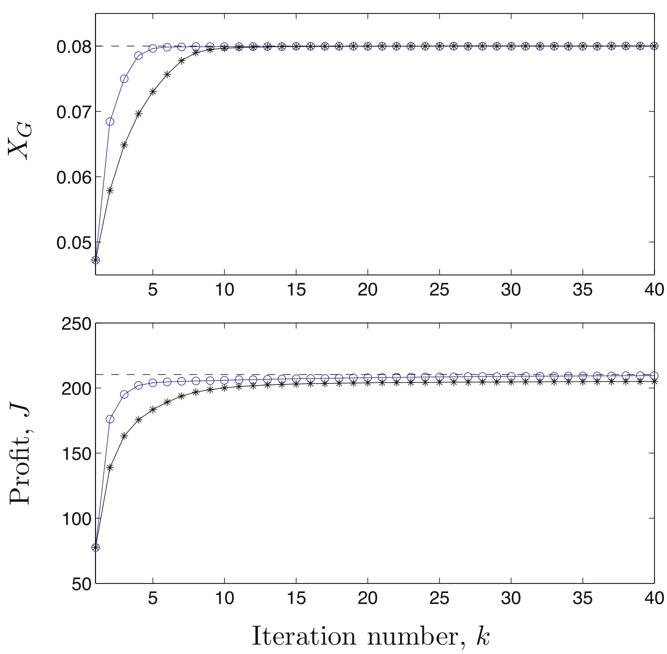

Fig. 4. Modifier adaptation augmented with quadratic terms, applied to the Williams–Otto reactor. Evolution of the plant concentration X_G and the plant profit J for Scenario A (-*-), and Scenario B (-o-). Dashed line: constraint bound (top plot); optimal profit (bottom plot).

5.3. Discussion

The case study shows that the values of δ^{Φ} and δ^{G_i} given by Proposition 1 may be conservative, resulting in slow convergence to the plant optimum. Conservatism is the price to pay for having global guarantees. In practice, feasible-side convergence can be enforced if the convex upper-bounding functions are valid in the region of the input space where the adaptation takes place. This calls for the use of an *adaptive* scheme to compute less conservative values of δ^{Φ} and δ^{G_i} based on the data collected during the iterative process. A first step in this direction is given in [27] for the case of quadratic surrogate models.

The Williams–Otto reactor example also shows that, if a constraint becomes (nearly) active before reaching the plant optimum, the iterative scheme may become extremely slow as the iterates start sliding on the constraint toward the plant optimum. In practice, this can prevent reaching the plant optimum. Nevertheless, significant improvement can be made by simply approaching the active constraints, which is done without constraint violations. For instance, consider the improvement made in the first 10 iterations in Fig. 4.

6. Conclusions

This paper has proposed a modifier-adaptation formulation that guarantees feasible-side global convergence to a KKT point of the plant. The approach is based on sequential convex inner-approximation methods for solving NLPs. We focused on *nominal* feasible-side global convergence guarantees that are valid in the *ideal* situation where the constraint values and the cost and constraint gradients of the plant are perfectly known at each RTO iteration. In practice, the Hessian upper bound parameters δ^{Φ} and δ^{G_i} that satisfy Proposition 1 are unknown. However, from a practical point of view, the useful implication of our result is that feasible-side convergence of modifier adaptation can be enforced by adding quadratic terms to the modified cost and constraint functions and increasing the values of δ^{Φ} and δ^{G_i}. In the presence of measurement noise, only estimates of the plant gradients can be obtained, and the guarantees presented in this paper do not apply. However, understanding how feasibility and convergence can be enforced in the nominal case is essential for studying more realistic scenarios that depart from the nominal case.

Future work will investigate two different directions to enforce the applicability of the proposed scheme: How to estimate suitable values for the Hessian upper bound parameters δ^{Φ} and δ^{G_i} from available process measurements, and how to adapt the parameters δ^{Φ} and δ^{G_i} by means of trust-region-like concepts.

References

[1] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, Nonlinear Programming: Theory and Algorithms, 3rd ed., John Wiley and Sons, New Jersey, 2006.
[2] A. Beck, A. Ben-Tal, L. Tetruashvili, A sequential parametric convex approximation method with applications to non convex truss topology design problems, J. Glob. Optim. 47 (2010) 29–51.
[3] D.P. Bertsekas, Nonlinear Programming, 2nd ed., Athena Scientific, 1999.
[4] L.T. Biegler, I.E. Grossmann, A.W. Westerberg, A note on approximation techniques used for process optimization, Comp. Chem. Eng. 9 (2) (1985) 201–206.
[5] M. Brdyś, J.E. Ellis, P.D. Robert, Augmented integrated system optimization and parameter estimation technique: derivation, optimality and convergence, in: Control Theory and Applications, IEE Proceedings D, vol. 134(3), 1987, pp. 201–209.
[6] G.A. Bunin, On the equivalence between the modifier-adaptation and trust-region frameworks, Comp. Chem. Eng. 71 (2014) 154–157.
[7] G.A. Bunin, G. François, D. Bonvin, Sufficient conditions for feasibility and optimality of real-time optimization schemes – I. Theoretical foundations, 2013, ArXiv:1308.2620.
[8] G.A. Bunin, Z. Wuillemin, G. François, A. Nakajo, L. Tsikonis, D. Bonvin, Experimental real-time optimization of a solid-oxide fuel cell stack via constraint adaptation, Energy 39 (1) (2012) 54–62.
[9] C.Y. Chen, B. Joseph, On-line optimization using a two-phase approach: an application study, Ind. Eng. Chem. Res. 26 (1987) 1924–1930.
[10] S. Costello, G. François, D. Bonvin, A directional modifier-adaptation algorithm for real-time optimization, J. Process Control 39 (2016) 64–76.
[11] M.L. Darby, M. Nikolaou, J. Jones, D. Nicholson, RTO: an overview and assessment of current practice, J. Process Control 21 (2011) 874–884.
[12] J.E. Ellis, C. Kambhampati, G. Sheng, P.D. Roberts, Approaches to the optimizing control problem, Int. J. Syst. Sci. 19 (10) (1988) 1969–1985.
[13] T. Faulwasser, D. Bonvin, On the use of second-order modifiers for real-time optimization, in: 19th IFAC World Congress, Cape Town, 2014.
[14] J.F. Forbes, T.E. Marlin, J.F. MacGregor, Model adequacy requirements for optimizing plant operations, Comp. Chem. Eng. 18 (6) (1994) 497–510.
[15] G. François, D. Bonvin, Use of convex model approximations for real-time optimization via modifier adaptation, Ind. Eng. Chem. Res. 52 (2013) 11614–11625.

[16] W. Gao, S. Engell, Iterative set-point optimization of batch chromatography, Comp. Chem. Eng. 29 (2005) 1401–1409.
[17] A. Marchetti, B. Chachuat, D. Bonvin, Modifier-adaptation methodology for real-time optimization, Ind. Eng. Chem. Res. 48 (13) (2009) 6022–6033.
[18] A. Marchetti, B. Chachuat, D. Bonvin, A dual modifier-adaptation approach for real-time optimization, J. Process Control 20 (2010) 1027–1037.
[19] A.G. Marchetti, A new dual modifier-adaptation approach for iterative process optimization with inaccurate models, Comp. Chem. Eng. 59 (2013) 89–100.
[20] A.G. Marchetti, M. Singhal, T. Faulwasser, D. Bonvin, Modifier adaptation with guaranteed feasibility in the presence of gradient uncertainty, Comp. Chem. Eng. 98 (2017) 61–69.
[21] B.R. Marks, G.P. Wright, A general inner approximation algorithm for nonconvex mathematical programs, Oper. Res. 26 (4) (1978) 681–683.
[22] D. Navia, G. Gutiérrez, C. de Prada, Nested modifier-adaptation for RTO in the Otto Williams reactor, in: 10th IFAC International Symposium on Dynamics and Control of Process Systems, vol. 10, Mumbai, India, 2013, pp. 123–128.
[23] P.D. Roberts, An algorithm for steady-state system optimization and parameter estimation, J. Syst. Sci. 10 (1979) 719–734.
[24] P.D. Roberts, T.W. Williams, On an algorithm for combined system optimisation and parameter estimation, Automatica 17 (1) (1981) 199–209.
[25] E.A. Rodger, B. Chachuat, Design methodology of modifier adaptation for on-line optimization of uncertain processes, in: IFAC World Congress, vol. 18, Milano, Italy, 2011, pp. 4113–4118.
[26] M. Singhal, T. Faulwasser, D. Bonvin, On handling cost gradient uncertainty in real-time optimization, in: 9th International Symposium on Advanced Control of Chemical Processes, Whistler, BC, Canada, 2015.
[27] M. Singhal, A.G. Marchetti, T. Faulwasser, D. Bonvin, Real-time optimization based on adaptation of surrogate models, in: 11th IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems, Trondheim, Norway, 2016.
[28] B. Srinivasan, S. Palanki, D. Bonvin, Dynamic optimization of batch processes: I. Characterization of the nominal solution, Comp. Chem. Eng. 27 (2003) 1–26.
[29] K. Svanberg, A class of globally convergent optimization methods based on conservative convex separable approximations, SIAM J. Optim. 12 (2) (2002) 555–573.
[30] P. Tatjewski, Iterative optimizing set-point control – the basic principle redesigned, in: 15th IFAC World Congress. Barcelona, Spain, 2002.
[31] T.J. Williams, R.E. Otto, A generalized chemical processing model for the investigation of computer control, AIEE Trans. 79 (1960) 458.
[32] Y. Zhang, J.F. Forbes, Extended design cost: a performance criterion for real-time optimization systems, Comp. Chem. Eng. 24 (2000) 1829–1841.