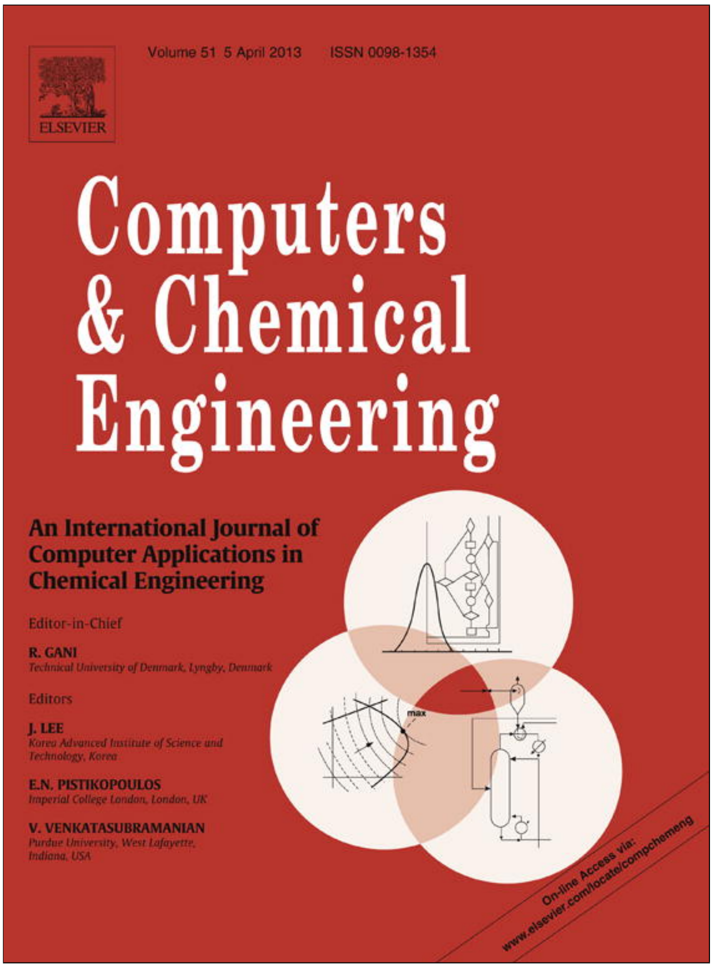
Volume 51 5 April 2013 ISSN 0098-1354
ELSEVIER
Computers & Chemical Engineering
An International Journal of Computer Applications in Chemical Engineering
Editor-in-Chief
R. GANI
Technical University of Denmark, Lyngby, Denmark
Editors
J. LEE
Korea Advanced Institute of Science and Technology, Korea
E.N. PISTIKOPOULOS
Imperial College London, London, UK
V. VENKATASUBRAMANIAN
Purdue University, West Lafayette, Indiana, USA
max
On-line Access via:
www.elsevier.com/locate/compchemeng

Computers and Chemical Engineering 51 (2013) 172–180

Contents lists available at SciVerse ScienceDirect

Computers and Chemical Engineering

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

On the role of the necessary conditions of optimality in structuring dynamic real-time optimization schemes

Dominique Bonvin[a,*], Bala Srinivasan[b]

[a] Laboratoire d'Automatique, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
[b] Department of Chemical Engineering, Ecole Polytechnique Montreal, Montreal, Canada

ARTICLE INFO

Article history:
Received 28 March 2012
Received in revised form 9 July 2012
Accepted 12 July 2012
Available online 2 August 2012

Keywords:
Dynamic real-time optimization
Plant-model mismatch
Modeling for optimization
Input parameterization
Plant model
Solution model

ABSTRACT

In dynamic optimization problems, the optimal input profiles are typically obtained using models that predict the system behavior. In practice, however, process models are often inaccurate, and on-line model adaptation is required for appropriate prediction and re-optimization. In most dynamic real-time optimization schemes, the available measurements are used to update the plant model, with uncertainty being lumped into selected uncertain plant parameters; furthermore, a piecewise-constant parameterization is used for the input profiles. This paper argues that the knowledge of the necessary conditions of optimality (NCO) can help devise more efficient and more robust real-time optimization schemes. Ideally, the structuring decisions involve the NCO as follows: (i) one measures or estimates the plant NCO, (ii) a NCO-based input parameterization is used, and (iii) model adaptation is performed to meet the plant NCO. The benefit of using the NCO in dynamic real-time optimization is illustrated in simulation through the comparison of various schemes for solving a final-time optimal control problem in the presence of uncertainty.

1. Introduction

Optimization is important in science and engineering as a way of finding the best solutions, designs or operating conditions. Optimization is typically performed on the basis of a mathematical model of the object of attention. For example, engineers might be interested in the optimal operation of processes that either operate at steady state or undergo transient changes. The object of attention, or reality, is called the "plant", whereas the "model" is a set of algebraic, differential or differential-algebraic equations.

In practice, optimization is complicated by the presence of uncertainty in the form of plant-model mismatch and unknown disturbances. Without uncertainty, one could use the model at hand, optimize it numerically off-line and implement the optimal inputs in an open-loop fashion. However, because of uncertainty, additional information such as uncertainty description or plant measurements must be included. In the former case, robust optimization computes a set of inputs that guarantees feasibility either for all possible realizations or with a desired probability level, however at the expense of a conservative solution (Srinivasan, Bonvin, Visser, & Palanki, 2003; Terwiesch, Agarwal, & Rippin, 1994). In the latter case, the inputs are updated in real-time based on measurements. This is the field of *real-time optimization*, which is labeled RTO for static optimization problems (Marlin & Hrymak, 1997) and DRTO for dynamic optimization problems (Biegler, 2009). This paper deals with two major implementation issues in DRTO, namely, model quality and computational aspects.

The issue of *model quality* raises an important question: Does good performance require a good model? This is not necessarily the case for control, since errors resulting from a poor model can be offset by the action of feedback. In optimization, without feedback to make up for modeling errors, the model needs to represent the reality accurately, in particular the optimality conditions of the plant. The situation is slightly different in real-time optimization since the measurements available on-line represent some form of feedback. However, this feedback is only partial as it is typically limited to output information. Furthermore, it is important to adapt the model appropriately, that is, there where it matters most for the purpose of optimization. These issues of measurement location and input update in the context of imperfect model are crucial for reaching optimality. It is argued in this paper that the necessary conditions of optimality (NCO) predicted by the model need to match those of the plant for plant optimality. We will discussed how the NCO measurements can be incorporated in a model so as to be most useful for optimization.

The *computational aspects* are also crucial for implementation. A very reliable optimization scheme is model predictive control (MPC), which incorporates state feedback, uses a receding horizon and carries out the optimization repeatedly at each sampling

* Corresponding author. Tel.: +41 21 693 3843; fax: +41 21 693 2574.
E-mail address: dominique.bonvin@epfl.ch (D. Bonvin).

0098-1354/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compchemeng.2012.07.012

time (Rawlings & Mayne, 2009). MPC was initially developed to track a reference trajectory by minimization of a quadratic error term. It has recently been extended to "economic MPC" that uses a non-quadratic cost function (Heidarinejad, Liu, & Christofides, 2012; Rawlings & Amrit, 2009). Furthermore, there has been considerable efforts in recent years to speed up the computations by formulating convex optimization problems and also using algorithms that exploit the structure of the problem (Diehl, Ferreau, & Haverbeke, 2009; Richter, Morari, & Jones, 2011; Wang & Boyd, 2010). On the other hand, recent trends in DRTO have included attempts to move the heavy computations off-line, where time and computational power are more available, and limit the on-line operations to quick decisions and easy computations. For example, multi-parametric programming generates off-line a lookup table of control laws, which are then used on-line based on the estimated states of the plant (Bemporad, Morari, Dua, & Pistikopoulos, 2002; Pistikopoulos, Georgiadis, & Dua, 2007; Zeilinger, Jones, & Morari, 2011). Also, "advanced step NMPC" strategies have been proposed, which solve the detailed optimization problem in background and apply sensitivity-based update on-line (D'Amato, Kumar, Lopez-Negrete, & Biegler, 2012; Zavala & Biegler, 2009). Another approach is the nonlinear real-time iteration scheme, which uses a continuation Newton-type framework and solves one QP at each iteration (Diehl, Bock, & Schlöder, 2005; Diehl et al., 2002). This allows for multiple active set changes and thus ensures that the nonlinear MPC algorithm cannot perform worse than a linear MPC controller. Yet a different approach is NCO tracking, which uses a NCO-based parameterization of the input profiles to design a multivariable feedback scheme that tracks the first-order optimality conditions, thereby pushing the system toward optimality (Srinivasan & Bonvin, 2007).

This paper deals with the model-quality issue in DRTO. In the presence of significant plant-model mismatch, the use of a fixed nominal model is typically insufficient to drive the *plant* to optimality. With MPC for example, the estimated states are often inaccurate, and one would need to update the model, which is difficult to do in closed-loop operation due to the so-called dual control problem (Aström & Wittenmark, 1995). This work adopts the viewpoint that, in real-time optimization, the model is a vehicle to process plant measurements and compute the optimal inputs. This step involves two major decisions, namely, the choice of the measured quantities and the choice of a finite number of decision variables via input parameterization. Note that these choices can benefit from knowledge of the NCO since the NCO are intimately linked to plant optimality. The structure of the optimal solution and the corresponding NCO can be determined off-line by numerical optimization. These measurement and input-parameterization issues are briefly addressed next.

Measurements and uncertainty description. The measurements are typically the plant outputs y_{plant}. The uncertainty, which is observed as the difference between the plant measurements and the corresponding model predictions, can be represented as parametric variations of the plant model. Alternatively, if the NCO elements can be measured, say y_{NCO}, the model uncertainty can be expressed as the difference between the measured and the predicted y_{NCO} values. It is interesting to notice the close relation between the type of measurements (the plant outputs y_{plant} vs. the NCO elements y_{NCO}) and the uncertainty description (the plant parameters θ_{plant} vs. the NCO deviations Δ_{NCO}).

Parameterization and update of the inputs. The traditional way of parameterizing infinite-dimensional inputs is control vector parameterization (CVP), whereby the inputs are approximated as piecewise-constant profiles. The main advantage is universality, that is, any solution can be closely approximated by introducing a sufficient number of pieces (barring certain numerical issues such as ringing around discontinuities). However, CVP typically contains

		Input Parameterization	
		π_{CVP}	π_{NCO}
Measurements	y_{plant}	$y_{plant} \xrightarrow{\text{Ident}} \theta_{plant} \xrightarrow{\text{Opt}} \pi_{CVP}$ (*two-step approach, CVP*)	$y_{plant} \xrightarrow{\text{Ident}} \theta_{plant} \xrightarrow{\text{Opt}} \pi_{NCO}$ (*two-step approach, NCO*)
	y_{NCO}	$y_{NCO} \xrightarrow{\text{Diff}} \Delta_{NCO} \xrightarrow{\text{Opt}} \pi_{CVP}$ (*modifier adaptation*)	$y_{NCO} \xrightarrow{\text{Diff}} \Delta_{NCO} \xrightarrow{\text{Control}} \pi_{NCO}$ (*NCO tracking*)

Fig. 1. Measurement and input-parameterization features of various DRTO schemes. The measured plant outputs are labeled y_{plant}, the measured NCO elements y_{NCO}; the inputs are parameterized via control vector parameterization, π_{CVP}, or via the elements of the NCO, π_{NCO}. Plant-model mismatch can be absorbed in the plant model parameters θ_{plant} or via an additive disturbance to the NCO values, Δ_{NCO}. "Ident" means the use of parameter identification, "Diff" the computation of a difference, "Opt" the use of numerical optimization, and "Control" the use of feedback control.

a large number of piecewise-constant input values, denoted here π_{CVP}. In contrast, a parsimonious input parameterization, π_{NCO}, can be obtained from the knowledge of the NCO, that is, the input elements correspond to switching times between arcs and input values associated with certain arcs. The way the inputs are parameterized impacts on the way there are updated. With π_{CVP}, the only efficient way to compute the inputs is through numerical optimization. With π_{NCO}, the few input parameters can be adjusted via feedback control to regulate the deviation Δ_{NCO} to zero.

Various DRTO schemes are possible based on the choice of the measurement and input-parameterization options, some of which are illustrated in Fig. 1 and discussed next.

- In the "two-step approach" of repeated parameter identification and performance optimization, the measurements are used to adapt the model parameters and estimate the current states. The estimated states serve as initial conditions for the optimization that is repeated on-line with the updated model (Chen & Joseph, 1987; Eaton & Rawlings, 1990). The input parameterization is of the CVP type.
- In the modifier-adaptation approach, modifier terms are added to the cost and constraint functions. Upon measurement of y_{NCO}, the modifiers are updated in order for the model and the plant to have matching first-order optimality conditions (Chachuat, Srinivasan, & Bonvin, 2009; Marchetti, Chachuat, & Bonvin, 2007). These schemes use the measurements y_{NCO} and the inputs π_{CVP} for optimization.
- It is also possible to perform numerical optimization using a NCO-based input parameterization. Such a scheme has been developed by Schlegel and Marquardt (2006a) and applied to an industrial polymerization process in Schlegel and Marquardt (2006b). The corresponding DRTO scheme consists in measuring y_{plant} and updating the model parameters accordingly, followed by numerical optimization using the π_{NCO} parameterization.
- Finally, NCO tracking uses the measurements of y_{NCO} to update π_{NCO} using feedback control to meet the plant NCO (Srinivasan & Bonvin, 2007).

This paper considers the implementation of optimal control in the presence of significant uncertainty in the form of plant-model mismatch, which requires some form of adaptation based on plant measurements. Three possible adaptation strategies are considered, namely adaptation of the process model, adaptation of the cost and constraint functions, and direct adaptation of the inputs. A similar investigation has already been proposed for the static optimization case (Chachuat et al., 2009). Note also that the emphasis in the present study is not on computational aspects as this has been the case in numerous recent investigations involving nonlinear MPC (Diehl et al., 2009), but rather on alternatives ways to compensate for plant-model mismatch and drive the *plant* to optimality.

The paper is organized as follows. Section 2 formulates the optimization problem and presents the corresponding NCO. Section 3 addresses numerical optimization using a plant model and touches upon the topics of "model adequacy" and "modeling for optimization". Optimizing control using a solution model is presented in Section 4. Section 5 discusses the choice of a DRTO scheme with regard to both measurements and input parameterization. The various ideas developed in this paper are illustrated through a simple dynamic example in Section 6, and Section 7 concludes the paper.

2. Dynamic optimization problem

2.1. Problem formulation

Consider the constrained dynamic optimization problems with finite operational time, where the objective is to determine the input profiles that optimize a final-time cost function. In addition to constraints corresponding to the dynamic system equations, there might be path constraints (involving inputs and states) as well as terminal constraints. Input bounds are dictated by actuator limitations, while state-dependent constraints typically result from safety and operability considerations. Terminal constraints usually arise from quality or performance considerations.

Since the inputs are infinite dimensional, numerical solution of the optimization problem requires input parameterization. Here the inputs are parameterized using a set of finite parameters π and numerical optimization then computes the optimal values of π.

The main challenge addressed in this paper is that of uncertainty, that is, the plant does not quite behave as predicted by the model. One way to incorporate the uncertainty in the model is via the set of uncertain parameters θ. The dynamic optimization problem can be formulated mathematically as follows (Bryson & Ho, 1975; Kirk, 1970):

$$\min_{\pi} \quad \phi(x(t_f), \rho, \theta) \tag{1}$$

$$\text{s.t.} \quad \dot{x}(t) = F(x(t), u(t), \rho, \theta), \quad x(0) = x_0(\rho, \theta) \tag{2}$$

$$S(x(t), u(t), \rho, \theta) \le 0 \tag{3}$$

$$T(x(t_f), \rho, \theta) \le 0 \tag{4}$$

$$[u(t), \rho] = \mathcal{U}(t, x, \pi, \theta), \tag{5}$$

where ϕ is the final-time cost functional to be minimized, x the states with the known initial conditions x_0, u the inputs, ρ the time-invariant decision variables, S the path constraints, T the terminal constraints, F the system dynamics, θ the uncertain plant model parameters, $\mathcal{U}$ the input parameterization, π the input parameters, and t_f the final time that is finite but can be either fixed or free. If t_f is free, it is part of ρ. Note that the initial conditions can also be considered as decision variables.

The solution of Problems (1)–(5) is typically discontinuous and consists of several intervals or arcs (Srinivasan, Palanki, & Bonvin, 2003). Yet, the input profiles are continuous and differentiable in each interval. Each interval is characterized by a different set of active path constraints, that is, this set changes between two successive intervals.

2.2. Necessary conditions of optimality

Let us define the following functions:

$$H(t) = \lambda^T(t) F(x(t), u(t), \rho, \theta) + \mu^T(t) S(x(t), u(t), \rho, \theta) \tag{6}$$

$$\Phi(t_f) = \phi(x(t_f), \rho, \theta) + \nu^T T(x(t_f), \rho, \theta) \tag{7}$$

$$\Psi(t_f) = \Phi(t_f) + \int_0^{t_f} H(t) dt \tag{8}$$

$$\dot{\lambda}^T(t) = -\frac{\partial H}{\partial x}(t), \quad \lambda^T(t_f) = \frac{\partial \Phi(t_f)}{\partial x(t_f)}, \tag{9}$$

where $H(t)$ is the Hamiltonian function, $\Phi(t_f)$ the augmented terminal cost, $\Psi(t_f)$ the total terminal cost, $\lambda(t) \neq 0$ the adjoint variables (Lagrange multipliers for the system equations), $\mu(t) \ge 0$ the Lagrange multipliers for the path constraints, and $\nu \ge 0$ the Lagrange multipliers for the terminal constraints.

The NCO for optimization problem (1)–(4), that is, without parameterization of the input, can be written as follows (Bryson & Ho, 1975; Srinivasan, Palanki, et al., 2003):

	Path	Terminal
Constraints	$\mu^T(t)\, S\Big(x(t), u(t), \rho, \theta\Big) = 0,\ \mu(t) \ge 0$	$\nu^T\, T\Big(x(t_f), \rho, \theta\Big) = 0,\ \nu \ge 0$
Sensitivities	$\frac{\partial H}{\partial u}(t) = 0$	$\frac{\partial \Psi(t_f)}{\partial \rho} = 0$

(10)

3. Numerical optimization using a plant model

The best representative of this class of methods for the case of plant-model mismatch is the two-step approach, which (i) uses a plant model and output measurements to identify uncertain model parameters and (ii) optimize the updated model using CVP. The plant model is of the form:

$$\dot{x}(t) = F(x(t), u(t), \rho, \theta), \quad y_{model}(t) = G(x(t), u(t), \rho, \theta), \tag{11}$$

with the state vector x and the output variables predicted by the model y_{model}. After a brief description of the model-building procedure, the two concepts of "model adequacy" and "modeling for optimization" will be developed. Special attention will be paid to the choice and adaptation of θ to deal with uncertainty.

3.1. Model building

To construct a mathematical model, the modeler typically uses both prior knowledge and measurements from the plant. The modeler goes through several steps that include (i) abstraction from the reality to define the "system", (ii) simplification to arrive at a mathematical model of manageable complexity, (iii) parameter identification to fit the model to the plant, and (iv) model validation to ensure that the model will be useful for its intended goal. Since the later steps influence the early ones, this procedure is typically iterative.

Model identification and validation are typically done by comparing the model prediction $y_{model}(t)$ with the observed plant outputs $y_{plant}(t)$. However, parameter identification is a difficult task in the presence of structural plant-model mismatch, that is, when the plant does not belong to the model set (Ljung, 1999). In this case, parameter identification requires appropriate experimental design (Montgomery, 2005) and persistency of excitation (Walter & Pronzato, 1997).

The model validation step is very important. How can one ensure that the model will be adequate for solving the optimization problem at hand? The criterion $J^{id} = \| y_{model}(t) - y_{plant}(t) \|$ is convenient because the outputs are, by definition, available. Furthermore, it is fully justified if the main purpose of the model is to predict the outputs, for example in a simulation study. But is it still justified if the model is used for optimization? This question is addressed next.

3.2. Model adequacy

It is well known that the two-step approach of repeated parameter identification and performance optimization works well provided that (i) there is no structural plant-model mismatch, that is, the plant lies in the model set and (ii) the operating conditions yield sufficient excitation for all the uncertain model parameters to be estimated (Chen & Joseph, 1987; Marlin & Hrymak, 1997). If these conditions are satisfied, convergence to the plant optimum can be achieved in one iteration (Forbes & Marlin, 1996). Yet, these conditions are rarely met in practice. Regarding the latter condition, in particular, the situation is somewhat similar to that found in the area of system identification and control, where the two tasks of identification and control are typically conflicting (dual control problem Aström & Wittenmark, 1995).

Hence, in the realistic case of plant-model mismatch and/or insufficient excitation, whether the scheme converges, or to which point it converges, becomes anyone's guess. This is due to the fact that the objective for parameter adaptation might be unrelated to the cost and constraints that drive optimality in the optimization problem. Hence, minimizing the mean-square error of the plant outputs, $\|y_{model}(t) - y_{plant}(t)\|$, may not help in our quest for feasibility and optimality.

On the other hand, convergence under plant-model mismatch has been addressed by Biegler, Grossmann, and Westerberg (1985) and Forbes, Marlin, and MacGregor (1994). It has been shown that optimal operation is reached if model adaptation leads to matched KKT conditions for the model and the plant. Hence, the basic idea is to adjust the value of θ in such a manner that the NCO predicted by the model match those of the plant, which naturally leads to the next section concerned with modeling for optimization.

3.3. Modeling for optimization

The basic idea is that the validity of a model depends on its intended use. This concept has been studied extensively in the 1990s in the area of system identification and control under the label "identification for control" (Aström, 1993; Gevers, 2004). The proposed solution has been to use matching criteria for the identification and control tasks. This is obtained by filtering the prediction error (identification) to make it resemble the closed-loop performance criterion (control).

Similarly in optimization, one can consider" modeling for optimization", whereby the synergy between the modeling and optimization steps is improved by reconciling the objective functions of the two problems. In concrete terms, the objective function of the identification problem can be set up to minimize the error between the plant and model NCOs (Srinivasan & Bonvin, 2002).

The idea can be formulated as follows. Let $\mathcal{N}$ be used to represent the components of the NCO in (10). Plant optimality under plant-model mismatch is given by,

$$\underbrace{0 = \mathcal{N}_{plant}}_{objective} = \underbrace{\mathcal{N}_{model}}_{\text{optimization}} + \underbrace{\mathcal{N}_{plant} - \mathcal{N}_{model}}_{\text{identification}}. \tag{12}$$

Plant optimality implies zeroing the necessary conditions, $\mathcal{N}_{plant} = 0$. For its part, numerical optimization of the model enforces $\mathcal{N}_{model} = 0$. Hence, if the identification step can minimize $J^{id} = \|\mathcal{N}_{plant} - \mathcal{N}_{model}\|$, the identified model will be suited for optimization. Such a choice of objective function is referred to as "modeling for optimization".

If the components of the NCO can be measured, y_{NCO}, then the identification criterion based on the minimization of the output error addresses the optimization objective directly. However, the evaluation of certain components of the NCO, such as the plant gradients, is difficult and requires special care with respect to both excitation and noise filtering.

4. Optimizing control using a solution model

As shown in Fig. 1, optimization can also be implemented via feedback control. NCO tracking is organized around a solution model, that is, qualitative knowledge of the optimal solution. This solution model is used to identify the NCO and select the corresponding measurements y_{NCO} and input elements π_{NCO}. To illustrate the NCO-based input parameterization, consider the input $u_1(t)$ in a given time interval. CVP would parameterize $u_1(t)$ using a number of constant values, which are then optimized numerically. However, if we know that the path constraint S_1 is active during that interval and u_1 is the input that pushes S_1 to its bound, an alternative parameterization would be $u_1(t) = \mathcal{G}_c(S_1)$, where $\mathcal{G}_c$ is an appropriate controller that keeps S_1 active. Hence, the parameterization looks for manipulated variables (MVs, the input parameters π_{NCO}) that can be used to track controlled variables (CVs, the elements of y_{NCO}) so as to satisfy the NCO. The solution model is not unique since the NCO depend on the parameterization that is used. Hence, the diversity in solution models can be exploited to ease up implementation.

The development of a solution model involves three steps:

(1) The optimal solution is first characterized in terms of the types and sequence of arcs. This step typically uses the available plant model and numerical optimization to compute the input profiles using CVP.
(2) A finite set of parameters π_{NCO} is selected to represent the input profiles (MVs), the corresponding NCO are formulated from which the CVs are obtained, and the MVs and CVs are paired to form a multivariable control problem.
(3) A robustness analysis is performed to ensure that the structure of the optimal solution is invariant in the presence of uncertainty and the nominal optimal solution are structurally the same. If this is not the case, it is necessary to modify the structure of the solution model and repeat the procedure.

The solution model considers the different parts of the NCO (path constraints, path sensitivities, terminal constraints and terminal sensitivities) that need to be enforced for optimality. The various NCO elements can be implemented with various degrees of ease or difficulty: a path constraint is often enforced on-line via constraint control; a path sensitivity is more difficult to implement as it involves the adjoint variables, which are not available on-line without the use of a plant model; the terminal constraints and sensitivities call for prediction, which requires a model, or else they can be met iteratively over several runs. To ease implementation, it is often possible to approximate the optimal inputs with simpler profiles. This represents the strength of the approach, as the approximations introduced at the solution level can be assessed in terms of optimality loss.

The MVs of the control problem are the handles available to reach optimality. The CVs are the NCO for the selected input parameterization. By definition, there are as many optimality conditions as there are degrees of freedom, thus resulting in a square control system. The pairing of MVs and CVs can be done in a centralized (multivariable control) or decentralized (multi-loop control) fashion. Note that there are different ways of implementing a given solution model, for example using alternative MVs via a change of variables, using different pairings of MVs and CVs, or using a plant model for prediction, each way defining a different NCO-tracking option (Srinivasan & Bonvin, 2007).

5. Choice of a DRTO scheme

The choice of a DRTO scheme is intimately linked to the measurements and the input parameterization as illustrated in Fig. 1 and discussed next.

5.1. Measurements

Measurement availability is key in real-time optimization and determines the way the uncertain model elements can be updated to account for the observed uncertainty:

(1) If the outputs y_{plant} are measured, the uncertainty is accounted for by varying the parameters θ_{plant} of the plant model. Note, however, that optimality cannot in general be guaranteed in the presence of structural plant-model mismatch.
(2) If the NCO elements y_{NCO} are measured, the uncertainty can be expressed as additive disturbances that represent the deviations of the NCO from their ideal values of zero. Hence, special effort is needed to measure or estimate the NCO, which requires excitation and specific measurements.

Which measurements should be used? If both y_{plant} and y_{NCO} are available, it seems easier to implement DRTO via modifier-adaptation or NCO tracking since the two-step approach requires system excitation and numerical optimization to estimate the parameters, another numerical optimization in the re-optimization step and, in many cases, a time-scale separation for the two iterative steps, namely, identification and optimization, to converge. Yet, the two-step approach has been used extensively since it represents the only available option when y_{NCO} is not available. Furthermore, measuring, estimating and approximating certain elements of the NCO is not trivial and represents an open field of research. Hence, for the time being, measuring y_{plant} and expressing the uncertainty through parametric variations of the plant model remains the main option.

5.2. Input parameterization

The input parameterization is important as it affects the choice of the adaptation strategy:

(1) The simplest and most generally valid way of parameterizing input profiles is via π_{CVP}, following which the optimal inputs are determined by numerical optimization.
(2) The alternative is to consider the NCO-based parameterization π_{NCO}, which lends itself to the design of control loops to reject the uncertainty in the NCO seen as additive disturbances.

Table 1
Model parameters and operating bounds.

m	1300	kg
f	0.5	$\frac{\mathrm{N\,s^2}}{\mathrm{m^2}}$
u_{min}	−8000	N
u_{max}	3600	N
v_{max}	40	$\frac{\mathrm{m}}{\mathrm{s}}$
x_{des}	1000	m

The parameterization of choice depends on the problem at hand. The π_{NCO} parameterization is initially more involved as it requires numerical optimization to compute the nominal optimal solution that is used to characterize the NCO. Yet, this parameterization may turn out to be very convenient as optimality can be reached using feedback control, that is, without having to repeat optimization on-line. The success of the NCO-based parameterization depends on its validity since a tailor-made parameterization is seldom universally valid.

5.3. Choice of a scheme

The choice of a scheme depends on several factors:

(1) Are the NCO known and can they be measured or estimated?
(2) Have the NCO been verified to be robust with respect to uncertainty?
(3) Is a multivariable feedback control scheme easily implementable?

In the case of a positive answer to all three questions, NCO tracking is clearly an option for implementing DRTO. Modifier adaptation is a valid option when only question (1) is verified. Otherwise, the two-step approach is still the method of choice because of its generality.

6. Illustrative example

The use of plant and solution models for real-time optimization will be illustrated on the simple car example that is presented next:

- *System*: Movement of a car from one point to another.
- *Uncertainty*: Slope of the road ±5%.
- *Objective*: Minimize final time.
- *Manipulated input*: Accelerating/braking force.
- *Path constraints*: Input bounds; speed limit.
- *Terminal constraints*: Zero velocity at final time; cover at least the prescribed distance.

6.1. Formulation of the optimization problem

6.1.1. Variables and parameters

x: position, v: velocity, u: accelerating/braking force, s: slope of the road, f: friction coefficient, g: gravitational constant, and m: mass of the car. The numerical values of the model parameters and operating bounds are given in Table 1.

6.1.2. Model equations

$$\dot{x} = v \quad x(0) = 0, \tag{13}$$

$$\dot{v} = \frac{u - f v^2}{m} - s(x)g, \quad v(0) = 0. \tag{14}$$

The nominal model assumes zero slope, that is $s(x) = 0$, while in reality the unknown elevation profile, which is the integral of the slope profile, is as shown in Fig. 2.

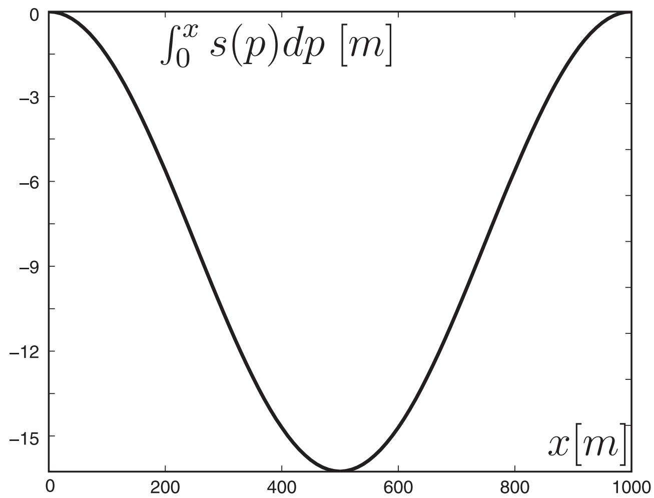

Fig. 2. Elevation profile.

6.1.3. Optimization problem

The optimization problem can be formulated mathematically as follows:

$$\begin{aligned} \min_{u(t),\, t_f} \quad & J = t_f \\ \text{s.t.} \quad & \text{dynamic system (13) and (14)} \\ & u_{min} \le u(t) \le u_{max} \\ & v(t) \le v_{max} \\ & v(t_f) = 0 \\ & x(t_f) \ge x_{des}\,. \end{aligned} \tag{15}$$

6.2. Characterization of the optimal solution

Fig. 3 shows that the optimal solution consists of three arcs, with the successive inputs u_{max}, u_{path} and u_{min}:

- The first arc u_{max} corresponds to maximum acceleration in order to reach v_{max} as quickly as possible. The duration of this arc, t_1, depends on the slope, which is uncertain. However, t_1 can be determined implicitly upon reaching v_{max}, that is $v(t_1) = v_{max}$.

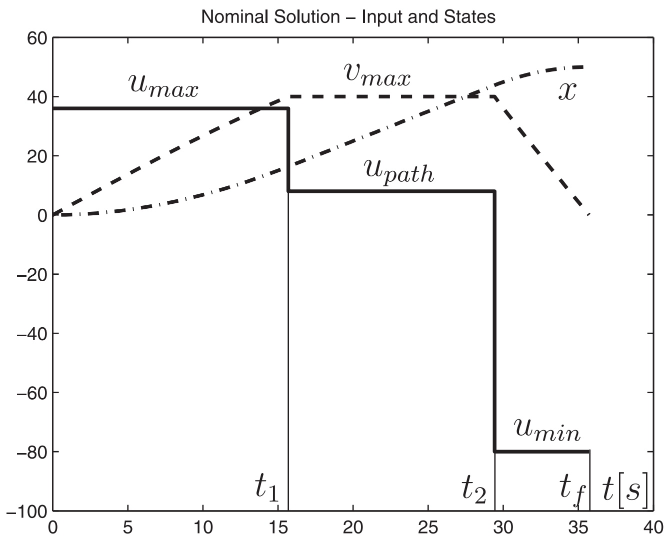

Fig. 3. Nominal optimal solution ($s = 0$): input $u/100$ [N], velocity v [m/s] and position $x/20$ [m] profiles.

Table 2
Pairing of MVs and CVs.

	Path objectives	Terminal objectives
Constraints	$t_1:\ v(t_1) = v_{max}$ $u_{path}(t):\ v(t) = v_{max}$	$t_2 \mapsto x(t_f) = x_{des}$ $t_f:\ v(t_f) = 0$
Sensitivities	–	–

- The second arc keeps the velocity at v_{max}, for which the corresponding input value u_{path} can be determined from (14) as $u_{path} = fv_{max}^2 + s(x)mg$. The value u_{path} is also a function of the uncertain slope.
- The third arc corresponds to full braking in order to achieve $v(t_f) = 0$. The switching time t_2 between the second and third arcs is chosen so that $x(t_f) = x_{des}$, that is, the desired distance will be exactly covered when the velocity goes to zero. The final time is determined upon reaching zero velocity, $v(t_f) = 0$.

6.2.1. Remarks

(1) This car example does not involve sensitivities, that is, the optimal solution is entirely determined by active constraints. The reader is referred to Srinivasan and Bonvin (2007) for cases where sensitivities are involved.

(2) The types and sequence of arcs (u_{max} followed by u_{path} and then u_{min}) hold for any car regardless of its weight and acceleration/braking power. They even hold generically for a bicycle, for which the second arc vanishes (t_1 does not exist). This generic aspect of the optimal solution provides much robustness to the solution-model approach.

6.3. NCO tracking

Input parameterization is straightforward in this problem, with the input parameters $\pi = [t_1\ u_{path}\ t_2\ t_f]^T$. The pairing of MVs ($t_1$, u_{path}, t_2 and t_f) and CVs ($v(t_1) = v_{max}$, $v(t) = v_{max}$, $x(t_f) = x_{des}$, and $v(t_f) = 0$) follows directly from the characterization of the optimal solution and is given in Table 2.

Since there is a prediction involved in the pairing $t_2 \mapsto x(t_f) = x_{des}$, meeting this constraint will either require a predictive model or be implemented over several runs. Next, we present two control strategies that correspond to different ways of adjusting t_2 to meet the terminal constraint $x(t_f) = x_{des}$.

6.3.1. Run-to-run adaptation of t_2

In the absence of a model to predict $x(t_f)$ during the run, NCO tracking will encompass on-line control to enforce $v(t) = v_{max}$ in the second arc and run-to-run control to adapt t_2 so as to satisfy $x(t_f) = x_{des}$ over several runs. The time instants t_1 and t_f are determined by the velocity reaching v_{max} and 0, respectively. It is assumed that the plant (the car with varying unknown slope) will have the same types and sequence of arcs, but different values of t_1, u_{path}, t_2 and t_f.[1]

6.3.2. On-line adaptation of t_2

Prediction of the final position $x(t_f)$ to initiate breaking at t_2 can be done during the run using the nominal plant model that assumes $s(x) = 0$. Since t_1, u_{path} and t_f can be determined from their

[1] This can be verified numerically off-line by perturbing the nominal model and computing the corresponding optimal inputs.

Table 3
Constraint and cost values for various optimization scenarios: $I_{v_+} = \int_{t_1}^{t_2} v_+(t)dt$, where $v_+(t) = v(t) - v_{max}$ if $v(t) > v_{max}$ and 0 otherwise; v_m: maximal velocity; x_f: final position; t_f: final time (cost); t_{faug}: cost penalized for deviation from constraints, each deviation being weighted 5 times the corresponding Lagrange multiplier.

Optimization scenario	$I_{v+}[m]$	$v_m[m/s]$	$x_f[m]$	$t_f[s]$	$t_{faug}[s]$
Ideal solution	0	40	1000	35.31	*35.31*
Open-loop use of nominal optimal input	34.57	43.87	1045.3	35.71	43.70
Re-optimization – no adaptation	1.85	40.87	1003.8	36.00	36.56
Re-optimization – adaptation	0.05	40.08	1004.4	35.44	35.90
NCO tracking – t_2 not adapted	0.003	40.001	1012.2	35.64	36.86
NCO tracking – run-to-run t_2 adaptation	0.002	40.001	1000	35.31	*35.31*
NCO tracking – on-line t_2 adaptation	0.002	40.001	996.5	35.22	35.57

corresponding NCO during simulation (integration) of the dynamic model, the optimization problem (15) can be rewritten as:

$$\min_{t_2} \quad J = t_f$$
$$\text{s.t. } \textit{dynamic} \text{ system (13) and (14)}$$
$$u(t) = \begin{cases} u_{max} & \text{for} \quad 0 \le t < t_1 \\ f v_{max}^2 & \textit{for} \quad t_1 \le t < t_2 \\ u_{min} & \text{for} \quad t_2 \le t < t_f \end{cases} \qquad (16)$$
$$t_1: \; v(t_1) = v_{max}$$
$$t_f: \; v(t_f) = 0.$$

Problem (16) is simpler to solve than the original problem (15) for at least two reasons: (i) the number of degrees of freedom has been reduced from ∞ to 1 and (ii) the discontinuities at the switching instants can be handled much more easily and without oscillations (Schlegel & Marquardt, 2006a). For implementation, the current state information is used as initial conditions for re-optimization at each sampling time. Note that the slope is not updated. The optimal value of t_2 is computed at each re-optimization instant, and breaking is implemented when the running time equals the value of t_2 computed last. As with run-to-run adaptation of t_2, the time instants t_1 and t_f are determined by the velocity reaching v_{max} and 0, respectively.

6.4. Optimization results

The performance of various DRTO schemes is summarized in Table 3 and discussed next. The case of no measurement noise is considered here to be able to focus on the distinguishing features of the various approaches.

The reference scenario is the ideal solution, which is computed with knowledge of the varying slope and is shown in Fig. 4. As expected, this scenario meets exactly all the path and terminal constraints. In comparison, the open-loop use of the nominal optimal input leads to significant violation of the speed constraint as there is no possibility to compensate for the fact that the road goes downhill initially.

The fives DRTO schemes analyzed below correspond to different re-optimization and NCO-tracking choices:

- *Re-optimization – no adaptation*: Uncertainty is considered as additive state disturbances, and CVP is used for numerical optimization. Re-optimization every two seconds without slope adaptation gives the profile shown in Fig. 5 (top-left). There is considerable chattering caused by the need for additional braking in the initial downhill part followed by additional power in the uphill part. The path constraint is not met accurately, as seen by the I_{v_+} values in Table 3. Note that this scheme corresponds to economic nonlinear MPC (Rawlings & Amrit, 2009) and shows the negative effect of not being able to adapt for plant-model mismatch.
- *Re-optimization – adaptation*: Uncertainty is considered as both parametric variations and additive state disturbances, and CVP is used for numerical optimization. With adaptation of the slope parameter, chattering and the deviation from v_{max} in the second arc are reduced (compare the corresponding I_{v_+} values in Table 3). Treating the uncertainty as parametric variations helps track the path constraints, but there is still some residual error. Note that it is very difficult to control the terminal position in a single run because of modeling errors. A comparison of the two input profiles for re-optimization shows that parameter adaptation helps approximate the ideal input (compare Figs. 4 and 5 (top row)).
- *NCO tracking – t_2 not adapted*: Uncertainty is considered as NCO deviations, and NCO tracking is implemented. Without t_2 adaptation, the path constraint $v(t) = v_{max}$ and the final velocity $v(t_f) = 0$ are met, but not $x(t_f) = x_{des}$.
- *NCO tracking – t_2 adapted over several runs*: Similar to the previous case, but t_2 is now adapted on a run-to-run basis. After 5 runs, all the constraints are met and optimality is reached, which in this problem is guaranteed by satisfaction of the constraints on maximum velocity and final position.
- *NCO tracking – on-line adaptation of t_2*: Even when t_2 is adapted, the terminal position constraint $x(t_f) = x_{des}$ is not met exactly since it is adapted on the basis of an inaccurate model. Such terminal objectives can only be met over several runs.

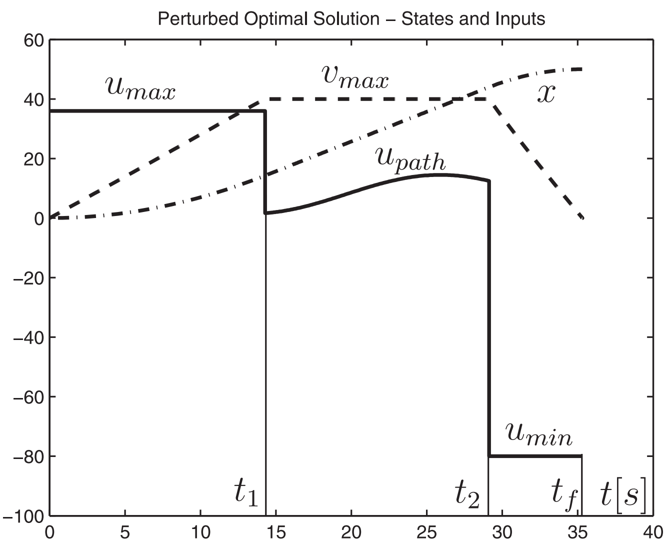

Fig. 4. Ideal solution for known varying slope: input $u/100$ [N], velocity v [m/s] and position $x/20$ [m] profiles.

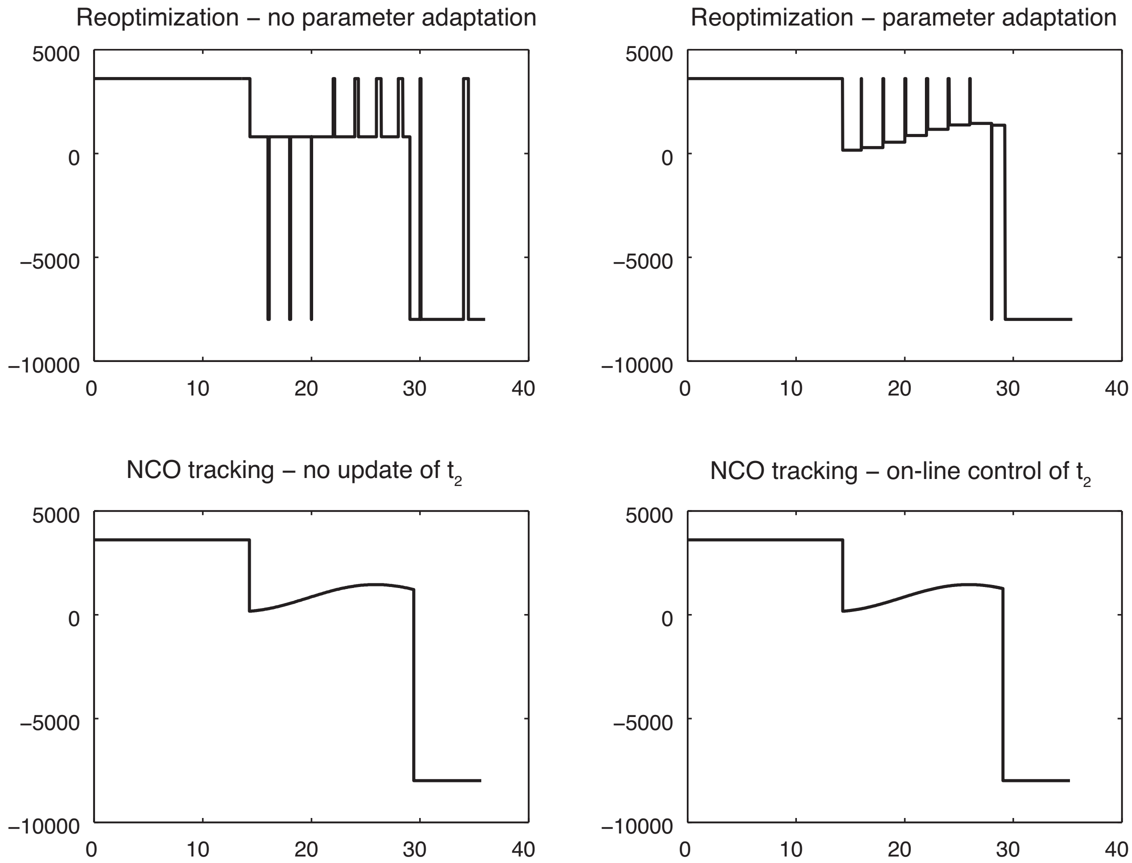

Fig. 5. Optimal input obtained with four DRTO schemes. The case of NCO tracking with t_2 adapted over 5 runs is not shown here as the converged solution corresponds to the ideal solution shown in Fig. 4.

7. Conclusions

This paper has addressed the quality of models used for real-time optimization. According to the quote “All models are wrong but some are useful” (Box, 1979), the model is not viewed as the “truth”, but rather as a tool that must be tailored to the optimization scheme. Modeling is really about making educated approximations to arrive at a model of acceptable complexity that is *adequate for optimization* in the presence of uncertainty. When dealing with a real plant, it is important to find a good way of introducing approximations. Is it at the plant-model level, before going through the optimization machinery? Or is it at the implementation level, when the user can see the implications of selected approximations? It has been argued that the parameters of the plant model adjusted using output measurements might not be the best handles to adjust in real time, since they might only be loosely connected to optimality. By contrast, if the NCO elements are measured or estimated, the optimization objective can be incorporated in the parameter identification step, thereby accommodating the objective of “modeling for optimization”.

The paper has also discussed the various elements of a solution model. Its primitive version is to use piecewise-constant profiles. However, one could devise a more appropriately parameterized solution structure by dissecting the nominal optimal input profiles and relating their elements to different parts of the NCO. Since the complexity of solution models depends on the number of inputs, and not on the number of states or the nonlinearity of the plant, the solution model is easier to obtain for problems with only a few arcs (and thus also only a few input elements π_{NCO}), and this regardless of the order of the system. Such a tailor-made parameterization has been shown to be more efficient than CVP.

One strength of NCO tracking is the possibility of combining off-line tasks (numerical optimization based on the nominal plant model to determine the set of active constraints) and on-line activities (optimizing control that adjusts the inputs on the basis of measurements). Another nice feature is the possibility, if necessary, of introducing approximations in the various profiles to ease implementation. This is particularly effective in dealing with sensitivity-seeking arcs, which are often difficult to compute but, at the same time, contribute only negligibly to the cost. Instead of building a model that will predict the plant performance from scratch, NCO tracking starts with a robust parameterized model of the solution and adjusts the few input parameters that are intimately linked to plant optimality. The synergy between off-line and on-line computations for real-time optimization needs to be studied in more details. Multi-parametric programming and NCO tracking represent two such initial attempts. A big push could come from researchers in the field of optimization once they realize that the available models are not sufficiently accurate to drive plants to optimality.

Acknowledgments

The authors would like to acknowledge insightful discussions with Benoît Chachuat and Grégory François.

References

Aström, K. J. (1993). Matching criteria for control and identification. In *European control conference* (pp. 248–251). Netherlands: Groningen.

Aström, K. J., & Wittenmark, B. (1995). *Adaptive control* (2nd ed.). Reading, MA: Addison-Wesley.

Bemporad, A., Morari, M., Dua, V., & Pistikopoulos, E. N. (2002). The explicit linear quadratic regulator for constrained systems. *Automatica*, *38*, 3–20.

Biegler, L. T. (2009). Technology advances for dynamic real-time optimization. *Computer Aided Chemical Engineering*, *27*, 1–6.

Biegler, L. T., Grossmann, I. E., & Westerberg, A. W. (1985). A note on approximation techniques used for process optimization. *Computers and Chemical Engineering*, *9*, 201–206.

Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer, & G. N. Wilkinson (Eds.), *Workshop on robustness in statistics*.

Bryson, A. E., & Ho, Y.-C. (1975). *Applied optimal control*. Washington DC: Hemisphere Publishing Corporation.

Chachuat, B., Srinivasan, B., & Bonvin, D. (2009). Adaptation strategies for real-time optimization. *Computers and Chemical Engineering*, *33*(10), 1557–1567.

Chen, C. Y., & Joseph, B. (1987). On-line optimization using a two-phase approach: An application study. *Industrial and Engineering Chemistry Research*, *26*, 1924–1930.

D'Amato, F., Kumar, A., Lopez-Negrete, R., & Biegler, L. T. (2012). Fast nonlinear model predictive control: Optimization strategy and industrial process applications. In *Proc. of the chemical process control conference VIII conference*. GA: Savannah Harbor.

Diehl, M., Bock, H. G., & Schlöder, J. P. (2005). A real-time iteration scheme for nonlinear optimization in optimal feedback control. *SIAM Journal on Control and Optimization*, *43*(5), 1714–1736.

Diehl, M., Bock, H. G., Schlöder, J. P., Findeisen, R., Nagy, Z., & Allgöwer, F. (2002). Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. *Journal of Process Control*, *12*(4), 577–585.

Diehl, M., Ferreau, H. J., & Haverbeke, N. (2009). Efficient numerical methods for nonlinear MPC and moving horizon estimation. In L. Magni, D. M. Raimondo, & F. Allgöwer (Eds.), *Nonlinear model predictive control* (pp. 391–417).

Eaton, J. W., & Rawlings, J. B. (1990). Feedback control of nonlinear processes using on-line optimization techniques. *Computers and Chemical Engineering*, *14*, 469–479.

Forbes, J. F., & Marlin, T. E. (1996). Design cost: A systematic approach to technology selection for model-based real-time optimization systems. *Computers and Chemical Engineering*, *20*, 717–734.

Forbes, J. F., Marlin, T. E., & MacGregor, J. F. (1994). Model adequacy requirements for optimizing plant operations. *Computers and Chemical Engineering*, *18*(6), 497–510.

Gevers, M. (2004). Identification for control: Achievements and open problems. In *IFAC DYCOPS-7* (pp. 401–412). MA: Boston.

Heidarinejad, M., Liu, J., & Christofides, P. D. (2012). Economic model predictive control of nonlinear process systems using Lyapunov techniques. *AIChE Journal*, *58*, 855–870.

Kirk, D. E. (1970). *Optimal control theory: An introduction*. London: Prentice-Hall.

Ljung, L. (1999). *System identification: Theory for the user* (2nd ed.). Upper Saddle River, NJ: PTR Prentice Hall.

Marchetti, A., Chachuat, B., & Bonvin, D. (2007). Batch process optimization via run-to-run constraints adaptation. In *European control conference*. Greece: Kos.

Marlin, T. E., & Hrymak, A. N. (1997). Real-time operations optimization of continuous processes. In *AIChE symposium series—Chemical process control conference V, Vol. 93* (pp. 156–164).

Montgomery, D. C. (2005). *Design and analysis of experiments* (6th ed.). New York: John Wiley & Sons.

Pistikopoulos, E. N., Georgiadis, M. C., & Dua, V. (2007). *Multi-parametric model-based control: Theory and applications*. Wiley-VCH.

Rawlings, J. B., & Amrit, R. (2009). Optimizing process economic performance using model predictive control. In L. Magni, D. M. Raimondo, & F. Allgöwer (Eds.), *Nonlinear model predictive control* (pp. 119–138).

Rawlings, J. B., & Mayne, D. Q. (2009). *Model predictive control: Theory and design*. Nob Hill.

Richter, S., Morari, M., & Jones, C. N. (2011). Towards computational complexity certification for constrained MPC based on Lagrange relaxation and the fast gradient method. In *IEEE conference on decision & control* Orlando, FL.

Schlegel, M., & Marquardt, W. (2006a). Detection and exploitation of the control switching structure in the solution of dynamic optimization problems. *Journal of Process Control*, *16*, 275–290.

Schlegel, M., & Marquardt, W. (2006b). Adaptive switching structure detection for the solution of dynamic optimization problems. *Industrial and Engineering Chemistry Research*, *45*, 8083–8094.

Srinivasan, B., & Bonvin, D. (2002). Interplay between identification and optimization in run-to-run optimization schemes. In *American control conference* Anchorage, AK, (pp. 2174–2179).

Srinivasan, B., & Bonvin, D. (2007). Real-time optimization of batch processes via tracking the necessary conditions of optimality. *Industrial and Engineering Chemistry Research*, *46*(2), 492–504.

Srinivasan, B., Bonvin, D., Visser, E., & Palanki, S. (2003). Dynamic optimization of batch processes: II. Role of measurements in handling uncertainty. *Computers and Chemical Engineering*, *27*, 27–44.

Srinivasan, B., Palanki, S., & Bonvin, D. (2003). Dynamic optimization of batch processes: I. Characterization of the nominal solution. *Computers and Chemical Engineering*, *27*, 1–26.

Terwiesch, P., Agarwal, M., & Rippin, D. W. T. (1994). Batch unit optimization with imperfect modeling—A survey. *Journal of Process Control*, *4*, 238–258.

Walter, E., & Pronzato, L. (1997). *Identification of parametric models from experimental data*. Berlin: Springer-Verlag.

Wang, Y., & Boyd, S. (2010). Fast model predictive control using online optimization. *IEEE Transactions on Control Systems Technology*, *18*(2), 267–278.

Zavala, V. M., & Biegler, L. T. (2009). The advanced step NMPC controller: Optimality, stability and robustness. *Automatica*, *45*, 86–93.

Zeilinger, M. N., Jones, C. N., & Morari, M. (2011). Real-time suboptimal model predictive control using a combination of explicit MPC and online optimization. *IEEE Transactions on Automatic Control*, *56*(7), 1524–1534.