Computers and Chemical Engineering 60 (2014) 242–259

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Dynamic optimization using adaptive direct multiple shooting

Fady Assassa[a,b], Wolfgang Marquardt[a,*]

[a] Aachener Verfahrenstechnik, RWTH Aachen University, Turmstr. 46, 52064 Aachen, Germany
[b] German Research School for Simulation Sciences GmbH, 52425 Jülich, Germany

ARTICLE INFO

Article history:
Received 13 December 2012
Received in revised form 4 June 2013
Accepted 20 September 2013
Available online 30 September 2013

Keywords:
Dynamic optimization
Optimal control
Direct multiple shooting
Adaptive grid refinement
Semi-batch reactor
HIPS polymerization

ABSTRACT

We present a wavelet-based grid refinement approach for direct multiple shooting applied to dynamic optimization problems. The algorithm, named adaptive multiple shooting, automatically generates a problem-dependent parameterization of the control profiles: Starting from an initially coarse parameterization, the control grid is refined iteratively using a wavelet analysis of the previously obtained optimal solution. Additional grid points are only inserted where required and redundant grid points are eliminated. Hence, the algorithm minimizes the number of grid points required to obtain accurate optimal control trajectories.

First, we demonstrate the superiority of adaptive grid refinement over an equidistant discretization for the Williams–Otto semi-batch reactor employing multiple and single shooting. Here, the accuracy is checked using an optimal solution obtained by an indirect optimization approach. Second, we successfully demonstrate the efficiency of adaptive grid refinement compared to an equidistant discretization employing multiple shooting to a dynamically unstable HIPS (*h*igh *i*mpact *p*oly*s*tyrene) polymerization reactor.

1. Introduction

In the last years, dynamic optimization problems have received steadily increasing scientific and industrial attention. Typical examples of such optimization problems include optimal control of batch and semi-batch reactors as well as optimal transitions in continuous processes (such as start-up, shutdown, and load or grade changes). The dynamic optimization problems describing these tasks are based on nonlinear process models. However, the derivation of a high-quality control trajectory is still a challenging task, especially for large-scale process models, e.g., those typically occurring in industrial applications.

Typically, dynamic optimization problems are solved by approximating the continuous problem by a finite-dimensional optimization problem. Naively, the control profiles are parameterized uniformly relying on a pragmatically chosen grid. Obviously, the approximation quality depends on the resolution determined by the chosen parameterization. A coarse resolution of the control variables results in low computational effort but might lead to a poor solution quality. However, even a fine resolution, though resulting in a significantly higher computational effort, might often not improve the solution quality satisfactorily. This behavior typically stems from over-parameterization which might lead to oscillation in the computed control profiles. Furthermore, the computed solution may not properly resolve the qualitative structure of the optimal control profile, which is determined by a series of so-called (continuous) arcs delimited by discontinuous jumps (Bryson & Ho, 1975). In such cases, the parameterization is unable to properly localize the switching points which separate the arcs of the optimal control profile (Schlegel & Marquardt, 2006b).

The solution quality can be significantly improved by an adaptive grid refinement of the control parameterization. Such adaptive (re-)parameterization of the control profiles during the optimization aim at an appropriate resolution of the characteristics of the controls, including a precise localization of the switching points, using a minimal number of parameters to avoid over-parameterization and ill-conditioning. Though such adaptation does often not improve the objective function significantly, it provides an accurate resolution of the structure of the control profile, which reflects the process phenomena correctly and hence favors interpretation of the optimal solution and consequently process understanding.

This paper presents a novel dynamic optimization algorithm which combines the adaptive refinement strategy suggested by Schlegel, Stockmann, Binder, and Marquardt (2005) with direct multiple shooting (originally introduced by Bock & Plitt, 1984), which is called *adaptive multiple shooting*. Thus, we separate the control grid from the shooting grid to locally enable a sufficiently fine control parameterization. At the same time, we rely on direct multiple shooting to explicitly address dynamic optimization

* Corresponding author. Tel.: +49 241 8096712.
E-mail address: wolfgang.marquardt@post.rwth-aachen.de (W. Marquardt).

0098-1354/$ – see front matter
http://dx.doi.org/10.1016/j.compchemeng.2013.09.017

embedding unstable initial value problems, which often cannot be solved successfully by (adaptive) single shooting.

The paper is structured as follows. In Section 2, we present a general formulation of dynamic optimization problems and introduce the numerical solution methods this work builds on. Section 3 first introduces some fundamentals of wavelet-based multi-scale analysis, which forms the basis for the subsequently presented refinement of the control parameterization. Section 4 demonstrates the potential of adaptive grid refinement for the Williams–Otto semi-batch reactor. In particular, we compare common equidistant parameterization with the suggested adaptively refined parameterization for single and multiple shooting, respectively. Section 5 compares adaptive and equidistant multiple shooting on dynamic optimization of an unstable polymerization reactor. Finally, Section 6 gives a summary and an outlook.

2. Preliminaries

In this section, we introduce the dynamic optimization problem and numerical solution strategies that are considered in this paper. We also give an overview on adaptive control grid refinement strategies available in the literature.

2.1. Dynamic optimization problem

We consider the continuous dynamic optimization (or optimal control) problem (CDYNOPT)

$$\min_{\boldsymbol{u}(t)} \quad \Phi = \Phi(\boldsymbol{y}(t_f), \boldsymbol{x}(t_f)) \tag{1a}$$

$$\text{s.t.} \boldsymbol{M}\dot{\boldsymbol{y}} = \boldsymbol{f}(\boldsymbol{y}(t), \boldsymbol{x}(t), \boldsymbol{u}(t)), \quad t \in (t_0, t_f], \tag{1b}$$

$$0 = \boldsymbol{g}(\boldsymbol{y}(t), \boldsymbol{x}(t), \boldsymbol{u}(t)), \quad t \in [t_0, t_f], \tag{1c}$$

$$\boldsymbol{y}(t_0) = \boldsymbol{y}_0, \tag{1d}$$

$$\boldsymbol{y}_{min} \le \boldsymbol{y}(t) \le \boldsymbol{y}_{max}, \quad t \in [t_0, t_f], \tag{1e}$$

$$\boldsymbol{x}_{min} \le \boldsymbol{x}(t) \le \boldsymbol{x}_{max}, \quad t \in [t_0, t_f], \tag{1f}$$

$$\boldsymbol{u}_{min} \le \boldsymbol{u}(t) \le \boldsymbol{u}_{max}, \quad t \in [t_0, t_f], \tag{1g}$$

$$0 \ge \boldsymbol{e}(\boldsymbol{y}(t_f), \boldsymbol{x}(t_f)). \tag{1h}$$

The semi-explicit differential algebraic (DAE) model is represented by the differential equations $\boldsymbol{f} : \mathbb{R}^{n_y} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_y}$ (1b) with the regular matrix $\boldsymbol{M}$ and the algebraic equations $\boldsymbol{g} : \mathbb{R}^{n_y} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ (1c). We pragmatically consider matrix $\boldsymbol{M}$ to be constant and the DAE system to have a differential index of 1 or less. Note that high-index DAE systems can be reduced to index-one DAE systems by one of the existing symbolical algorithms (e.g. Bachmann, Brüll, Mrziglod, & Pallaske, 1990; Mattsson & Söderlind, 1993; Pantelides, 1988; Unger, Kröner, & Marquardt, 1995). Here, $\boldsymbol{y}(t) \in \mathbb{R}^{n_y}$ denotes the vector of differential and $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$ algebraic states with $\boldsymbol{y}_0$ referring to the vector of consistent initial conditions. The time-dependent control variables $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$ are the decision variables of the continuous optimization problem. The final time t_f can either be fixed or free. Optimal control problems with a free final time can be easily transformed into an equivalent optimal control problem with fixed final time. In this case, the DAE system is scaled with a time-invariant parameter (a decision variable), which represents the final time t_f. This generalizes the formulation of CDNYOPT and allows treating the sensitivity of the final time parameter similarly to time-invariant controls. Without loss of generality, the objective function $\Phi : \mathbb{R}^{n_y} \times \mathbb{R}^{n_x} \to \mathbb{R}$ is of Mayer-type, since an integral cost can be reformulated as a terminal cost by including additional differential variables and equations to the DAE system (Hazewinkel, 2001, cf. Bolza problem). Furthermore, state and control constraints, Eqs. (1e), (1f) and (1g), respectively, are defined on the entire optimization horizon $I := [t_0, t_f]$ whereas the terminal constraints $\boldsymbol{e} : \mathbb{R}^{n_y} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_e}$, i.e. Eq. (1h), are only formulated at final time t_f. Note that combined state and control constraints can be converted easily to the considered formulation by including additional variables and equations to the DAE system (e.g., Hannemann, 2012).

For brevity and without loss of generality, we neither consider additional time-invariant system parameters nor the initial conditions $\boldsymbol{y}_0$ as additional decision variables of the optimization problem.

2.2. Numerical solution strategies

The solution strategies for dynamic optimization problems of type CDYNOPT can be classified into indirect and direct methods. Binder, Blank, Bock, et al. (2001) and Srinivasan, Palanki, and Bonvin (2003) have reviewed the specific advantages and disadvantages of the two approaches in detail. In particular, indirect methods require knowledge on the structure of the optimal control profiles in order to derive the boundary value problem representing the first-order necessary conditions of optimality (Srinivasan et al., 2003). In case of nonlinear path constrained problems involving states and controls, as they are frequently encountered in chemical engineering applications (Biegler, 2010), this solution structure is not known a priori; thus, the application of the indirect approach is rather cumbersome or even impossible (Binder, Blank, Bock, et al., 2001). In this contribution, we therefore focus on direct methods, which solve CDYNOPT by adaptive control-vector parameterization.

In case of direct methods, the infinite-dimensional dynamic optimization problem is approximated by a finite-dimensional (algebraic) optimization problem, which is solved by means of a suitable nonlinear programming (NLP) algorithm. Three different types of direct approaches have been established: single shooting, also referred to as control-vector parameterization (Sargent & Sullivan, 1978), multiple shooting (Bock & Plitt, 1984), and simultaneous state and control-vector parameterization, often called full discretization approach (Biegler, 1984, 2010; Kraft, 1985). These approaches differ in the way the variables of CDYNOPT are discretized.

In this work, we employ direct multiple shooting to solve the problem CDYNOPT because it can cope with DAE models that show unstable dynamical behavior.

2.3. Transcription of CDYNOPT by direct multiple shooting

In direct multiple shooting (Bock & Plitt, 1984), the time horizon $I := [t_0, t_f]$ of CDYNOPT is first divided into N equidistant shooting intervals with $I_j := [t_j, t_{j+1}]$, $j = 0, \ldots, N-1$ and $t_0 < t_1 < t_2 < \cdots < t_N = t_f$. Next, the control variables $\boldsymbol{u}(t)$ for each shooting interval are explicitly parameterized. This can, for example, be done using B-splines (de Boor, 1978) of order r such that the parameterized controls $\tilde{u}_i^j$ on shooting interval j are represented by

$$\tilde{u}_i^j(\boldsymbol{c}_i^j, t) = \sum_{l \in \Delta_{u_i}^j} c_{i,l}^j \varphi_l^{r,j}(t), \quad i = 1, \ldots, n_u, \tag{2}$$