Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

Hybrid shooting - a new optimization method for unstable dynamical systems

Arndt Hartwich [*,**,⋆] **Fady Assassa** [*,**] **Wolfgang Marquardt** [*]

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

Abstract: Chemical processes are modeled by large-scale, highly-nonlinear process models often governed by unstable dynamics. Dynamic optimization is required to exploit economical performance of these processes in their unstable regions. On the one hand, direct single shooting is able to solve large-scale dynamic optimization problems, but lacks the ability to cope properly with unstable dynamical systems. On the other hand, the multiple shooting method is capable of dealing with instabilities but results in larger optimization problems. This work deals with a novel parameterization approach, termed *hybrid shooting*, which combines the advantages of single and multiple shooting into one approach. Similar to multiple shooting, stages are introduced and states with unstable properties are parameterized by free initial values at the beginning of each stage. This approach significantly enhances the behavior of the optimization problem by improving the condition of the underlying initial value problem (IVP).

Keywords: hybrid shooting, direct single shooting, direct multiple shooting, dynamic optimization, unstable dynamical systems, initial value problem

1. INTRODUCTION

Multi-product and multi-purpose plants comprising batch processes and continuous processes with frequent product specification changes have become more common in the chemical industry. A promising technique to cope with these transient nonlinear processes are model-based methods, such as off-line optimization, model-predictive control, etc. which ensure flexible and profitable operation with economic objectives. Chemical processes, in particular reactors, often exhibit unstable dynamics coupled with strong nonlinearities between the reaction kinetics and the cooling system. However, the optimal operating point is often located inside or close to unstable regions of operation. Hence, an efficient solution procedure for large-scale unstable problems is required.

Dynamic optimization builds the foundation of model-based optimal control methods. The single shooting method (Sargent and Sullivan, 1978) solves large-scale dynamic optimization problems efficiently, but lacks the ability to cope with unstable dynamics. In single shooting the process model is simulated on the entire optimization time horizon as one initial value problem (IVP). Contrary, multiple shooting (Bock and Plitt, 1984) separates the optimization time horizon into several smaller independent IVPs on each stage. The initial conditions of the IVPs, except of the first one, are treated as additional decision variables of the optimization problem. In order to link the stages, each differential state has to match the final value of the preceding stage at the stage boundaries on convergence.

However, the drawback for both shooting approaches is that in the presence of unstable IVPs, the objective function and the constraints are sensitive towards small changes in the initial values, which impacts on the condition of the mathematical program. Following Ascher et al. (1995), a problem is well-conditioned if a small change in the data, i.e. the decision variables, leads to a small change in the results, i.e. state variables. An ill-conditioned ODE may cause severe problems for the numerical integrator when solving the IVP. This issue can be resolved in multiple shooting by inserting more stages into the parameterization. However, this well-conditioning is achieved through an increase in the number of decision variables by introducing additional initial values on the additional stages.

The main idea of the novel hybrid shooting algorithm is based on a mixture of single and multiple shooting. This approach aims at minimizing the number of initial values i.e. decision variables, which are introduced to cope with unstable dynamic behavior. Thus, a subpart of the state vector is solved according to multiple shooting, whereas the remaining part of the state vector is solved according to single shooting. The choice of the subspaces has to be performed such that the resulting IVP is well-conditioned. Hence, we leave those states continuous which do not affect unstable modes and discretize the states that have a strong impact on unstable modes. The subspaces are separated pragmatically via multiple local linearizations along the trajectory followed by an analysis of eigenvalues and eigenvectors.

⋆ Present address: Bayer Technology Services GmbH, 51368 Leverkusen

In this paper, hybrid shooting is demonstrated for a case study, which aims at the optimal control of an unstable CSTR presented by Uppal et al. (1974). The dynamic model exhibits unstable modes.

The paper is organized as follows. First, the problem formulation along with established solution methods is briefly sketched. In Section 3, the nonlinear hybrid shooting approach is introduced. Section 4 illustrates numerical experiments. Finally, Section 5 concludes the paper and gives directions for future work.

The algorithms presented, have been implemented into DyOS (Dynamic Optimization Software), which is developed and maintained by AVT - Process Systems Engineering, RWTH Aachen University. Most recently, efforts are being made to include DyOS in the software package gPROMS, a commercial model builder of Process Systems Enterprise (PSE).

2. DYNAMIC OPTIMIZATION ALGORITHMS FOR NONLINEAR SYSTEMS

This section introduces the optimal control problem for nonlinear systems and, then, briefly reviews established numerical solution methods. We have recently provided a wider overview on the available methods (Hartwich, 2010), with particular emphasis on a condition analysis of direct single and multiple shooting. Here, we only provide the information necessary for this work.

2.1 Problem formulation

We consider the continuous dynamic optimization problem

$$\min_{\mathbf{u}(t)} \Phi = \Phi(\mathbf{y}(t_f), \mathbf{u}(t_f)) \qquad \text{(CDYNOPT)}$$

$$s.t.\ \dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}(t), \mathbf{u}(t)), \qquad t \in (t_0, t_f], \qquad (1)$$

$$0 \geq \mathbf{d}(\mathbf{y}(t), \mathbf{u}(t)), \qquad t \in [t_0, t_f], \qquad (2)$$

$$0 \geq \mathbf{e}(\mathbf{y}(t_f), \mathbf{u}(t_f)), \qquad (3)$$

$$\mathbf{u}_{min}(t) \leq \mathbf{u}(t) \leq \mathbf{u}_{max}(t), \qquad t \in [t_0, t_f], \qquad (4)$$

$$\mathbf{y}(t_0) = \mathbf{y}_0. \qquad (5)$$

The dynamic process model (1) is represented as an ODE. Correspondingly, $\mathbf{y} \in \mathbb{R}^{n_y}$ denote the differential state variables and $\mathbf{u} \in \mathbb{R}^{n_u}$ the control variables, respectively. Path constraints (2) are valid on the entire optimization horizon $I = [t_0, t_f]$, whereas the terminal constraints (3) are only formulated for t_f. The input bounds are described by Eq. (4). The initial state $\mathbf{y}(t_0)$ is only valid for t_0 as denoted in Eq. (5). For the sake of notational brevity and without loss of generality, we omit to consider time-invariant parameters and a free final time t_f.

2.2 Numerical solution methods

In chemical engineering applications, three different dynamic optimization approaches have been established, i.e. single (Sargent and Sullivan, 1978) and multiple (Bock and Plitt, 1984) shooting methods and the simultaneous approach (Biegler, 1984). We restrict our review to the shooting in this work.

Single shooting. This *direct method* solves the dynamic optimization problem (CDYNOPT) numerically according to the *control vector parameterization approach.* CDYNOPT is transcribed into an NLP by discretizing the controls $\mathbf{u}(t)$ by some suitable approximation. A control vector parameterization is for example performed by

$$u_i(t) = \sum_{l=1}^{N} c_{il} \varphi_l^{(r)}(t),\ i = 1, \ldots, n_u, \qquad (6)$$

with constant parameters $c_{il} \in \mathbb{R}^{n_u \times N}$ parameterizing the n_u controls by N time-variant basis functions $\varphi_l^{(r)}(t) \in \mathbb{R}^N$, e.g. B-splines of order r are utilized. Choosing the vector of control parameters as the n_D decision variables, i.e. $\boldsymbol{\zeta} = [\mathbf{c}]$, the transcription into an NLP results in

$$\min_{\boldsymbol{\zeta}} \Phi = \Phi(\boldsymbol{\zeta}) \qquad \text{(NLP)}$$

$$\text{s.t.}\ \ 0 \geq \mathbf{d}(\boldsymbol{\zeta}), \qquad (7)$$

$$0 \geq \mathbf{e}(\boldsymbol{\zeta}), \qquad (8)$$

$$\boldsymbol{\zeta}_{min} \leq \boldsymbol{\zeta} \leq \boldsymbol{\zeta}_{max}. \qquad (9)$$

The path constraints (2) have been relaxed by a pointwise approximation on the discretization grid. The typically small-scale and dense NLP is solved by a standard SQP algorithm (Nocedal and Wright, 1999). Since the optimization algorithm requires repetitive function and gradient evaluations, the objective function $\Phi(\cdot)$, the constraints (2) and (3) as well as their gradients need to be evaluated.

Multiple shooting. In contrast to single shooting, the time horizon is divided into a set $\Xi = [1, ..., J]$ of subintervals or stages on which the dynamic model is numerically integrated independently. For the sake of notational brevity and without loss of generality, we omit to regard the length of each stage j as decision variable and consider only an equidistant discretization with $\Delta\boldsymbol{\tau} = (t_f - t_0)/J$. In order to link these stages, additional variables $\mathbf{y}_0^{(j)}$ are introduced that describe the initial values of each stage. The n_D decision variables of the NLP arising from multiple shooting are the discretized controls and the initial states, i.e. $\boldsymbol{\zeta} = [\mathbf{c}^{(1)^T}, \mathbf{c}^{(2)^T}, \mathbf{y}_0^{(2)^T}, \ldots, \mathbf{c}^{(J)^T}, \mathbf{y}_0^{(J)^T}]^T$. Additionally, the problem (NLP) has to be expanded by the constraints

$$\mathbf{y}_0^{(j)} = \mathbf{y}^{(j-1)}(t^{(j)}),\ j \in \Xi \backslash 1. \qquad (10)$$

The states at the exit point $\mathbf{y}^{(j-1)}(t^{(j)})$ of stage $(j-1)$ are matched with the initial values of the state variables $\mathbf{y}_0^{(j)}$ of the stage j at the entry point $t^{(j)},\ j = 2, \ldots, J$. The free initial conditions $\mathbf{y}_0^{(j)}$ only fulfill the state continuity condition at the optimal solution, hence, the method follows an infeasible path strategy. However, this property enhances the behavior of the associated IVPs in the presence of unstable dynamics.

3. HYBRID SHOOTING

The fundamental idea of hybrid shooting is to discretize as few states as possible $\mathbf{y}_{\mathcal{U}} \in \mathbb{R}^{n_{\mathcal{U}}}$, whereas the remaining states $\mathbf{y}_{\mathcal{S}} \in \mathbb{R}^{n_{\mathcal{S}}}$ are continuous over the time horizon with $\mathbf{y} = \left[\mathbf{y}_{\mathcal{S}}^T, \mathbf{y}_{\mathcal{U}}^T\right]^T$. Roughly speaking, $\mathbf{y}_{\mathcal{S}}$ is treated as in the single shooting approach while $\mathbf{y}_{\mathcal{U}}$ is treated as in the multiple shooting approach. The separation of the state vector is pragmatically guided by an analysis of the eigenvalues and eigenvectors at multiple local linearizations along the trajectory. Few matching conditions are introduced at the

stage boundaries. This combination is expected to yield an NLP that is easier to solve compared to multiple shooting.

Figure 1 illustrates and compares the ideas of all three shooting approaches. In all cases, two states y_1 and y_2 are depicted on the time horizon $I = [t_0, t_f]$:

- In case of single shooting (cf. Fig. 1.a), all states are continuous.
- In case of multiple shooting (cf. Fig. 1.b) both states are discontinuous at all stage boundaries ($t = t^{(2)},\ t^{(3)}$) with the initial values $y_{0,i}^{(j)},\ i = 1, 2,\ j = 2, 3$, as additional decision variables. The chosen number of three stages is arbitrary.
- In case of hybrid shooting (cf. Fig. 1.c), state y_1 is discontinuous at the stage boundaries ($t = t^{(2)},\ t^{(3)}$), with the initial values $y_{0,1}^{(j)},\ j = 2, 3$, as additional decision variables. However, state y_2 remains continuous.

The challenges of hybrid shooting are an appropriate choice of the stage length and the detection of states belonging to the subspace $\mathcal{U}$ and $\mathcal{S}$, respectively. In addition, an appropriate choice of the control grid is paramount to allow for the proper control of the unstable modes.

3.1 Determination of subspaces

The idea to split the state vector stems from the requirement to generate a well-conditioned IVP.

We define the condition number κ as the amplification of the perturbation $\boldsymbol{\beta}$ by the system according to

$$||\mathbf{y}(\boldsymbol{\xi}) - \mathbf{y}(\boldsymbol{\xi} + \boldsymbol{\beta})||_2 \leq \kappa\epsilon \tag{11}$$

where $\boldsymbol{\xi} = [\mathbf{y}_0^T, \mathbf{c}^T]^T$ and $||\boldsymbol{\beta}||_2 \leq \epsilon$. Applying the mean value theorem for multi-variable functions on Eq. (11), we obtain

$$||\mathbf{y}(\boldsymbol{\xi}) - \mathbf{y}(\boldsymbol{\xi} + \boldsymbol{\beta})||_2 \leq ||\nabla\mathbf{y}(\boldsymbol{\xi})||_2 \cdot ||\boldsymbol{\beta}||_2 \leq \kappa\epsilon \tag{12}$$

with

$$\kappa = ||\nabla\mathbf{y}(\boldsymbol{\xi})||_2 = \left|\left|\left(\frac{\partial\mathbf{y}}{\partial\mathbf{y}_0}^T \ \frac{\partial\mathbf{y}}{\partial\mathbf{c}}^T\right)^T\right|\right|_2. \tag{13}$$

For a detailed analysis of the condition number the reader is referred to Hartwich (2010). Here, we will only illustrate the idea, how to separate the state space of a time-invariant linear system.

With the linear system

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u}, \quad \mathbf{y}(t_0) = \mathbf{y}_0, \tag{14}$$

with $\mathbf{A} \in \mathbb{R}^{n_y \times n_y}$ and $\mathbf{B} \in \mathbb{R}^{n_y \times n_u}$, we will carry out a stability analysis. We assume $\mathbf{B}$ to be bounded as $||\mathbf{B}||_2 \leq K$. Furthermore, using the transformation $\mathbf{z} = \mathbf{V}^{-1}\mathbf{y}$, where $\mathbf{V}$ is the eigenvector matrix of $\mathbf{A}$, we define the modal coordinates

$$\dot{\mathbf{z}}(t) = \boldsymbol{\Lambda}\mathbf{z}(t) + \mathbf{V}^{-1}\mathbf{B}\mathbf{u}, \quad \mathbf{z}(0) = \mathbf{z}_0. \tag{15}$$

For brevity, we assume here only real eigenvalues of multiplicity one ordered as

$$\boldsymbol{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{P} & 0 & 0 \\ 0 & \mathbf{N} & 0 \\ 0 & 0 & \mathbf{Q} \end{bmatrix} \tag{16}$$

with $\mathbf{P}$ being a diagonal matrix of n^- negative eigenvalues, $\mathbf{Q}$ a diagonal matrix having n^+ positive eigenvalues and $\mathbf{N}$ a diagonal matrix having n^0 zero eigenvalues. The vector $\mathbf{z}$ can be split into a vector $\mathbf{z_P}$ associated to $\mathbf{P}$, a vector $\mathbf{z_N}$ associated to $\mathbf{N}$ and a vector $\mathbf{z_Q}$ associated to $\mathbf{Q}$, i.e. $\mathbf{z} = [\mathbf{z}_\mathbf{P}^T,\ \mathbf{z}_\mathbf{N}^T,\ \mathbf{z}_\mathbf{Q}^T]^T$. All three sub-vectors are obviously independent of each other. Thus, $\mathbf{V}$ and $\mathbf{V}^{-1}$ can be split accordingly:

$$\mathbf{V} = [\mathbf{V_P}\ \mathbf{V_N}\ \mathbf{V_Q}], \quad \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{S_P} \\ \mathbf{S_N} \\ \mathbf{S_Q} \end{bmatrix}. \tag{17}$$

The eigenvectors are chosen such that $||\mathbf{V}||_2 = 1$ and $\left|\left|\mathbf{V}^{-1}\right|\right|_2 \leq \theta$ are bounded. Next, the states are separated into

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_\mathcal{S} \\ \mathbf{y}_\mathcal{U} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_\mathcal{S} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{y}_\mathcal{U} \end{bmatrix}, \tag{18}$$

with $\mathbf{y}_\mathcal{U}$ in subspace $\mathcal{U}$ and $\mathbf{y}_\mathcal{S}$ in subspace $\mathcal{S}$, respectively, with $n_y = n_\mathcal{U} + n_\mathcal{S}$. Thus, the eigenvector matrix $\mathbf{V}$ and its inverse $\mathbf{V}^{-1}$ are split accordingly:

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_\mathbf{P}^\mathcal{S} & \mathbf{V}_\mathbf{N}^\mathcal{S} & \mathbf{V}_\mathbf{Q}^\mathcal{S} \\ \mathbf{V}_\mathbf{P}^\mathcal{U} & \mathbf{V}_\mathbf{N}^\mathcal{U} & \mathbf{V}_\mathbf{Q}^\mathcal{U} \end{bmatrix}, \quad \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{S}_\mathbf{P}^\mathcal{S} & \mathbf{S}_\mathbf{P}^\mathcal{U} \\ \mathbf{S}_\mathbf{N}^\mathcal{S} & \mathbf{S}_\mathbf{N}^\mathcal{U} \\ \mathbf{S}_\mathbf{Q}^\mathcal{S} & \mathbf{S}_\mathbf{Q}^\mathcal{U} \end{bmatrix}. \tag{19}$$

Similar to multiple shooting, the time horizon is divided into a set $\Xi = [1, ..., J]$ of stages, with an equidistant stage length $\Delta\tau = (t_f - t_0)/J$. Then, we can write the condition number κ for stage j as

$$\kappa(\Delta\tau) \leq \eta\theta \left(\Upsilon^{(j-1)}(2 + D) + \sum_{i=1}^{j} \left(\Upsilon^{(j-i)} \left[2 + D + K\sqrt{N\Delta\tau} \left[\frac{1}{||\mathbf{P}||_2} + \Delta\tau + \frac{D}{||\mathbf{Q}||_2} \right] \right] \right) \right) \tag{20}$$

with $\eta = \sqrt{n_\mathcal{S} + jn_\mathcal{U} + jn_u N}$, $\Upsilon = \left|\left|\mathbf{S}_\mathbf{P}^\mathcal{S}\right|\right|_2 + \left|\left|\mathbf{S}_\mathbf{N}^\mathcal{S}\right|\right|_2 + D\left|\left|\mathbf{S}_\mathbf{Q}^\mathcal{S}\right|\right|_2$, $D = \left|\left|\mathrm{tr}\left[\mathrm{e}^{\mathbf{Q}\Delta\tau}\right]\right|\right|_2$. The condition number κ strongly depends on Υ. Except for $D = \left|\left|\mathrm{tr}\left[\mathrm{e}^{\mathbf{Q}\Delta\tau}\right]\right|\right|_2$ all entries of Υ are bounded and do not lead to ill-conditioning. However, the entry $D\left|\left|\mathbf{S}_\mathbf{Q}^\mathcal{S}\right|\right|_2$ is raised to the power of $(j-1)$ in Eq. (20). If positive eigenvalues $\mathbf{Q}$ are present in $\boldsymbol{\Lambda}$ and $\left|\left|\mathbf{S}_\mathbf{Q}^\mathcal{S}\right|\right|_2 \neq 0$, terms of type D^j occur and dominate the numerical value of the condition number κ of hybrid shooting. In this case hybrid shooting is ill-conditioned. If, however $\left|\left|\mathbf{S}_\mathbf{Q}^\mathcal{S}\right|\right|_2 = 0$, hybrid shooting provides a condition number which only depends on D. Depending on the problem at hand, if $D\left|\left|\mathbf{S}_\mathbf{Q}^\mathcal{S}\right|\right|_2 < \mu$ then Υ is moderate in value and the IVP to be solved in hybrid shooting possesses a moderate condition number.

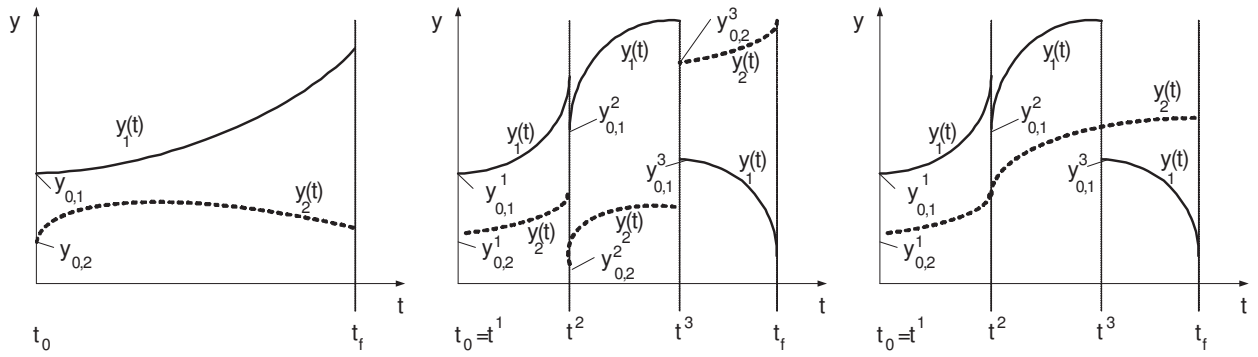
We pragmatically apply this analysis to the locally linearized model

$$\Delta\dot{\mathbf{y}}_{\tilde{j}} = \mathbf{A}_{\tilde{j}}\Delta\mathbf{y}_{\tilde{j}} + \mathbf{B}_{\tilde{j}}\Delta\mathbf{u}_{\tilde{j}}, \quad \Delta\mathbf{y}(t_{\tilde{j}}) = \Delta\mathbf{y}_{\tilde{j}}, \tag{21}$$

of the nonlinear system (1) along the considered trajectory, which is carried out at multiple reference points $\tilde{j} = [1, ..., n_{\tilde{l}}]$.

3.2 Algorithm

We want to minimize the number of boundary conditions and initial values to the extent possible. Therefore, we



(a) single shooting (b) multiple shooting (c) hybrid shooting

Fig. 1. Single shooting, multiple shooting and hybrid shooting: continuity of states during iterations of the NLP and initial conditions of the stages.

solve an optimal control problem by means of single shooting if the modeled system shows stable dynamics. For unstable systems we have to apply a discretization of the states as it is performed for multiple shooting. However, the dynamics of the system at hand are unknown a priori. Even if a system is known to have positive eigenvalues in short parts of the considered time horizon, it is not straightforward to see whether any intermediate iterations or the optimal solution would show unstable behavior on most of the time horizon. Hence, we have to adapt the discretization to the problem while the optimization algorithm is carried out.

During the initialization of the hybrid shooting algorithm we first parameterize the problem according to single shooting and start the optimization. If the problem is stable the optimal solution would be found by means of single shooting. However, single shooting often does not converge if the problem is unstable. Therefore, we need to verify, whether the problem is stable or not after the initialization and during the optimization. The local stability of the ODE system at $t = t_{\tilde{j}}$ is determined by checking the eigenvalues $\mathbf{\Lambda}_{\tilde{j}}$ of the Jacobian matrix $\boldsymbol{A}_{\tilde{j}} = \mathbf{f}_y(\mathbf{y}, \mathbf{u})|_{t=t_{\tilde{j}}}$, which is the system matrix of the local linearization (21) of the considered system. If one positive eigenvalue $\lambda_{i,\tilde{j}} > 0$ appears in one of the vectors $\mathbf{\Lambda}_{\tilde{j}}$, we have detected a locally (at $t = t_{\tilde{j}}$) unstable mode and the optimization using single shooting is terminated.

Obviously, this local analysis lacks precision and the stability properties of the nonlinear dynamic system cannot be assessed rigorously. In addition, positive eigenvalues may emerge or vanish along the trajectory. Therefore, the largest positive eigenvalues of each linearization along the trajectory are averaged by calculating the measure

$$\lambda^* = \sum_{\tilde{j}}^{n_{\tilde{l}}} = \frac{\max_k(\lambda_k(t_j), 0)}{n_{\tilde{l}}}, \; k = 1, ..., n_y, \tag{22}$$

with $n_{\tilde{l}}$ indicating the number of linearization points. The problem is then transcribed into the hybrid shooting formulation relying on the eigenvalue and eigenvector analysis (cf. Section 3.1) and the optimization is resumed.

In order to allow for a good comparison of hybrid and multiple shooting in the numerical experiments, we have to employ comparable initial guesses to the decision variables. In case of hybrid shooting, we use the current intermediate solution gained from single shooting to interpolate the initial guess of the decision variables for hybrid shooting. The same procedure is applied to multiple shooting as well. However, the computational statistics neglect the initialization obtained with single shooting, since the regarded example comprises positive eigenvalues instantly.

3.3 Determination of the stage length $\Delta\tau$

Usually, determining the stage length of multiple and hybrid shooting approaches requires model knowledge. However, the eigenvalue analysis yields information about the system that is reused to automatically determine the stage length $\Delta\tau$. The following heuristic requirement for the stage length is formulated

$$\frac{z_i(t^{(j+1)})}{z_i(t^{(j)})} = 2 = \exp(\lambda^* \Delta\tau), \tag{23}$$

which means that the mode z_i, that belongs to the averaged largest eigenvalue λ^* (cf. Eq. (22)), may not more than double its value on one stage. Thus,

$$\Delta\tau = \ln(2)/\lambda^* \tag{24}$$

is obtained.

3.4 Reparameterization

Now, with the stable and the unstable subspace known along with the stage length, the problem can be reparameterized. Choosing the discretized controls and the initial values, i.e. $\boldsymbol{\zeta} := [\mathbf{c}^{(1)^T}, \mathbf{c}^{(2)^T}, \mathbf{y}^{(2)^T}_{\mathcal{U},0}, \ldots, \mathbf{c}^{(J)^T}, \mathbf{y}^{(J)^T}_{\mathcal{U},0}]^T$, as the n_D decision variables, the dynamic optimization problem can be transcribed into an NLP.

After reparameterization of the new problem, initial values for the decision variables $\mathbf{c}^{(j)}$ and $\mathbf{y}^{(j)}_{\mathcal{U},0}$ need to be obtained. The initial values of the control variables and the initial values of the states are obtained by interpolation from the last iteration of the preceding problem.

3.5 Computation of sensitivities

The predominant computational effort in shooting methods is spent on the computation of the NLP gradients, termed sensitivities $\mathbf{s}^y_i(t) = \frac{\partial \mathbf{y}(t)}{\partial \zeta_i}$ (Vassiliadis et al.,

1994a,b; Hartwich, 2010). For all shooting approaches, the highest sensitivity load is located in the final interval of the control grid of the final stage. However, in case of multiple shooting all stages have the same sensitivity load (assuming equidistant discretization of the controls) since the IVPs on the stages are decoupled. For single and hybrid shooting the sensitivities of any previous stage is propagated through the entire time horizon due to the presence of states, which are continuous over the stage boundaries (see Figure 1). Hence, the number of sensitivity systems that have to be solved simultaneously depends on the number of control parameters on each stage n_c^j, the number of unstable states $n_{\mathcal{U}}$, the total number of states n_y and the number of stages J.

For multiple shooting the maximum number of sensitivity systems amounts to $N_{s,ms} = n_c^j + n_y$. In case of single shooting, we obtain $N_{s,ss} = n_c^j \cdot J$, whereas hybrid shooting possesses $N_{s,hs} = n_c^j \cdot J + n_{\mathcal{U}} \cdot (J-1)$ sensitivity systems. Thus, the sensitivity load of hybrid shooting is like that of single shooting extended by an additional term for the initial values of the unstable states $\mathbf{y}_{\mathcal{U}}$.

These three approaches can be particularly preferable for different settings of n_c^j, $n_{\mathcal{U}}$, n_y and the number of stages J. In practical applications, it is assumed that $n_{\mathcal{U}} << n_y$, such that the sensitivity load of hybrid shooting is moderate for unstable systems. However, it seems to be reasonable to entirely switch to multiple shooting if $n_{\mathcal{U}} \sim n_y$.

4. NUMERICAL EXPERIMENTS

In the following, a nonlinear problem is examined to illustrate the properties of the novel hybrid shooting (HS) method in comparison to single (SS) and multiple shooting (MS). We consider an unstable cooled CSTR with an exothermic first order reaction $A \rightarrow B$ (Uppal et al., 1974), perfect level and temperature control. In this study the inner cascade loop of the temperature control is open-loop to obtain an unstable system. The CSTR model consists of nonlinear state equations for material and energy balances including reaction kinetics and heat transfer. The goal of the optimization is to make a step change from a stable operating point at $T = 330K$ to an unstable operating point at $T = 353K$. The overall optimization problem results in

$$\min_{T_{c,set}(t)} \Psi = \int_0^{t_f=5} (T-353)^2 dt$$

$$\dot{c_A} = \frac{q}{V}(c_{Af} - c_A) - \Delta k_0 \exp\left(-\frac{E}{RT}\right) c_A$$

$$\dot{T} = \frac{q}{V}(T_f - T) + \frac{UA}{V\rho C_p}(T_c - T) - \frac{\Delta H k_0}{\rho C_p} \exp\left(-\frac{E}{RT}\right) c_A$$

$$\dot{T}_c = V_c \rho_c c_{p,c}(T_{in,c} - T_c) - \frac{UA}{V\rho C_p}(T_c - T)$$

$$\dot{T}_{in,c} = K\left(V_c \rho_c c_{p,c}\left(T_{in,c} - T_c\right) - \frac{UA}{V\rho C_p}(T_c - T) - \frac{\mathrm{d}}{\mathrm{d}t}T_{c,set}\right) + \frac{1}{\tau}(T_c - T_{c,set})$$

$$T_{c,set} \in [270, 350K].$$

The control $T_{c,set}$ is parameterized equidistantly by continuous piecewise linear approximations. Hence, the time derivative $\frac{\mathrm{d}}{\mathrm{d}t}T_{c,set}$ exists on $[0, t_f]$ and can be calculated analytically since the control values and the time intervals are known. The derivative is not required to be formulated as part of the ODE.

Contrary to single shooting, hybrid and multiple shooting successfully complete the desired set point change and are able to keep the reactor in the unstable region (cf. Fig. 2). However, hybrid shooting converged faster and required 44% less NLP iterations than multiple shooting. The obtained solution of both approaches are of good quality but the optimal control of multiple shooting oscillates slightly when the steady state is reached. The automatic detection, illustrated in Section 3.1, determines c_A and T as subspace $\mathcal{U}$ ($n_{\mathcal{U}} = 2$). Furthermore, the number of stages for hybrid and multiple shooting are calculated to $J = 33$ stages, by means of eq. (24) since the mean of the largest eigenvalues is $\lambda^* = 2.2873$.

Table 1 reveals that hybrid shooting and multiple shooting are equal in the objective. Hybrid shooting requires less NLP iterations and is faster than multiple shooting, although the sensitivity load of hybrid shooting (see Section 3.5) is disadvantageous. Exemplarily, we calculate the sensitivity loads for the chosen equidistant control grid of $n_c^j = 3$ points per stage as $N_{s,ms} = 8$ for multiple shooting, $N_{s,ss} = 99$ for single shooting and $N_{s,hs} = 163$ for hybrid shooting. The increased number of sensitivity systems for hybrid shooting is obviously unfavorable. However, for larger models, the sensitivity system is expected to increase only moderately for hybrid shooting, due the more advantageous ratio of unstable states to the total number of states.

5. CONCLUSIONS

This work deals with a novel optimization approach for nonlinear dynamical systems. The solution procedure commences the optimization by employing single shooting. Only if positive eigenvalues are detected the problem is reparameterized, where states leading to instability are discretized on the time horizon, whereas the other states remain continuous. Following this approach the rather small NLP of single shooting can be preserved and at the same time, the abilities of this approach are enlarged by the solution to unstable problems.

The proposed algorithm performs well for the studied example resulting in better convergence properties than for multiple shooting. Depending on the problem at hand in some cases hybrid shooting and in some cases multiple shooting will be more efficient. If the majority of the differential states belong to the unstable subspace, we expect that multiple shooting performs more efficient rather than hybrid shooting. However, for processes where only few states can be related to unstable dynamics, the sensitivity system of hybrid shooting is expected to increase only moderately.

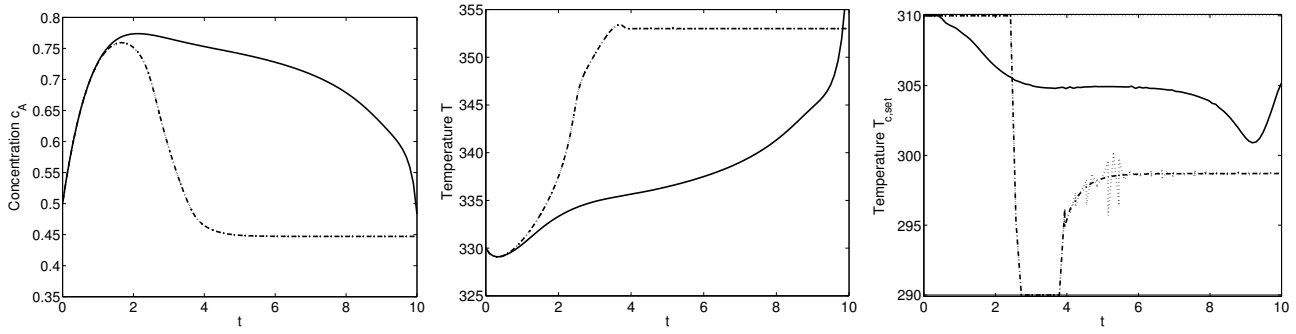


Fig. 2. Set point change of CSTR; left: concentration c_A, middle: control T, right: temperature $T_{c,set}$; SS (solid), HS (dashed) and MS (dotted); the trajectories of hybrid and multiple shooting are congruent.

Table 1. Computational statistics for CSTR

Appr.	no. of NLP iterations	decision variables n_D	CPU-time [sec]	objective Ψ	discretized variables
SS	176	99	42.28	2.73E0	–
MS	446	259	152.1	1.01E0	$c_A, T, T_c, \Psi, T_{in,c}$
HS	196	163	143.7	1.01E0	c_A, T

In the example, hybrid shooting required less NLP iterations. The authors emphasize at this point, that the chosen small examples are more or less unfavorable for hybrid shooting, since a major part of the state vector belongs to the unstable subspace.

The efficiency of the solution method can be increased by employing rSQP-methods (Leineweber et al., 2003). In another work, we have presented an algorithm for parallel sensitivity analysis (Hartwich et al., 2010), which has not yet been combined with hybrid shooting. The application of this algorithm for hybrid shooting can compensate for the increased numerical effort for sensitivities in comparison to multiple shooting.

The calculation of the eigenvalues is computationally demanding and could be replaced by an Arnoldi iteration (Arnoldi, 1951). The algorithm is able to calculate iteratively certain ranges of eigenvalues of the system matrix. Thus, not all eigenvalues have to be obtained which decreases the computational load for the detection of instabilities, while applying single shooting. This algorithm is available within the software package ARPACK (Lehoucq et al., 1997). In case one positive eigenvalue is detected, the remaining eigenvalues and eigenvectors have to be calculated.

ACKNOWLEDGEMENTS

The authors gratefully thank the German Research School, Jülich, for their financial support. The software implementation was largely supported by K. Stockmann and T. Hoffmann, RWTH Aachen University.

REFERENCES

Arnoldi, W.E. (1951). The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of Applied Mathematics*, 9, 17–29.

Ascher, U., Mattheij, R., and Russell, R. (1995). *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. SIAM Publications, Philadelphia, PA, 2nd edition.

Biegler, L. (1984). Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Comp. Chem. Eng.*, 8, 243–248.

Bock, H. and Plitt, K.J. (1984). A multiple shooting algorithm for direct solution of optimal control problems. In *Proc. of the 9th IFAC World Congress, Budapest*. Pergamon Press.

Hartwich, A. (2010). Ph.D. thesis, RWTH Aachen University, AVT Prozesstechnik. In preparation.

Hartwich, A., Stockmann, K., Terboven, C., Feuerriegel, S., and Marquardt, W. (2010). Parallel sensitivity analysis for efficient large-scale dynamic optimization. *Optimization and Engineering*. doi:10.1007/s11081-010-9104-4.

Lehoucq, R.B., Sorensen, D.C., and Yang, C. (1997). Arpack users' guide: Solution of large scale eigenvalue problems with implicitly restarted arnoldi methods.

Leineweber, D., Schäfer, A., Bock, H., and Schlöder, J. (2003). An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization part II: Software aspects and applications. *Comp. Chem. Eng.*, 27, 167–174.

Nocedal, J. and Wright, S. (1999). *Numerical Optimization*. Springer-Verlag.

Sargent, R. and Sullivan, G. (1978). The development of an efficient optimal control package. In J. Stoer (ed.), *Proceedings of the 8th IFIP Conference on Optimization Technique*, 158–168. Springer-Verlag.

Uppal, A., Ray, W., and Poore, A. (1974). On the dynamic behavior of continuous stirred tank reactors. *Chem. Eng. Sci.*, 29(4), 967–985.

Vassiliadis, V., Sargent, R., and Pantelides, C. (1994a). Solution of a class of multistage dynamic optimization problems: 1. Problems without path constraints. *Ind. Eng. Chem. Res.*, 33(9), 2111–2122.

Vassiliadis, V., Sargent, R., and Pantelides, C. (1994b). Solution of a class of multistage dynamic optimization problems: 2. Problems with path constraints. *Ind. Eng. Chem. Res.*, 33(9), 2123–2133.

[1] The optimization was halted manually.