Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-7 (2016) 412–417

Real-Time Optimization Based on Adaptation of Surrogate Models

Martand Singhal *, **Alejandro G. Marchetti** *, **Timm Faulwasser** *,** **and Dominique Bonvin** *

* *Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*
e-mail: {martand.singhal, alejandro.marchetti, timm.faulwasser, dominique.bonvin}@epfl.ch
** *Institute for Applied Informatics, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany*

Abstract: Recently, different real-time optimization (RTO) schemes that guarantee feasibility of all RTO iterates and monotonic convergence to the optimal plant operating point have been proposed. However, simulations reveal that these schemes converge very slowly to the plant optimum, which may be prohibitive in applications. This note proposes an RTO scheme based on second-order surrogate models of the objective and the constraints, which enforces feasibility of all RTO iterates, i.e., plant constraints are satisfied at all iterations. In order to speed up convergence, we suggest an online adaptation strategy of the surrogate models that is based on trust-region ideas. The efficacy of the proposed RTO scheme is demonstrated in simulations via both a numerical example and the steady-state optimization of the Williams-Otto reactor.

Keywords: real-time optimization, plant-model mismatch, plant feasibility, online adaptation

1. INTRODUCTION

To ensure profitability in the process industries, one typically maximizes the economic performance, while respecting safety and environmental constraints. Hence, maximizing plant performance in real time, known as real-time optimization (RTO), has gathered commendable industrial attention. The goal of RTO is to enforce plant optimality in the presence of uncertainty such as plant-model mismatch and process disturbances. RTO methods rely on the available (not necessarily very accurate) plant model and measurements to push the *plant* towards optimality.

In the design of RTO schemes, it is desirable to achieve the following properties: (i) plant optimality and feasibility upon convergence, (ii) acceptable number of RTO iterations, and (iii) plant feasibility throughout the optimization process. The works of Gao and Engell (2005), Chachuat et al. (2009), Marchetti et al. (2009, 2010) are tailored to (i). However, the proposed schemes depend on tuning parameters to enforce (ii) and cannot guarantee (iii). Furthermore, it has been proposed to estimate gradients via surrogate models (Bunin et al., 2013a; Gao et al., 2015).

Recently, Bunin et al. (2013b) suggested to rely on Lipschitz constants of the constraints and a Hessian upper bound of the cost to enforce plant feasibility and monotonic convergence in some kind of post-processing of RTO iterates. Furthermore, it has been shown in Singhal et al. (2015) that similar ideas can also be used to design data-driven RTO schemes based on linear-quadratic surrogate models, where the plant constraints are approximated in a linear-affine fashion and the plant objective is modeled as a quadratic function. Both Bunin et al. (2013b) and Singhal et al. (2015) guarantee plant feasibility of all RTO iterates, that is, they enforce (iii). However, simulations have shown that the (conservative) first-order approximation of plant constraints using Lipschitz constants often leads to slow convergence of the RTO algorithm, which may be prohibitive in applications.

The present paper investigates data-driven RTO based on surrogate models. We propose an RTO scheme based on quadratically-constrained quadratic programs (QCQP), which can be regarded as an extension of the scheme proposed in Singhal et al. (2015) based on quadratic programs (QP). The contributions of the present work are as follows: We sketch the feasibility and optimality properties of the proposed RTO scheme. Furthermore, we analyze why first-order constraint approximations based on Lipschitz constants may lead to overly slow convergence. Finally, we present a scheme for adapting second-order surrogate models and illustrate via simulation that this leads to considerably faster RTO convergence.

The paper is structured as follows. Section 2 briefly formulates the RTO problem. Section 3 introduces the QCQP-based surrogate model and draws a comparison between the QCQP and the QP surrogate models. The performance of the QCQP-based surrogate model is tested on a numerical example. Section 4 introduces the online adaptation algorithm for the upper bounding term and presents the gradient-estimation method. Section 5 presents the Williams-Otto reactor case study, illustrating the performance of the proposed RTO scheme. Finally, the conclusions are presented in section 6.

10.1016/j.ifacol.2016.07.377

2. PLANT OPTIMIZATION PROBLEM

Steady-state performance improvement can be formulated mathematically as a nonlinear program

$$\min_{u} \phi(u,d) \tag{1a}$$
$$\text{s.t.} \quad g_j(u,d) \leq 0, \qquad j{=}1,..,n_g \tag{1b}$$
$$u^L \preceq u \preceq u^U, \tag{1c}$$

where u is the n_u-dimensional input vector, d is the n_d-dimensional vector of disturbances, $\phi : \mathbb{R}^{n_u} \times R^{n_d} \to \mathbb{R}$ is the plant cost, $g_j : \mathbb{R}^{n_u} \times R^{n_d} \to \mathbb{R}$ is the j^{th} plant constraint. The symbol ($\prec$) denotes componentwise inequality of vectors. The disturbance d models the fact that the plant cost and constraints are affected by disturbances. In this paper, however, we do not deal with the disturbance d explicitly. Instead, we assume that the plant cost ϕ and the plant constraints g_j are not exactly known.

3. RTO BASED ON SURROGATE MODELS

Next, we focus on RTO based on surrogate models to find the plant optimum. The surrogate models will enforce plant feasibility and guarantee monotonic plant cost decrease provided that the available measurements are noise free. To this end, we recall two technical lemmas presented in Bunin et al. (2013b).

Lemma 1. (Lipschitz upper bound). Let $f : \mathbb{R}^{n_u} \to \mathbb{R}$ be continuously differentiable over the compact set $\mathcal{U} \subset \mathbb{R}^{n_u}$ such that

$$-\lambda_i < \left.\frac{\partial f}{\partial u_i}\right|_u < \lambda_i, \quad \forall u \in \mathcal{U}, \quad i = 1, ..., n_u, \tag{2}$$

where λ are the univariate Lipschitz constants of f. Then, the evolution of f between the two successive inputs u_k and u_{k+1} is bounded by

$$f(u_{k+1}) \leq f(u_k) + \sum_{i=1}^{n_u} \lambda_i |u_{k+1,i} - u_{k,i}|. \tag{3}$$

$\square$

Lemma 2. (Hessian upper bound). Let $f : \mathbb{R}^{n_u} \to \mathbb{R}$ be twice continuously differentiable over the compact set $\mathcal{U} \subset \mathbb{R}^{n_u}$ such that

$$-M_{ij} < \left.\frac{\partial^2 f}{\partial u_i \partial u_j}\right|_u < M_{ij}, \quad \forall u \in \mathcal{U}, \quad i,j = 1, ..., n_u.$$

Let $\Delta_{k+1} := u_{k+1} - u_k$. Then, the change in f between u_k and u_{k+1} can be bounded as

$$f(u_{k+1}) \leq f(u_k) + \nabla f(u_k)^T \Delta_{k+1} + \frac{1}{2} \Delta_{k+1}^T \bar{Q} \Delta_{k+1}, \tag{4}$$

where $\bar{Q} \succ 0$ is a diagonal matrix with the diagonal elements $\bar{Q}_{ii} = \sum_{j=1}^{n_u} M_{ij},\ i = 1, ..., n_u$. $\square$

While the first lemma provides a first-order affine upper bound for any continuously differentiable function over any compact set, the second lemma gives a quadratic approximation that upper bounds twice continuously differentiable functions. Note that, in Lemma 2, the matrix $\bar{Q}$ can be seen as a Hessian upper bounding matrix for the function f over the compact set $\mathcal{U}$. We refer to Bunin et al. (2013b) for the proofs of these two lemmas.

3.1 RTO with QP surrogate model

Based on these two lemmas, the following QP-based surrogate model is proposed in Singhal et al. (2015)

$$\min_{\Delta_{k+1}} \quad \nabla \phi_k^T \Delta_{k+1} + \frac{1}{2} \Delta_{k+1}^T \bar{Q} \Delta_{k+1} \tag{5a}$$

subject to

$$g_j(u_k) + \sum_{i=1}^{n_u} \lambda_{i,j} |\Delta_{k+1,i}| \leq 0, \quad j = 1, ..., n_g, \tag{5b}$$
$$\nabla g_{j,k}^T \Delta_{k+1} \leq -\delta_j, \ \forall j \in \mathcal{J}_k, \tag{5c}$$
$$u^L - u_k \preceq \Delta_{k+1} \preceq u^U - u_k, \tag{5d}$$

where $\nabla \phi_k$ is the gradient of the plant cost function ϕ at u_k and $\nabla g_{j,k}$ is the gradient of the plant constraint g_j at u_k. The positive definite matrix $\bar{Q}$ is the Hessian upper bounding matrix for the cost function (computed as in Lemma 2 for $f := \phi$), $\lambda_{i,j}$ are the univariate Lipschitz constants of the plant constraint g_j. The set $\mathcal{J}_k$ is the set of ϵ-active constraints defined as (Bunin et al., 2013b)

$$\mathcal{J}_k = \{ j \in \{1, \ldots, n_g\} : -\epsilon_j \leq g_j(u_k) \leq 0 \},$$

where ϵ is a small positive scalar. The next RTO iterate u_{k+1} is given by

$$u_{k+1} = u_k + \Delta_{k+1}^\star, \tag{6}$$

where $\Delta_{k+1}^\star$ is the solution to Problem (5). The constraints (5b) ensure that u_{k+1} is a feasible point for the plant.

Singhal et al. (2015) showed that the RTO scheme based on (5)-(6) guarantees plant feasibility at all RTO iterations. However, this scheme may converge very slowly because of the conservatism induced by both the Hessian upper bound $\bar{Q}$ and the univariate Lipschitz constants $\lambda_{i,j}$. Furthermore, note that the surrogate model does not use information on the constraint gradient for ensuring feasibility in (5b), but only for keeping a distance from the active constraints in (5c), which may also contribute to the slow convergence.

3.2 RTO with QCQP surrogate model

In order to achieve faster convergence, we propose the following RTO scheme based on a QCQP surrogate model

$$\min_{\Delta_{k+1}} \nabla \phi_k^T \Delta_{k+1} + \frac{1}{2} \Delta_{k+1}^T \bar{Q} \Delta_{k+1} \tag{7a}$$

subject to

$$g_j(u_k) + \nabla g_{j,k}^T \Delta_{k+1} + \frac{1}{2} \Delta_{k+1}^T \bar{Q}_j \Delta_{k+1} \leq 0, \quad j = 1, ..., n_g, \tag{7b}$$
$$u^L - u_k \preceq \Delta_{k+1} \preceq u^U - u_k, \tag{7c}$$

where $\bar{Q}_j$ is the Hessian upper bounding matrix of the plant constraint function g_j. The input update is also given by (6) with $\Delta_{k+1}^\star$ being in this case the solution to Problem (7).

Remark 1. (Nominal plant feasibility).
The scheme (7) can be understood as an RTO-specific translation of the nominal optimization method suggested in (Auslender et al., 2010). Therein, it is shown that the QCQP formulation leads to recursively feasible iterates provided that the gradient information is exact and the following assumptions hold: