Obscured by terminology: Hidden parallels in direct methods for open-loop optimal control

Susanne Sass[1] and Alexander Mitsos[2,1,3,*]

[1] Process Systems Engineering (AVT.SVT), RWTH Aachen University, 52074 Aachen, Germany
[2] JARA-CSD, 52056 Aachen, Germany
[3] Institute of Climate and Energy Systems: Energy Systems Engineering (ICE-1), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Abstract: Active research on optimal control methods comprises the developments of research groups from various fields, including control, mathematics, and process systems engineering. Although there is a consensus on the classification of the main solution methods, different terms are often used for the same method. For example, solving optimal control problems with control discretization and embedded state integration may be called sequential method or direct single shooting. Equally severely, the same term may be used ambiguously: Is control vector parameterization a synonym for control discretization or for direct single shooting? Both misleading distinctions and ambiguity complicate the scientific discourse. Thus, we delineate standard terms from open-loop optimal control in this tutorial. More precisely, we formulate and challenge hypotheses on the terminology of direct methods, i.e., solution methods using control discretization combined with state integration and/or state discretization. In particular, we point out the parallel of the embedded state integration with a numerical integration scheme and the reduced-space formulation of approaches using state discretization. As another parallel between two apparently distinct methods, we investigate the similarities and differences between the discrete-time solution of optimal control problems and optimal quasi-steady operation. In this context, we also hint on the discrete-time representation in scheduling which refers to the handling of controls rather than the handling of process dynamics. This tutorial concludes with recommendations on how to avoid misunderstandings in the versatile research community.

Keywords: *dynamic optimization; control vector parametrization; collocation; single shooting; multiple shooting; steady-state assumption*

1 Introduction

The dynamic behavior of processes where the future state of the system depends on the current state and control values is typically modeled with differential equations. Consequently, dynamic optimization addresses optimization problems with differential or differential-algebraic equations (DAEs) embedded. We, in particular, focus on dynamic optimization problems determining control *functions*. Such infinite-dimensional dynamic optimization problems are also known as optimal control problems.

In our context, both terms "dynamic optimization problem" [e.g., 4, 5, 53] and "optimal control problem" [e.g., 10, 18, 28, 34, 70, 90] are ambiguous: often, "dynamic optimization" is also used as an umbrella term including both finite dimensional problems like dynamic parameter estimation problems and infinite dimensional optimal control problems [e.g., 12, 79]. Similarly, optimization with differential equations or DAEs embedded may comprise both finite and infinite dimensional problems. However, "optimal control" may refer to closed-loop control or model-predictive control, i.e., optimization problems with a strong focus on online applications and, typically, least-squares objectives [e.g., 73]. In this tutorial, we focus on the offline optimization of open-loop optimal control problems. In other words, our hidden conjecture is that the control functions, i.e., the system inputs which can be manipulated from a technical or human controller, depend on time but not explicitly on feedback from the system states.

Throughout this tutorial, we use an example rather than a general optimal control problem for our discussions. Without loss of generality, we refer to the optimal control problem

$$\begin{aligned} \min_{u(\cdot)} \quad & \int_0^{t_f} \mathcal{L}(x(t))\,\mathrm{d}t \\ \text{s.t.} \quad & \frac{\mathrm{d}x}{\mathrm{d}t}(t) = \frac{u(t)}{x(t)} \quad \forall t \in [0, t_f] \\ & x(0) = x_0 \\ & x(t_f) = x_f \end{aligned} \tag{OCP}$$

with infinitely many optimization variables $u(t),\ t \in [0, t_f]$, a Lagrange-type objective function, a nonlinear but control-affine ordinary differential equation (ODE), an explicitly known initial value x_0 and

Corresponding author: *A. Mitsos
AVT Process Systems Engineering, RWTH Aachen University, 52074 Aachen, Germany
E-mail: amitsos@alum.mit.edu

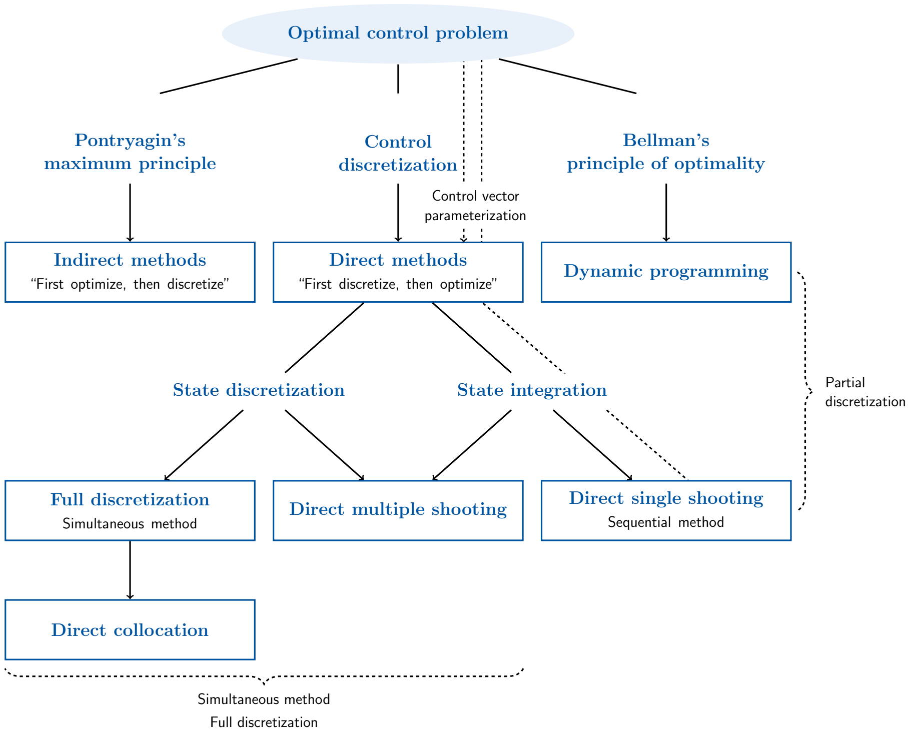

Figure 1: Overview on methods for solving optimal control problems and their common terminology. The bold, colored terms are used throughout this tutorial.

an endpoint constraint. The feasible control domain $u(t) \in [u^{\min}, u^{\max}],\ t \in [0, t_f]$ and the initial value x_0 are chosen such that we can ensure $x(t) > 0$ for any $t \in [0, t_f]$. Together with essentially bounded controls $u \in L^\infty([0, t_f])$, we obtain a well-posed optimization problem. Note that (OCP) is designed to be as simple as possible for showcasing hidden parallels between direct methods.

Optimal control problems like (OCP) can essentially be solved with three principles [e.g., 5, 10, 13, 15], see Figure 1: Firstly, Pontryagin's maximum principle gives indirect methods [71]. Secondly, control discretization yields direct methods. Thirdly, Bellman's principle of optimality results in dynamic programming [9, 61]. Note that indirect and direct methods are also called "First optimize, then discretize" and "First discretize, then optimize" methods, respectively [e.g., 80]. However, "First optimize, then discretize" suggests that the boundary value problem obtained by Pontryagin's maximum principle (*optimize*) is solved with discretization methods. In principle, we may determine the control *function* via control vector iteration [e.g., 60] or, for comparatively simple problems, the analytical solution. In other words, it is a hidden conjecture that discretization is required for the solution of optimal control problems.

In the following, we formulate and challenge hidden conjectures for standard direct methods as hypotheses. We validate or disprove (parts of) each hypothesis based on the literature as well as reformulations showcased with (OCP). In line with the objectives of a tutorial, we recall the standard approaches of direct methods depicted in the bottom of Figure 1 based on (OCP) right before the related discussions. We may conjecture that an external integration software provides solution trajectories for the continuous-time differential equations, while numerical integration schemes with a fixed step size yield discrete-time solutions. Consequently, we also study the usage of "discrete-time" and "continuous-time" as well as the parallels and differences between discrete-time solutions and quasi-steady operation. A detailed examination of control methods apart from direct methods for offline open-loop optimal control as well as of solution methods for DAEs, partial differential equations, boundary value problems, or mixed-integer problems lies beyond the scope of this tutorial.

The remainder of this article is structured as follows. In Section 2, we investigate hidden conjectures on direct methods for the solution of optimal control problems like (OCP). In particular, we focus on methods using state discretization in Section 2.1, methods using state integration in Section 2.2, and methods combining both of these approaches in Section 2.3. In Section 3, we link and compare these methods with approaches postulating quasi-steady behavior. Finally, we summarize and formalize our conclusions in the form of recommendations for (mostly) unambiguous terminology in Section 4.

2 Hidden conjectures on direct methods

In all direct methods, we start with *discretizing* the infinite-dimensional function space $L^\infty([0,t_f])$ which hosts the control function $u(\cdot)$ into a finite-dimensional vector space $\mathbb{R}^{n_u}$ which hosts the control coefficients $\hat{\boldsymbol{u}} = (\hat{u}_i)_{i=1,\dots,n_u}$. This is also known as *transcription* in the context of trajectory optimization [10]. In detail, we choose finitely many basis functions $\varphi_i : [0,t_f] \to \mathbb{R},\ i = 0,\dots,n_u-1$ and *parameterize* the control function via $u(t) \approx \sum_{i=0}^{n_u-1} \hat{u}_i \cdot \varphi_i(t),\ t \in [0,t_f]$. As an example, we exploit the indicator function $\mathbf{1}_T : [0,t_f] \to \{0,1\}$ with $\mathbf{1}_T(t) = 1$ if $t \in T$ and $\mathbf{1}_T(t) = 0$ otherwise for defining piecewise constant controls.

Approach 1 (Control discretization)**.** *For solving* (OCP) *with direct methods, we use piecewise constant controls given by specific basis functions* $\varphi_i(\cdot) = \mathbf{1}_{[t_i,t_{i+1})}(\cdot)$ *for any* $i = 1,\dots,n_u$ *defined on subintervals given by time discretization* $0 = t_0 < t_1 < \dots < t_{n_u} = t_f$ *which yields the discretized optimal control problem*

$$\begin{aligned} \min_{\hat{\boldsymbol{u}}} \quad & \int_0^{t_f} \mathcal{L}(x(t))\ dt \\ s.t. \quad & \frac{\mathrm{d}x}{\mathrm{d}t}(t) = \frac{\hat{u}_i}{x(t)} && \forall t \in [t_i,t_{i+1})\ \forall i = 0,\dots,n_u-1 \\ & x(0) = x_0 \\ & x(t_f) = x_f && . \end{aligned} \tag{discOCP}$$

Note that piecewise constant discretization is often used without an explicit function rule for the control values $u(t)$ in the interior of the subintervals $t \in (t_i,t_{i+1})$. Another common discretization is the choice of piecewise polynomials [e.g., 13, Sec. 10]. Note further that the solution obtained by direct methods is, in general, only optimal for the discretized problem (discOCP). Some optimization algorithms guarantee convergence to the solution of (OCP) by an adaptive control and/or state discretization [e.g., 11, 48].

Direct methods mainly use two approaches for solving the differential equations embedded: a-priori state discretization and state integration, cf. Figure 1. In the following subsections, we therefore successively focus on approaches using state discretization, state integration, or both. In particular, we argue that state integration may be performed with an external integration software or an embedded numerical integration scheme hidden from the optimizer, see Hypothesis 6 and Section 2.3.

2.1 Distinct terminology for approaches using state discretization

When using state discretization, we determine finitely many state values which are the solution of the differential equations or the related boundary value problem. Collocation [39, 55] is a standard approach for the solution of boundary value problems where the states are interpolation polynomials on subintervals $[t_i,t_{i+1}],\ i = 0,\dots,n$, the so-called collocation elements. The differential equations and the boundary values are enforced at the n_C collocation points $c_j \in [0,1],\ j = 1,\dots,n_C$, i.e., at time points $t_i + c_j \cdot h_i$. Alternative choices for the interpolation polynomials and collocation points result in specific collocation approaches like trapezoidal collocation and Hermite-Simpson collocation [e.g., 10, 43, 54]. In *orthogonal* collocation, the collocation points are the roots of orthogonal polynomials, e.g., Legendre or Chebyshev polynomials [92]. Since direct methods using state discretization ("full discretization") typically apply collocation, we formulate the hidden conjecture:

Hypothesis 1. *"Full discretization" and "direct collocation" refer to the same method.*

Already in the late 1960's, Tabak and Kuo proposed to discretize both controls and states where the discretized state values satisfy the differential equations according to a numerical integration scheme [90]. For example, we may apply the explicit Euler scheme.

Approach 2 (Full discretization with the explicit Euler)**.** *We discretize the states of* (discOCP) *according to the explicit Euler scheme with step width* $h_i := t_{i+1} - t_i,\ i = 0,\dots,n_u-1$ *and obtain*

$$\begin{aligned} \min_{\hat{\boldsymbol{u}},\hat{\boldsymbol{x}}} \quad & \sum_{i=0}^{n_u} \hat{\mathcal{L}}(\hat{x}_i) \\ s.t. \quad & \hat{x}_{i+1} = \hat{x}_i + h_i \cdot \frac{\hat{u}_i}{\hat{x}_i} && \forall i = 0,\dots,n_u-1 \\ & \hat{x}_0 = x_0 \\ & \hat{x}_{n_u} = x_f && . \end{aligned} \tag{FullDisc:Euler}$$

Subsequently, Reddien and Biegler et al. proposed to use orthogonal collocation for an accurate and efficient state discretization [12, 32, 51, 75].

As an example, we interpolate the first derivative of the state based on one collocation point c_1 and, thus, with constant interpolation polynomials. With this, the differential equation of (OCP) reads as

$$x(t_i + \tau \cdot h_i) = x(t_i) + h_i \cdot \frac{u(t_i + c_1 \cdot h_i)}{x(t_i + c_1 \cdot h_i)} \cdot \int_0^\tau 1 \, d\sigma \quad \forall i = 0, \dots, n_u - 1$$

resulting in the following fully discretized optimal control problem.

Approach 3 (Direct collocation)**.** *We discretize the states of* (discOCP) *according to collocation with the single collocation point* $c_1 = 0$ *and obtain*

$$\begin{aligned} \min_{\hat{\boldsymbol{u}},\hat{\boldsymbol{x}}} \quad & \int_0^{t_f} \mathcal{L}(x(t))\,dt \\ s.t. \quad x(t_i + \tau h_i) &= \hat{x}_i + \tau \cdot h_i \cdot \frac{\hat{u}_i}{\hat{x}_i} && \forall \tau \in [0,1]\ \forall i = 0,\dots,n_u - 1 \\ \hat{x}_{i+1} &= x(t_i + h_i) && \forall i = 0,\dots,n_u - 1 \\ \hat{x}_0 &= x_0 \\ \hat{x}_{n_u} &= x_f && . \end{aligned} \tag{Coll}$$

In comparison, (FullDisc:Euler) results in the same state values $\hat{x}_i = x(t_i)$ at the grid points but does not explicitly define state values $x(t_i + \tau \cdot h_i)$ at the interior of the collocation elements $\tau \in (0, 1)$. Note that *direct* indicates that we are solving the problem with discretized controls, i.e., using a direct method, compared to solving the boundary value problem obtained from Pontryagin's maximum principle, i.e., using an indirect method, cf. Figure 1.

Conclusion 1. Direct collocation is a specific variant of the full discretization approach. Since (orthogonal) collocation is the state-of-the-art approach for handling state discretization, "direct collocation" is sometimes equated with the full discretization approach [e.g., 34, 79, 86]. □

In the literature, we find further synonyms for full discretization:

Hypothesis 2. *"Full discretization" and "simultaneous method" refer to the same method.*

Literally, both states and controls are discretized by "full discretization" and simultaneously optimized by the "simultaneous method". We can imagine to decouple these two approaches: we discretize both states and controls (*full discretization*) but evaluate the states sequentially while optimizing only the controls (*sequential method*), see Hypotheses 3 and 7. However, the standard approach is to use a combination, namely to optimize an algebraic optimization problem where the differential equations are reformulated as equality constraints (*simultaneous method*) based on discretized states (*full discretization*), see Approaches 2 and 3. "Full discretization", "simultaneous method", and "all-at-once method" are therefore often used synonymously [e.g., 13, 15, 79].

In direct multiple shooting [20], the states are partially discretized and optimized simultaneously to the discretization and optimization of controls, see Approach 5. Thus, direct multiple shooting may also be classified as "simultaneous method" or "full discretization method" [e.g., 15, 34], although it is more often regarded as hybrid method, see Section 2.3.

Conclusion 2. The use of "full discretization" and "simultaneous method" is ambiguous: On the one hand, both terms may synonymously refer to a simultaneous full discretization approach. On the other hand, either of the terms may be used as umbrella term for the other as well as hybrid approaches like direct multiple shooting. □

In this tutorial, we use the term "full discretization" for the specific direct method to stress the usage of state discretization as opposed to state integration, see Figure 1.

2.2 Distinct terminology for approaches based on state integration

Apart from collocation, single shooting is a common approach for the solution of boundary value problems [e.g., 89]. When combining single shooting with control discretization, we obtain a direct method for optimal control where the integration is hidden in the evaluation of a black-box function.

Approach 4 (Direct single shooting)**.** *We assume that function* $F^{int}(\cdot;\hat{\boldsymbol{u}},x_0):[0,t_f]\to\mathbb{R}$ *evaluates the solution* $x(\cdot)$ *of the differential equation of* (discOCP) *with control values* $\hat{\boldsymbol{u}}$ *initialized at* $x(0)=x_0$*. Thus, we can replace the differential equation with a function evaluation and obtain*

$$\begin{aligned}&\min_{\hat{\boldsymbol{u}}} \int_0^{t_f} \mathcal{L}\big(F^{int}(t;\hat{\boldsymbol{u}},x_0)\big)\ dt\\&s.t.\ \ F^{int}(t_f;\hat{\boldsymbol{u}},x_0)=x_f\end{aligned} \quad . \tag{SiSh}$$

Since the state integration is embedded in the optimization when using direct single shooting, we may conjecture:

Hypothesis 3. *"Direct single shooting" and "sequential method" refer to the same method.*

Literally, "sequential method" comprises all methods where state integration and optimization of controls is performed sequentially such that the states are dependent variables. Apart from direct single shooting, see Approach 4, this is also true for hybrid methods like direct multiple shooting where all state values except those at the grid points are dependent variables, see Approach 5. However, in the literature, direct multiple shooting is classified as a hybrid or simultaneous method, compare Hypothesis 2, while direct single shooting is called a (pure) sequential method [e.g., 5, 16]. In parts of the literature, direct methods using embedded state integration, which we call "direct single shooting", are explicitly introduced as "sequential methods" [e.g., 15].

Conclusion 3. Indeed, "direct single shooting" and "sequential method" refer to the same method in the literature. □

Since the controls but not the states are discretized in direct single shooting, we may also conjecture:

Hypothesis 4. *"Direct single shooting" and "partial discretization" refer to the same method.*

While "full discretization" refers to the discretization of both controls and states, "partial discretization" is an umbrella term indicating any approach solely discretizing the controls. This includes, e.g., direct single shooting and a dynamic programming approach [15]. However, "full discretization" and "partial discretization" may also be used for the opposed direct methods applying state discretization and state integration, respectively.

Conclusion 4. The use of the term "partial discretization" is ambiguous: it may comprise or equate direct single shooting. □

In large parts of the literature, the counterpart of full discretization is called "control vector parameterization", leading to the following hidden conjecture:

Hypothesis 5. *"Direct single shooting" and "control vector parameterization" refer to the same method.*

In fact, the usage of "control vector parameterization" is ambiguous. Literally, "control vector parameterization" means to parametrize the control vector, i.e., to discretize the controls. Control discretization was originally proposed for the solution of boundary value problems [77]. Since the late 1960's and early 1970's, it has also been increasingly used for optimal control, first by applying it to the solution of the boundary value problems obtained from indirect methods [50, 56, 78, 87]. Subsequently, control discretization was combined with embedded state integration giving rise to a novel direct method [45, 70, 94]. With these and further advancements, control discretization with embedded state integration became a capable state-of-the method and was shortly referred to as "control vector parameterization".

Conclusion 5. Some groups [e.g., 14, 36, 81] use "control vector parameterization" as a synonym for direct single shooting, while others [e.g., 20] use the term literally, i.e., as a synonym for control discretization. □

Either way, the embedded state integration is often related to an external integrator:

Hypothesis 6. *In direct shooting methods, the differential equations are solved with an external integration software.*

Nowadays, direct shooting methods apply external ODE or DAE solvers which provide accurate solutions of the differential equations by limiting the integration error via adaptive grid refinement and the like [e.g., 46, 69, 85]. However, when control discretization with embedded state integration was proposed for optimal control, there had hardly been any capable ODE or DAE solvers, yet. Still, the differential equations were not treated as equality constraints to be handled by the optimizer but hidden within the optimizer's function evaluations.

Originally, integration and optimization were decoupled by directly including the sensitivity equations of the states within the optimization problem [e.g., 70, 94]. Subsequently, state integration was included within the evaluation of the objective function, e.g., in combination with derivative-free optimization methods or penalty formulations [e.g., 28, 35, 45]. A further advancement was the proposal of forward and backward differentiation based on a numerical integration scheme for the state integration embedded in the optimization of the controls [81]. Only then, capable solvers were available for integrating the states externally [e.g., 18, 91].

Conclusion 6. In direct shooting methods, using an external integration software for the embedded state integration is the standard. This standard allows for exploiting all recent advancements for the efficient and accurate solution of differential equations. However, judging from the origins of direct single shooting, there are also alternative methods for the embedded state integration like the use of sensitivity equations or numerical integration schemes. □

The use of numerical integration schemes points towards a hidden parallel between direct collocation and direct single shooting.

Hypothesis 7. *Direct single shooting handles the differential equations substantially differently compared to the full discretization method.*

In the standard approaches, the optimizer implicitly handles the integration by calling an external integration software like a black-box function F^{int} within direct single shooting, see (SiSh), while it explicitly handles an algebraic equation system given by orthogonal collocation within full discretization, see (FullDisc:Euler). Following our conclusions for Hypothesis 6, we may also include the evaluation of the numerical integration scheme hidden in F^{int} in the optimization problem.

As an example, we use an explicit Euler scheme on the time grid used for discretizing the controls. With this, (SiSh) transforms into

$$\begin{aligned}
\min_{\hat{\boldsymbol{u}}} \quad & \sum_{i=0}^{n_u} \hat{\mathcal{L}}(\hat{x}_i) \\
\text{s.t.} \quad & \hat{x}_{i+1} := \hat{x}_i + h_i \cdot \frac{\hat{u}_i}{\hat{x}_i} && \forall i = 0, \ldots, n_u - 1 \\
& \hat{x}_0 := x_0 \\
& \hat{x}_{n_u} = x_f && .
\end{aligned} \tag{SiSh:Euler}$$

Note that we distinguish between equality constraints "=" handled by the optimizer, see the endpoint constraint, and assignments ":=", see the numerical integration scheme. The assignments refer to simple function evaluations triggered by the optimizer. In particular, the assignments can be substituted into the objective and constraints, cf. (SiSh), when using an explicit numerical integration scheme. Thus, (SiSh:Euler) is a reduced-space formulation of (FullDisc:Euler): we reduce the number of optimization variables by identifying and determining the states as dependent variables.

If the objective function or constraints require knowledge of intermediate state values, we can add an interpolation scheme for the states. A piecewise linear interpolation connecting the state values determined based on the explicit Euler scheme yields

$$\begin{aligned}
\min_{\hat{\boldsymbol{u}}} \quad & \int_0^{t_f} \mathcal{L}(x(t))\mathrm{d}t \\
\text{s.t.} \quad & x(t_i + \tau h_i) := \hat{x}_i + \tau \cdot h_i \cdot \frac{\hat{u}_i}{\hat{x}_i} && \forall \tau \in [0,1] \ \forall i = 0, \ldots, n_u - 1 \\
& \hat{x}_{i+1} := x(t_i + h_i) && \forall i = 0, \ldots, n_u - 1 \\
& \hat{x}_0 := x_0 \\
& \hat{x}_{n_u} = x_f
\end{aligned} \tag{SiSh:interpol}$$

which is the reduced-space formulation of (Coll).

Reduced-space formulations [23, 65] are similar to decomposition approaches: the structure of the model equations is exploited for reducing the dimensionality of the optimization problem. However, the dimensionality reduction is performed during modeling with reduced-space formulations and within the optimization algorithm with decomposition algorithms [22, Sec. 3.4.3]. The parallel between direct collocation combined with such reduction approaches and direct single shooting has been already hinted in the 1990's:

> *"In fact, the use of decomposition techniques for the solution of the large NLP arising from the complete discretization approach leads to methods that are, in general, very similar to the control parameterization approach. The main remaining difference between the two approaches is the precise method used for approximating the infinite dimensional DAEs by a finite number of constraints. It can readily be shown that the collocation methods on finite elements (that are primarily used in the context of the complete discretization approach) are equivalent to fully implicit Runge-Kutta single-step integration methods. On the other hand, most implementations of control vector parametrization rely on multistep backward-difference formula (BDF) integration methods (Gear, 1971). However, in principle, there is no reason why either integration method could not be used in either approach although one would expect multistep methods to lead to more efficient (i.e., smaller) approximations especially at high accuracies."* (Vassiliadis, Sargent, and Pantelides 1994 [91], p. 2114)

The equivalency of specific collocation and Runge-Kutta methods has been shown around 1970 [42, 95]. For the proof, the first derivative of the states are interpolated with Lagrange polynomials. Consequently, the right-hand side of the explicit ODE equals the interpolation polynomial which gives the well-known formulas for the Runge-Kutta coefficients based on Lagrange polynomials defined with the collocation points [e.g., 27]. Still, collocation is slightly distinct from Runge-Kutta methods:

> *"The only difference is that in collocation one works in terms of polynomial coefficients and in implicit Runge-Kutta in terms of values of the function f."* (Wright 1970 [95], p. 219)

We may close this gap with a natural continuous extension [96] which basically defines an interpolation scheme for obtaining a continuous-time trajectory based on the determined state values. As an example, compare the linear interpolation used in (Coll) and (SiSh:interpol).

Conclusion 7. Briefly speaking, the state-of-the-art approaches of direct single shooting and full discretization are substantially different since these methods follow different paradigms: the former exploits the capabilities of state-of-the-art ODE or DAE solvers for an efficient integration, while the latter formulates an algebraic optimization problem and exploits its structure for an efficient optimization. Both paradigms intersect, in particular, when using Runge-Kutta methods within direct single shooting and collocation within the full discretization approach. At this point of intersection, the differences between the two approaches come down to the comparison of reduced-space and full-space formulations. □

In the following section, we investigate reduced-space formulations and other hybrid methods combining state discretization and state integration in more detail.

2.3 Multiple shooting and other hybrid methods

Direct multiple shooting is commonly referred to as hybrid method [e.g., 3, 5, 16, 53] since it combines both (partial) state discretization and state integration, see the corners of Figure 2. More precisely, state integration is performed independently on subintervals, the so-called *shooting intervals*. For this, *shooting variables* are introduced which serve as initial values for the integration. As for single shooting, see Approach 4, the integration is typically performed by an external solver, see Hypothesis 6. Closing constraints enforce the equality of the shooting variables with the last state value obtained in the previous shooting intervals such that we obtain continuous state trajectories.

As an example, we use the time grid used for the control discretization to define the shooting intervals.

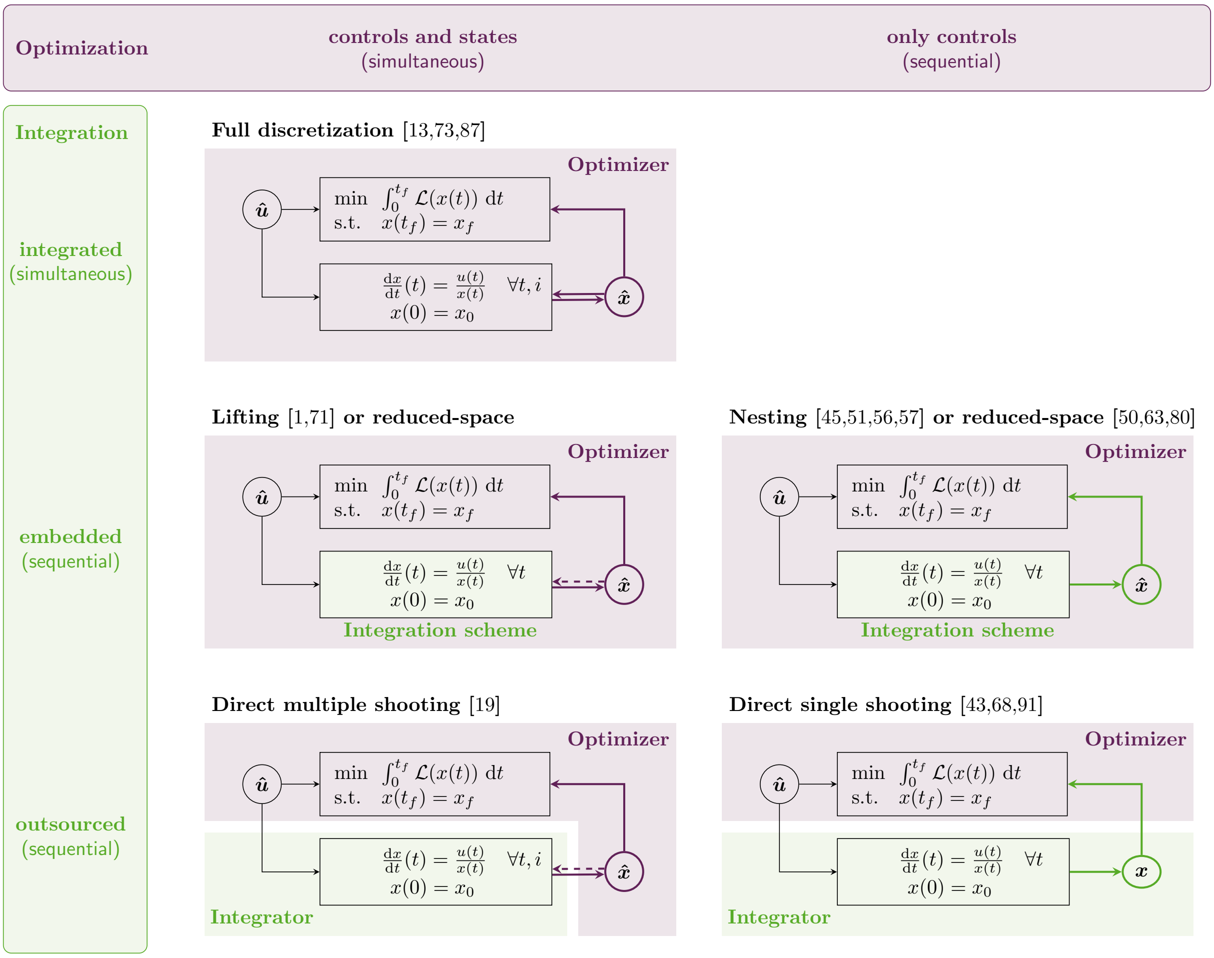

Figure 2: Schematic overview of standard and hybrid methods for solving optimal control problems like (OCP). Such optimal control problems extend an algebraic optimization problem (upper box) by differential equations (lower box). From top to bottom, the handling of the differential equations changes, namely from transforming it into constraints handled by the optimizer ( ) via using an integration scheme hidden from the optimizer ( ) to using an external integration software ( ). Arrows from controls and/or states to the model boxes indicate optimization variables. In the schemes on the left, both controls and states are optimized after applying state discretization (⇆). The dashed arrow indicates that only a part of the state values are optimized (⇠⇢). In the schemes on the right, the state values are dependent variables since state integration is used (→).

Approach 5 (Direct multiple shooting)**.** *We assume that function $F^{int}(\cdot;\hat{u}_i,s_i):[t_i,t_{i+1}]\to\mathbb{R}$ evaluates the solution $x_i(\cdot)$ of the differential equation of* (discOCP) *with fixed control value $\hat{u}_i$ initialized at $x_i(t_i)=s_i$. Thus, we can replace the differential equation within each subinterval with a function evaluation and obtain*

$$
\begin{aligned}
\min_{\hat{\boldsymbol{u}},\boldsymbol{s}} \quad & \int_0^{t_f} \mathcal{L}\big(x(t)\big)\ dt \\
s.t.\ x(t) &:= \sum_{i=0}^{n_u-1} \mathbf{1}_{[t_i,t_{i+1})}(t)\cdot x_i(t) && \forall i=0,\dots,n_u-1 \\
x_i(t) &:= F^{int}(t;\hat{u}_i,s_i) && \forall t\in[t_i,t_{i+1}]\ \forall i=0,\dots,n_u-1 \\
x_i(t_i) &:= s_i && \forall i=0,\dots,n_u-1 \\
s_0 &:= x_0 \\
x_i(t_{i+1}) &= s_{i+1} && \forall i=0,\dots,n_u-1 \\
s_{n_u} &:= x_f && .
\end{aligned}
\tag{MuSh}
$$

Analogously to (SiSh:Euler), we can replace black-box function $F^{\text{int}}(\cdot)$ with an integration scheme which is hidden from the optimizer as depicted in the middle row in Figure 2. When using the explicit Euler scheme with step width $h_{i,j}:=t_{i,j+1}-t_{i,j}$ defined for intermediate time points $t_i=t_{i,0}<t_{i,1}<\ldots<t_{i,n_i}=t_{i+1},\ i=0,\dots,n_u-1$, (MuSh) transforms into

$$
\begin{aligned}
\min_{\hat{\boldsymbol{u}},\boldsymbol{s}} \quad & \sum_{i=0}^{n_u}\sum_{j=0}^{n_i}\hat{\mathcal{L}}(\hat{x}_{i,j}) \\
\text{s.t.}\quad \hat{x}_{i,j+1} &:= \hat{x}_{i,j}+h_{i,j}\cdot\frac{\hat{u}_i}{\hat{x}_{i,j}} && \forall j=0,\dots,n_i-1\ \forall i=0,\dots,n_u-1 \\
\hat{x}_{i,0} &:= s_i && \forall i=0,\dots,n_u-1 \\
s_0 &:= x_0 \\
\hat{x}_{i,n_i} &= s_{i+1} && \forall i=0,\dots,n_u-1 \\
s_{n_u} &:= x_f && .
\end{aligned}
\tag{MuSh:Euler}
$$

From direct single shooting via direct multiple shooting to full discretization, the number of optimization variables and the sparsity of the model equations increases whereas the sensitivity to integration errors or unstable differential equations decreases [11, 13–16, 20, 30]. For stabilizing the optimization, specialized integrators can be used, e.g., for handling stiff differential equations, within direct shooting methods. The decision whether to use state integration and/or a reduced-space formulation affects also the accessibility of intermediate state values for imposing constraints or for exploiting knowledge on the solution, e.g., state values at specific time points. The optimizer can read and change all state values for full discretization but only the values for the shooting variables in standard direct multiple shooting. With standard direct single shooting, the optimizer can only call for state values. When using reduced-space formulations, i.e., direct shooting methods where the state values are assigned via integration schemes, the optimizer can easily impose constraints on the intermediate state values although it can adapt at most the shooting variables.

Compared to (SiSh:Euler), we may interpret (MuSh:Euler) as a partial reduced-space formulation: a subset of the state values is assigned while the rest, the shooting variables, are optimization variables. Similarly, we can design different reduced-space formulations for handling the objective function. In fact, we can replace the Lagrange-type objective $\int_0^{t_f}\mathcal{L}(x(t))\ dt$ with a Mayer-type objective $y(t_f)$, where $\frac{\mathrm{d}y}{\mathrm{d}t}(t)=\mathcal{L}\left(x(t)\right),\ t\in[0,t_f]$ and $y(0)=\mathcal{L}\left(x(0)\right)$. This additional differential equations may be handled equivalently or differently to the other differential equations. In particular, we obtain a partial reduced-space formulation when reformulating a quadrature formula for the numerical integration of the Lagrange-type objective within full discretization, see (FullDisc:Euler), as assignments for evaluating a Mayer-type

(a) Continuous-time system (b) Discrete-time system

Figure 3: Schematic overview of time representation in scheduling problems.

objective

$$
\begin{aligned}
\min_{\hat{\boldsymbol{u}},\hat{\boldsymbol{x}}} \quad & y_{n_u} & \\
\text{s.t.} \quad & y_{i+1} := y_i + h_i \cdot \mathcal{L}(\hat{x}_i) & \forall i = 0, \dots, n_u - 1 \\
& y_0 := \mathcal{L}(\hat{x}_0) & \\
& \hat{x}_{i+1} = \hat{x}_i + h_i \cdot \frac{\hat{u}_i}{\hat{x}_i} & \forall i = 0, \dots, n_u - 1 \\
& \hat{x}_0 = x_0 & \\
& \hat{x}_{n_u} = x_f & ,
\end{aligned}
$$

where we observe $\hat{\mathcal{L}}(\hat{x}_0) = \mathcal{L}(\hat{x}_0)$ and $\hat{\mathcal{L}}(\hat{x}_i) = h_i \cdot \mathcal{L}(\hat{x}_i)$ for any $i = 1, \dots, n_u$.

Direct single or multiple shooting have been combined with full discretization approaches to obtain more efficient solution methods. Similar to the assignments used in the reduced-space formulation [e.g., 52, 65, 83], state integration may be nested within the optimization by applying collocation [47, 58, 59] or the implicit Euler scheme [53] in each iteration of the optimizer, see the middle right scheme in Figure 2. In the same way, the state integration within direct multiple shooting may be performed with collocation which, in turn, is hidden from the optimizer with the help of the lifted Newton algorithm [1, 73], see the middle left scheme in Figure 2.

Overall, full discretization, direct single shooting, and hybrid methods like direct multiple shooting differ in the handling of optimization variables and constraints. However, all of these methods may use the same approximate solution of the embedded differential equations no matter whether the integration is integrated or embedded within the optimization or outsourced to an external integration software, see also Hypotheses 6 and 7.

We may conjecture that an external integrator gives an accurate solution for the continuous-time differential equations, whereas a numerical integration scheme embedded in the optimization problem gives only an approximation, namely the solution of the related discrete-time difference equations. However, external integration software are just implementations of numerical integration schemes. Either way, we obtain a discrete-time trajectory. We therefore investigate further implications for discrete-time solutions in the following section.

3 Hidden parallels between discrete-time dynamics and quasi-steady operation

We clarify the terminology before we compare discrete-time and quasi-steady solutions. With direct methods in mind, we formulate a first hidden conjecture:

Hypothesis 8. *A "discrete-time system" is obtained by discretizing differential equations.*

Commonly, differential equations are solved by transforming them into algebraic equations with the help of time discretization and a numerical integration scheme. In control theory, the resulting equation system is called "discrete-time model" [e.g., 67, p. 107]. In particular for linear ODE's, we can exploit the knowledge of the analytical solution to obtain an equivalent reformulation by the discretization.

In the context of scheduling problems, we also distinguish between discrete-time and continuous-time systems [e.g., 38, 62, 97]. Scheduling problems aim at the optimal timing of processes including the determination of starting and end time points. When using continuous-time systems, these time points can take any real value in the time horizon $[t_0, t_f]$, see Figure 3a. When using discrete-time systems, these time points are chosen from a set of a-priori given time points t_i, $i = 0, \dots, n$ with $t_n = t_f$, see Figure 3b. In fact, the time representation, namely the choice between continuous-time or discrete-time systems, refers to the handling of the scheduling controls rather than the handling of differential equations potentially embedded.

We interpret scheduling problems as optimal control problems: a control variable decides whether to run the process. When solving this optimal control problem with direct methods, in particular, a control discretized on a fixed time grid, we obtain a discrete-time system in terms of scheduling. This is still true if the underlying process is characterized by differential equations which we solve with the analytical continuous-time solution.

Conclusion 8. "Discrete-time" is used differently in different fields of research. While it refers to the handling of differential equations in control theory, it refers to the handling of (scheduling) controls in scheduling. □

For the term "quasi-steady", we first check whether there is a distinction to a similar term:

Hypothesis 9. *"Quasi-steady" and "pseudo-steady" are synonyms.*

Indeed, there is no universal distinction between "pseudo-steady" and "quasi-steady" in the literature. As an example, the steady-state approximations for chemical reaction kinetics [21, 26, 63] is called quasi-steady state assumption (QSSA) [24, 74, 88], pseudo-steady state hypothesis (pssh) [44, 88] or, by the International Union of Pure and Applied Chemistry (IUPAC), steady-state approximation [66]. Note also that in these established terms, the words "assumption", "hypothesis" and "approximation" are used with the same meaning, namely for an approximation whose validity is rarely checked but which typically results in a small error. Still, these terms as well as "quasi-steady" and "pseudo-steady" may refer to distinct terms.

Conclusion 9. In the literature, "quasi-steady" and "pseudo-steady" are mainly treated as synonyms. □

The term "quasi-steady" itself is ambiguous. We conjecture:

Hypothesis 10. *"Quasi-steady" operation refers to a time series of steady-state operating points.*

In the literature, we primarily find two different usages of the term "quasi-steady". On the one hand, "quasi-steady" refers to almost steady-state processes, namely if

(i) a state subject to a dynamic process is (approximately) constant in time [e.g., 49, 57];
Example: The fluid level in a tank is constant if the inflow equals the outflow.

(ii) the rate of change is small compared to other changes in time [e.g., 17, 21];
Example: The concentration of a highly reactive intermediate substance has a negligible rate of change compared to less reactive products of chemical reactions.

(iii) states describe (almost) time invariant properties [e.g., 37]
Example: The amplitude of a time dependent harmonic oscillation is constant.

(iv) the steady state is temporary [e.g., 2].
Example: Between two transient phases, the system remains in a steady state for a specific time period.

On the other hand, "quasi-steady" refers to piece-wise constant state trajectories. This includes both trajectories switching between a fixed number of modes [e.g., 31, 98, 99] and trajectories with continuous-valued operating points [e.g., 41, 84, 93]. In the former, the optimizer decides which of the discrete operation modes is chosen at any point in time, while it can choose any real value within given bounds for the current operating point in the latter.

Conclusion 10. States are called "quasi-steady" if they are almost constant for a specific time period and, in particular, if they are governed by piece-wise constant operating points. □

In the following, quasi-steady operation refers to a time series of (almost) steady-state operating points. This description is also valid for discrete-time solutions of differential equations obtained with the help of numerical integration schemes. Thus, we challenge the following hidden conjecture:

Hypothesis 11. *Quasi-steady operation and discrete-time dynamics are distinct approaches for approximating the continuous-time dynamics embedded in optimal control problems.*

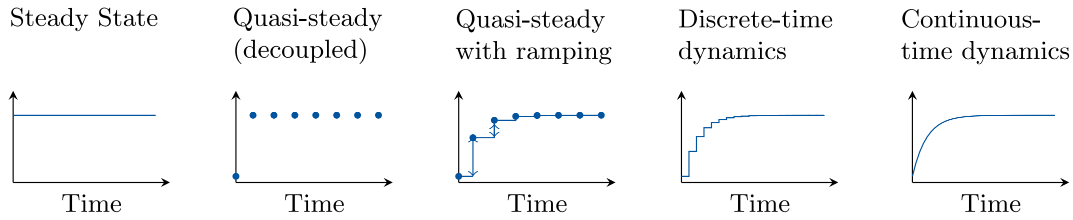

Figure 4: Schematic overview of different paradigms for modeling process dynamics.

The continuous-time solution of differential equations embedded in an optimal control problem can be approximated with different levels of accuracy, see Figure 4. We achieve the smallest but least accurate model when replacing the transient state trajectory with a steady state. The most accurate but computationally costly approximation is given by a discrete-time trajectory obtained from an accurate integration scheme. As a compromise, we may use steady-state assumptions resulting in quasi-steady behavior. The discussion of whether the system dynamics can be adequately represented by decoupled operating points is not in scope of this tutorial but discussed elsewhere [e.g., 33, 72, 82, 93]. Instead, we focus on quasi-steady operation where the dynamic system behavior is explicitly considered by coupling subsequent state values in time.

In quasi-steady operation, we may introduce a coupling in time via transition modes [e.g., 100], ramping constraints limiting the difference of subsequent state values [e.g., 25, 29, 72] or a combination of these approaches [e.g., 64]. Pointing to a parallel with discrete-time solutions of differential equations, we focus on first-order static ramping where the difference between subsequent state values is constrained by a constant ramp

$$\begin{aligned} \hat{x}_{i+1} &\leqslant \hat{x}_i + h_i \cdot \frac{\Delta x^{\text{ramp}}}{\Delta t^{\text{ramp}}} && \forall i = 0, \ldots, n_u - 1 \\ \hat{x}_{i+1} &\geqslant \hat{x}_i - h_i \cdot \frac{\Delta x^{\text{ramp}}}{\Delta t^{\text{ramp}}} && \forall i = 0, \ldots, n_u - 1 \quad . \end{aligned} \tag{staticRamp}$$

Moving the ramping into bounds of an auxiliary variable yields

$$\begin{aligned} \hat{x}_{i+1} &= \hat{x}_i + h_i \cdot \hat{u}_i && \forall i = 0, \ldots, n_u - 1 \\ \hat{u}_i &\in \left[-\frac{\Delta x^{\text{ramp}}}{\Delta t^{\text{ramp}}}, \frac{\Delta x^{\text{ramp}}}{\Delta t^{\text{ramp}}} \right] && \forall i = 0, \ldots, n_u - 1 \quad . \end{aligned} \tag{explEuler}$$

For operational optimization, the main difference between (staticRamp) and (explEuler) is that we exclusively optimize for the state values in the former while we optimize for the control and, when using full discretization, also the state values in the latter.

The difference equation of (explEuler) resembles the explicit Euler scheme for differential equation $\frac{\mathrm{d}x}{\mathrm{d}t} = u$. We extend (explEuler) to general explicit ODE's $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, u)$ based on a dynamic ramping approach proposed for integrated scheduling and control [6, 7]. Therein, Baader et al. replace the process dynamics with time-dependent bounds on the scheduling control which corresponds to the backward direction of our reformulation, namely from (explEuler) to (staticRamp). For this, the existence of at least two distinct controls is exploited: at least one for scheduling, i.e., for determining the optimal operating point, and at least one for control, i.e., for maintaining the operating point.

We adapt the dynamic ramping approach from Baader et al. [6, 7] to our example which contains only one control. In fact, we replace the process dynamics with time-dependent bounds on subsequent state values by applying the explicit Euler scheme and introducing an auxiliary variable

$$\hat{x}_{i+1} = \hat{x}_i + h_i \cdot \underbrace{f(\hat{x}_i, \hat{u}_i)}_{=:\nu(t_i)} \qquad \forall i = 0, \ldots, n_u - 1 \quad .$$

The process dynamics and the control bounds determine the bounds on ramping variable ν. We extend our example (OCP) with control bounds $u(t) \in [-u^{\max}, u^{\max}]$ and state bounds $x(t) \in [x^{\min}, x^{\max}]$ for any $t \in [0, t_f]$. We obtain state-dependent ramping limits $\nu\,(x(t)) \in [\frac{-u^{\max}}{x(t)}, \frac{u^{\max}}{x(t)}]$, $t \in [t_i, t_{i+1}]$. However, we have only limited information on the transient state values at the interior of the subintervals $t \in (t_i, t_{i+1})$ in quasi-steady operation: We only know the neighboring state values x_i and x_{i+1} as well as the general state bounds. Thus, we propose three different ramping strategies:

(R1) *Maximum flexible ramping* allowing for the largest deviations possible with the given differential equation, i.e.,

$$\nu(t) \in [\min_{x\in[x^{\min},x^{\max}]} \frac{-u^{\max}}{x}, \max_{x\in[x^{\min},x^{\max}]} \frac{u^{\max}}{x}] = [\frac{-u^{\max}}{x^{\min}}, \frac{u^{\max}}{x^{\min}}] \quad \forall t \in [t_i, t_{i+1})$$

(R2) *Minimum flexible ramping* guaranteeing that we cannot exceed the inertia given by the differential equation, i.e.,

$$\nu(t) \in [\max_{x\in[x^{\min},x^{\max}]} \frac{-u^{\max}}{x}, \min_{x\in[x^{\min},x^{\max}]} \frac{u^{\max}}{x}] = [\frac{-u^{\max}}{x^{\max}}, \frac{u^{\max}}{x^{\max}}] \quad \forall t \in [t_i, t_{i+1})$$

(R3) *Dynamic ramping* with time-dependent bounds assuming that we have a steady state in the interval, i.e.,

$$\nu(t) \in [\frac{-u^{\max}}{x(t_i)}, \frac{u^{\max}}{x(t_i)}] \quad \forall t \in [t_i, t_{i+1})$$

Maximum and minimum flexible ramping are static, i.e., independent from time discretization. When using dynamic ramping with a refined time grid, we can, however, obtain a more accurate approximation of the reachable state differences.

Quasi-steady operational optimization with dynamic ramping

$$\begin{aligned} \min_{\hat{\boldsymbol{x}}} \quad & \sum_{i=0}^{n_u} \hat{\mathcal{L}}(\hat{x}_i) \\ \text{s.t.} \quad & \hat{x}_{i+1} \leqslant \hat{x}_i + h_i \cdot \frac{u^{\max}}{\hat{x}_i} && \forall i = 0, \dots, n_u - 1 \\ & \hat{x}_{i+1} \geqslant \hat{x}_i - h_i \cdot \frac{u^{\max}}{\hat{x}_i} && \forall i = 0, \dots, n_u - 1 \\ & \hat{x}_0 = x_0 \\ & \hat{x}_{n_u} = x_f \end{aligned} \tag{dynRamp}$$

shows strong parallels to direct methods using the explicit Euler scheme for solving the differential equations. In fact, dynamic ramping (R3) is an equivalent formulation for the system of (in)equalities consisting of the related difference equation and the control bounds. Thus, the feasible set for the discretized states $\hat{\boldsymbol{x}}$ is the same in optimization problems (dynRamp), (FullDisc:Euler), and (SiSh:Euler). Despite this similarity, there are three structural differences between the quasi-steady optimization problem (dynRamp) and the discrete-time optimal control problems (FullDisc:Euler) and (SiSh:Euler). Firstly, the subsequent state values are coupled via inequality constraints in (dynRamp), while they are coupled via equality constraints and assignments in (FullDisc:Euler) and (SiSh:Euler), respectively. Which of these modeling alternatives results in more favorable convergence properties depends on the optimization algorithm used and its specific implementation. Secondly, the piecewise constant states and controls share the same time discretization in (dynRamp), while (FullDisc:Euler) and (SiSh:Euler) allow for two different time grids similar to (MuSh:Euler). Thirdly, the control values are chosen implicitly in (dynRamp), while (FullDisc:Euler) and (SiSh:Euler) explicitly provide the optimal control values. In fact, the quasi-steady problem formulation determines the optimal state values such that the difference of subsequent values stays in the limits given by any control value within $[u^{\min}, u^{\max}]$. However, we may need explicit information on the control values for implementing the optimal control strategy in reality. In the best case, we can reformulate the ramping constraints such that we can calculate the control values from the optimal state values. In the worst case, we need to solve another optimal control problem, e.g., where the objective value is the deviation of the current state prediction to the optimal state trajectory. For (OCP), we obtain

$$\begin{aligned} \min_{u(\cdot)} \quad & \int_0^{t_f} \left(x(t) - x^{\text{opt}}(t)\right)^2 \, \mathrm{d}t \\ \text{s.t.} \quad & \frac{\mathrm{d}x}{\mathrm{d}t}(t) = \frac{u(t)}{x(t)} \quad \forall t \in [0, t_f] \\ & x(0) = x_0 \end{aligned}$$

when using the integral least squares deviation as the error measure.

Conclusion 11. Both discrete-time dynamics and quasi-steady operation provide the same approximate solution to the differential equations if the numerical integration scheme used in the former is captured by an adequate ramping in the latter. The resulting optimization problems show nevertheless structural differences in the handling of constraints and optimization variables as well as the specific information provided for the optimal control strategy. □

4 Conclusions

In this tutorial, we investigated ambiguous terminology for standard optimal control methods to clear potential misunderstandings between different research groups. We will now use this analysis to dare to extend some warnings and recommendations.

In the introduction, we started with briefly delineating "optimal control" from which we conclude:

!1 "Optimal control" refers to the class of infinite-dimensional optimization problems with differential (–algebraic) equations embedded *or* closed-loop control problems.
When using "optimal control", specify whether it is open-loop or closed-loop optimal control.

In the main part, we focused on direct methods for offline open-loop optimal control. In detail, we argued that direct collocation is a specific variant of (simultaneous) full discretization which, in turn, may be considered as a specific simultaneous method. Note, however, that pairs of these three terms are used synonymously in parts of the literature. Analogously, "direct single shooting" is handled as synonym or specification of "partial discretization", while "direct single shooting" and "sequential method" mainly refer to the same method. Literally, "control vector parameterization" means discretization of controls and is therefore a paradigm shared by all direct methods. Historically, however, it became a synonym for direct single shooting in parts of the literature and, thus, may just refer to a specific direct method. In brief:

!2 Although collocation is the standard approach used within the full discretization approach, there are alternative state discretization approaches.
Distinguish between full discretization and direct collocation.

!3 The terms "full discretization", "simultaneous method", and "control vector parameterization" are ambiguous, cf. Figure 1.
Specify the intended meaning when using any of these terms.

Commonly, the embedded state integration within direct single shooting is performed by an external ODE or DAE solver. Alternatively, a numerical integration scheme may be used which is evaluated within the optimizer's iterations. This alternative has been used in the origins of direct single shooting and is used in hybrid methods combining full discretization and direct shooting. The numerical integration scheme can be hidden from the optimizer when using assignments rather than equality constraints. This yields a reduced-space formulation of the full discretization approach. Based on the usage of numerical integration schemes, we continued to elaborate a strong parallel between the discrete-time solution obtained by direct methods and the optimal quasi-steady operation considering adequate ramping constraints. Note, however, that there are distinct definitions of "quasi-steady" and "discrete-time". We conclude:

!4 Although the use of an external integration software is the standard for the state integration embedded in direct single shooting, there are alternative approaches.
Recall that most external integrators implement a numerical integration scheme when studying and extending direct shooting approaches.

!5 "Quasi-steady operation" may refer to a(n) (almost) constant operating point *or* a time-series of (almost) constant operating points.
Specify the intended meaning when using "quasi-steady".

!6 "Discrete-time system" refers to a reformulation of dynamic systems by discretizing the differential equations *or* a specific time representation used in scheduling problems.
Specify the intended meaning when using "discrete-time system".

Following the motivation of this tutorial, investigating hidden parallels between solution methods for mixed-integer nonlinear programming, mixed-integer optimal control, and scheduling with finitely many alternative control values merits careful attention. For example, continuous-time systems in scheduling [e.g., 38] show parallels to switching time optimization in mixed-integer optimal control [e.g., 4, 40, 76]: in both approaches the optimal time point of control switches, e.g., for changing modes or starting a new process, is chosen freely from a continuous-valued time interval. Besides, both scheduling and optimal control problems may contain discrete-continuous dynamic systems, i.e., differential equations which switch their behavior at specific time points. For both, methods for solving "hybrid systems" may be of interest [e.g., 8, 19, 68, 80].

Acknowledgement

Funding: This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant MI 1851/10-1 "Parameter estimation with (almost) deterministic global optimization".

Susanne Sass is grateful for her association to the International Research Training Group (DFG) IRTG-2379 "Hierarchical and Hybrid Approaches in Modern Inverse Problems" under grant 333849990/GRK2379.

CRediT authorship contribution statement

S. Sass: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original draft, Writing – Review & Editing, Funding acquisition. **A. Mitsos:** Conceptualization, Resources, Writing – Review & Editing, Supervision, Funding acquisition.

Bibliography

[1] J. Albersmeyer and M. Diehl. "The Lifted Newton Method and Its Application in Optimization". In: *SIAM J. on Optimization* 20.3 (2010), pp. 1655–1684. doi: 10.1137/080724885.

[2] D. C. Alsop. "The sensitivity of low flip angle RARE imaging". In: *Magnetic resonance in medicine* 37.2 (1997), pp. 176–184. doi: 10.1002/mrm.1910370206.

[3] R. Amrit, J. B. Rawlings, and L. T. Biegler. "Optimizing process economics online using model predictive control". In: *Computers & Chemical Engineering* 58 (2013), pp. 334–343. doi: 10.1016/j.compchemeng.2013.07.015.

[4] M. P. Avraam, N. Shah, and C. C. Pantelides. "Modelling and optimisation of general hybrid systems in the continuous time domain". In: *Computers & Chemical Engineering* 22 (1998), S221 –S228. doi: 10.1016/S0098-1354(98)00058-1.

[5] E. Aydin, D. Bonvin, and K. Sundmacher. "Toward Fast Dynamic Optimization: An Indirect Algorithm That Uses Parsimonious Input Parameterization". In: *Industrial & Engineering Chemistry Research* 57.30 (2018), pp. 10038–10048. doi: 10.1021/acs.iecr.8b02109.

[6] F. J. Baader, P. Althaus, A. Bardow, and M. Dahmen. "Dynamic ramping for demand response of processes and energy systems based on exact linearization". In: *Journal of Process Control* 118 (2022), pp. 218–230. doi: 10.1016/j.jprocont.2022.08.017.

[7] F. J. Baader, P. Althaus, A. Bardow, and M. Dahmen. "Demand response for flat nonlinear MIMO processes using dynamic ramping constraints". In: *Computers & Chemical Engineering* 172 (2023), p. 108171. doi: 10.1016/j.compchemeng.2023.108171.

[8] P. I. Barton and C. K. Lee. "Modeling, Simulation, Sensitivity Analysis, and Optimization of Hybrid Systems". In: *ACM Transactions on Modeling and Computer Simulation* 12.4 (2002), pp. 256–289. doi: 10.1145/643120.643122.

[9] R. Bellman. "The theory of dynamic programming". In: *Bulletin of the American Mathematical Society* 60.6 (1954), pp. 503–515. doi: 10.1090/s0002-9904-1954-09848-8.

[10] J. T. Betts. *Practical methods for optimal control using nonlinear programming*. Philadelphia: Society for Industrial and Applied Mathematics, 2010. doi: 10.1137/1.9780898718577.

[11] J. T. Betts and W. P. Huffman. "Mesh refinement in direct transcription methods for optimal control". In: *Optimal Control Applications and Methods* 19.1 (1998), pp. 1–21. doi: 10.1002/(SICI)1099-1514(199801/02)19:1⟨1::AID-OCA616⟩3.0.CO;2-Q.

[12] L. T. Biegler. "Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation". In: *Computers & Chemical Engineering* 8.3-4 (1984), pp. 243–247. doi: 10.1016/0098-1354(84)87012-X.

[13] L. T. Biegler. *Nonlinear programming: Concepts, algorithms, and applications to chemical processes*. Philadelphia: MOS-SIAM, 2010. doi: 10.1137/1.9780898719383.

[14] L. T. Biegler. "Nonlinear programming strategies for dynamic chemical process optimization". In: *Theoretical Foundations of Chemical Engineering* 48.5 (2014), pp. 541–554. doi: 10.1134/S0040579514050157.

[15] L. T. Biegler and I. E. Grossmann. "Retrospective on optimization". In: *Computers & Chemical Engineering* 28.8 (2004), pp. 1169–1192. doi: 10.1016/j.compchemeng.2003.11.003.

[16] T. Binder, L. Blank, H. G. Bock, R. Bulirsch, W. Dahmen, M. Diehl, T. Kronseder, W. Marquardt, J. P. Schlöder, and O. von Stryk. "Introduction to Model Based Optimization of Chemical Processes on Moving Horizons". In: *Online Optimization of Large Scale Systems*. Ed. by M. Grötschel, S. O. Krumke, and J. Rambau. Berlin, Heidelberg: Springer, 2001, pp. 295–339. doi: 10.1007/978-3-662-04331-8_18.

[17] K. B. Bischoff. "Accuracy of the pseudo steady state approximation for moving boundary diffusion problems". In: *Chemical Engineering Science* 18.11 (1963), pp. 711–713. doi: 10.1016/0009-2509(63)85050-2.

[18] H. G. Bock. "Numerical Treatment of Inverse Problems in Chemical Reaction Kinetics". In: *Modelling of Chemical Reaction Systems*. Ed. by V. I. Goldanskii, F. P. Schäfer, R. Gomer, J. P. Toennies, K. H. Ebert, P. Deuflhard, and W. Jäger. Vol. 18. Springer Series in Chemical Physics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 102–125. doi: 10.1007/978-3-642-68220-9_8.

[19] H. G. Bock, C. Kirches, A. Meyer, and A. Potschka. "Numerical solution of optimal control problems with explicit and implicit switches". In: *Optimization Methods and Software* 33.3 (2018), pp. 450–474. doi: 10.1080/10556788.2018.1449843.

[20] H. G. Bock and K.-J. Plitt. "A Multiple Shooting Algorithm for Direct Solution of Optimal Control Problems". In: *IFAC Proceedings Volumes* 17.2 (1984), pp. 1603–1608. doi: 10.1016/S1474-6670(17)61205-9.

[21] M. Bodenstein. "Eine Theorie der photochemischen Reaktionsgeschwindigkeiten". In: *Zeitschrift für Physikalische Chemie* 85U.1 (1913), pp. 329–397. doi: 10.1515/zpch-1913-8512.

[22] D. Bongartz. "Deterministic global flowsheet optimization for the design of energy conversion processes". Dissertation. Aachen, Germany: RWTH Aachen University, 2020. doi: 10.18154/RWTH-2020-06052.

[23] D. Bongartz and A. Mitsos. "Deterministic global optimization of process flowsheets in a reduced space using McCormick relaxations". In: *Journal of Global Optimization* 69.4 (2017), pp. 761–796. doi: 10.1007/s10898-017-0547-4.

[24] J. R. Bowen, A. Acrivos, and A. K. Oppenheim. "Singular perturbation refinement to quasi-steady state approximation in chemical kinetics". In: *Chemical Engineering Science* 18.3 (1963), pp. 177–188. doi: 10.1016/0009-2509(63)85003-4.

[25] L. C. Brée, K. Perrey, A. Bulan, and A. Mitsos. "Demand side management and operational mode switching in chlorine production". In: *AIChE Journal* 65.7 (2019). doi: 10.1002/aic.16352.

[26] G. E. Briggs and J. B. Haldane. "A Note on the Kinetics of Enzyme Action". In: *The Biochemical journal* 19.2 (1925), pp. 338–339. doi: 10.1042/bj0190338.

[27] H. Brunner. *Collocation methods for Volterra integral and related functional differential equations*. Vol. 15. Cambridge monographs on applied and computational mathematics. Cambridge, UK and New York: Cambridge University Press, 2004. isbn: 9780511263613.

[28] R. G. Brusch and R. H. Schappelle. "Solution of Highly Constrained Optimal Control Problems Using Nonlinear Programing". In: *AIAA Journal* 11.2 (1973), pp. 135–136. doi: 10.2514/3.50443.

[29] M. Carrion and J. M. Arroyo. "A Computationally Efficient Mixed-Integer Linear Formulation for the Thermal Unit Commitment Problem". In: *IEEE Transactions on Power Systems* 21.3 (2006), pp. 1371–1378. doi: 10.1109/TPWRS.2006.876672.

[30] A. M. Cervantes and L. T. Biegler. "A stable elemental decomposition for dynamic process optimization". In: *Journal of Computational and Applied Mathematics* 120.1 (2000), pp. 41–57. doi: 10.1016/S0377-0427(00)00302-2.

[31] G. Cimini, Y. Kim, B. McCain, J. Siegel, and A. Stefanopoulou. "Model Predictive Control for Real-time Position Tracking of a Catenary-free Tram". In: *IFAC-PapersOnLine* 50.1 (2017), pp. 1000–1005. doi: 10.1016/j.ifacol.2017.08.206.

[32] J. E. Cuthrell and L. T. Biegler. "On the optimization of differential-algebraic process systems". In: *AIChE Journal* 33.8 (1987), pp. 1257–1270. doi: 10.1002/aic.690330804.

[33] A. D'Amato, Y. Wang, D. Filev, and E. Remes. "On the Tradeoffs between Static and Dynamic Adaptive Optimization for an Automotive Application". In: *SAE International Journal of Commercial Vehicles* 10.1 (2017), pp. 346–352. doi: 10.4271/2017-01-0605.

[34] M. Diehl, H. G. Bock, H. Diedam, and P.-B. Wieber. "Fast Direct Multiple Shooting Algorithms for Optimal Robot Control". In: *Fast Motions in Biomechanics and Robotics*. Ed. by M. Diehl and K. Mombaur. Vol. 340. Lecture Notes in Control and Information Sciences. Berlin Heidelberg: Springer-Verlag GmbH, 2006, pp. 65–93. DOI: 10.1007/978-3-540-36119-0_4.

[35] J. W. Evans, J. Szekely, W. H. Ray, and Y. K. Chuang. "On the optimum temperature progression for irreversible non-catalytic gas—solid reactions". In: *Chemical Engineering Science* 28.3 (1973), pp. 683–690. DOI: 10.1016/0009-2509(77)80002-X.

[36] W. F. Feehery and P. I. Barton. "Dynamic simulation and optimization with inequality path constraints". In: *Computers & Chemical Engineering* 20 (1996), S707–S712. DOI: 10.1016/0098-1354(96)00127-5.

[37] L. N. G. Filon. "The forces on a cylinder in a stream of viscous fluid". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 113.763 (1926), pp. 7–27. DOI: 10.1098/rspa.1926.0136.

[38] C. A. Floudas and X. Lin. "Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review". In: *Computers & Chemical Engineering* 28.11 (2004), pp. 2109–2129. DOI: 10.1016/j.compchemeng.2004.05.002.

[39] R. A. Frazer, W. N. P. Jones, and S. W. Skan. "Approximations to functions and to the solutions of differential equations". In: *Reports and Memoranda of the Aeronautical Research Committee* 1799 (1937).

[40] M. Gerdts. "A variable time transformation method for mixed-integer optimal control problems". In: *Optimal Control Applications and Methods* 27.3 (2006), pp. 169–182. DOI: 10.1002/oca.778.

[41] A. Ghobeity and A. Mitsos. "Optimal time-dependent operation of seawater reverse osmosis". In: *Desalination* 263.1 (2010), pp. 76–88. DOI: 10.1016/j.desal.2010.06.041.

[42] A. Guillou and J. L. Soulé. "La résolution numérique des problèmes différentiels aux conditions initiales par des méthodes de collocation". In: *Revue française d'informatique et de recherche opérationnelle. Série rouge* 3.R3 (1969), pp. 17–44. DOI: 10.1051/m2an/196903R300171.

[43] C. R. Hargraves and S. W. Paris. "Direct trajectory optimization using nonlinear programming and collocation". In: *Journal of Guidance, Control, and Dynamics* 10.4 (1987), pp. 338–342. DOI: 10.2514/3.20223.

[44] F. G. Heineken, H. M. Tsuchiya, and R. Aris. "On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics". In: *Mathematical Biosciences* 1.1 (1967), pp. 95–113. DOI: 10.1016/0025-5564(67)90029-6.

[45] G. A. Hicks and W. H. Ray. "Approximation methods for optimal control synthesis". In: *The Canadian Journal of Chemical Engineering* 49.4 (1971), pp. 522–528. DOI: 10.1002/cjce.5450490416.

[46] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward. "SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers". In: *ACM Transactions on Mathematical Software* 31.3 (2005), pp. 363–396. DOI: 10.1145/1089014.1089020.

[47] W. Hong, S. Wang, P. Li, G. Wozny, and L. T. Biegler. "A quasi-sequential approach to large-scale dynamic optimization problems". In: *AIChE Journal* 52.1 (2006), pp. 255–268. DOI: 10.1002/aic.10625.

[48] B. Houska and B. Chachuat. "Branch-and-Lift Algorithm for Deterministic Global Optimization in Nonlinear Optimal Control". In: *Journal of Optimization Theory and Applications* 162.1 (2014), pp. 208–248. DOI: 10.1007/s10957-013-0426-1.

[49] R. X. Huang. "Oceanographic Topics: Thermohaline Circulation". In: *Encyclopedia of Atmospheric Sciences (Second Edition)*. Ed. by G. R. North, J. Pyle, and F. Zhang. Oxford: Academic Press, 2015, pp. 315–328. DOI: 10.1016/B978-0-12-382225-3.00281-4.

[50] R. Jackson. "An approach to numerical solution of time-dependent optimisation problems in 2-phase contacting devices". In: *Transactions of the Institution of Chemical Engineers* 45.4 (1967), T160.

[51] S. Kameswaran and L. T. Biegler. "Convergence rates for direct transcription of optimal control problems using collocation at Radau points". In: *Computational Optimization and Applications* 41.1 (2008), pp. 81–126. DOI: 10.1007/s10589-007-9098-9.

[52] C. D. Kappatou, D. Bongartz, J. Najman, S. Sass, and A. Mitsos. "Global dynamic optimization with Hammerstein–Wiener models embedded". In: *Journal of Global Optimization* 84.2 (2022), pp. 321–347. DOI: 10.1007/s10898-022-01145-z.

[53] M. T. Kelley, R. Baldick, and M. Baldea. "A direct transcription-based multiple shooting formulation for dynamic optimization". In: *Computers & Chemical Engineering* 140 (2020), p. 106846. DOI: 10.1016/j.compchemeng.2020.106846.

[54] M. Kelly. "An Introduction to Trajectory Optimization: How to Do Your Own Direct Collocation". In: *SIAM Review* 59.4 (2017), pp. 849–904. DOI: 10.1137/16M1062569.

[55] C. Lanczos. "Trigonometric Interpolation of Empirical and Analytical Functions". In: *Journal of Mathematics and Physics* 17.1-4 (1938), pp. 123–199. DOI: 10.1002/sapm1938171123.

[56] G. J. Lastman. "Sub-optimal open-loop control of non-linear systems using approximations for the controls". In: *International Journal of Control* 20.2 (1974), pp. 289–303. DOI: 10.1080/00207177408932738.

[57] M. A. LeMone. "Boundary Layer (Atmospheric) and Air Pollution: Convective Boundary Layer". In: *Encyclopedia of Atmospheric Sciences (Second Edition)*. Ed. by G. R. North, J. Pyle, and F. Zhang. Oxford: Academic Press, 2015, pp. 250–257. DOI: 10.1016/B978-0-12-382225-3.00085-2.

[58] P. Li, H. A. Garcia, G. Wozny, and E. Reuter. "Optimization of a Semibatch Distillation Process with Model Validation on the Industrial Site". In: *Industrial & Engineering Chemistry Research* 37.4 (1998), pp. 1341–1350. DOI: 10.1021/ie970695l.

[59] J. S. Logsdon and L. T. Biegler. "Decomposition strategies for large-scale dynamic optimization problems". In: *Chemical Engineering Science* 47.4 (1992), pp. 851–864. DOI: 10.1016/0009-2509(92)80272-E.

[60] R. Luus. "Control Vector Iteration CVI". In: *Encyclopedia of optimization*. Ed. by C. A. Floudas and P. M. Pardalos. New York: Springer, 2009, pp. 509–513. DOI: 10.1007/978-0-387-74759-0_93.

[61] D. Mayne. "A Second-order Gradient Method for Determining Optimal Trajectories of Non-linear Discrete-time Systems". In: *International Journal of Control* 3.1 (1966), pp. 85–95. DOI: 10.1080/00207176608921369.

[62] C. A. Méndez, J. Cerdá, I. E. Grossmann, Iiro Harjunkoski, and Marco Fahl. "State-of-the-art review of optimization methods for short-term scheduling of batch processes". In: *Computers & Chemical Engineering* 30.6–7 (2006), pp. 913–946. DOI: 10.1016/j.compchemeng.2006.02.008.

[63] L. Michaelis and M. L. Menten. "Die Kinetik der Invertinwirkung". In: *Biochemische Zeitschrift* 49 (1913), pp. 333–369.

[64] S. Mitra, I. E. Grossmann, J. M. Pinto, and N. Arora. "Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes". In: *Computers & Chemical Engineering* 38 (2012), pp. 171–184. DOI: 10.1016/j.compchemeng.2011.09.019.

[65] A. Mitsos, B. Chachuat, and P. I. Barton. "McCormick-Based Relaxations of Algorithms". In: *SIAM Journal on Optimization* 20.2 (2009), pp. 573–601. DOI: 10.1137/080717341.

[66] P. Muller. "Glossary of terms used in physical organic chemistry (IUPAC Recommendations 1994)". In: *Pure and Applied Chemistry* 66.5 (1994), pp. 1077–1184. DOI: 10.1351/pac199466051077.

[67] B. A. Ogunnaike and W. H. Ray. *Process dynamics, modeling, and control*. Vol. 1. Oxford University Press New York, 1994. ISBN: 9780195091199.

[68] J. Oldenburg and W. Marquardt. "Disjunctive modeling for optimal control of hybrid systems". In: *Computers & Chemical Engineering* 32.10 (2008), pp. 2346–2364. DOI: 10.1016/j.compchemeng.2007.12.002.

[69] L. R. Petzold. "A Description of DASSL: A Differential/Algebraic System Solver". In: *Sandia report SAND82-8637*. Livermore, CA: Sandia National Laboratories, 1982.

[70] G. P. Pollard and R. Sargent. "Off line computation of optimum controls for a plate distillation column". In: *Automatica* 6.1 (1970), pp. 59–76. DOI: 10.1016/0005-1098(70)90075-0.

[71] L. S. Pontryagin, V. G. Boltyanski, R. V. Gamkrelidze, and E. F. Miscenko. *The Mathematical Theory of Optimal Processes*. New York: John Wiley & Sons, 1962. ISBN: 978-0470693810.

[72] Y. Qin, P. Liu, and Z. Li. "Multi-timescale hierarchical scheduling of an integrated energy system considering system inertia". In: *Renewable and Sustainable Energy Reviews* 169 (2022), p. 112911. DOI: 10.1016/j.rser.2022.112911.

[73] R. Quirynen, S. Gros, and M. Diehl. "Lifted implicit integrators for direct optimal control". In: *2015 54th IEEE Conference on Decision and Control (CDC)*. 2015, pp. 3212–3217. DOI: 10.1109/CDC.2015.7402701.

[74] W. H. Ray. "The quasi–steady–state approximation in continuous stirred tank reactors". In: *The Canadian Journal of Chemical Engineering* 47.6 (1969), pp. 503–508. DOI: 10.1002/cjce.5450470601.

[75] G. W. Reddien. "Collocation at Gauss Points as a Discretization in Optimal Control". In: *SIAM Journal on Control and Optimization* 17.2 (1979), pp. 298–306. DOI: 10.1137/0317023.

[76] M. Ringkamp, S. Ober-Blöbaum, and S. Leyendecker. "On the time transformation of mixed integer optimal control problems using a consistent fixed integer control function". In: *Mathematical Programming* 161.1 (2017), pp. 551–581. doi: 10.1007/s10107-016-1023-5.

[77] W. Ritz. "Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik". In: *Journal für die reine und angewandte Mathematik* 135 (1909), pp. 1–61. doi: 10.1515/crll.1909.135.1.

[78] H. H. Rosenbrock and C. Storey. *Computational techniques for chemical engineers*. London: Pergamon Press, 1966. doi: 10.1016/C2013-0-01841-5.

[79] S. Sager. *Numerical methods for mixed-integer optimal control problems*. Der andere Verlag Tönning, Lübeck, Marburg, 2005.

[80] S. Sager. "Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control". In: *Journal of Process Control* 19.8 (2009), pp. 1238–1247. doi: 10.1016/j.jprocont.2009.03.008.

[81] R. W. H. Sargent and G. R. Sullivan. "The development of an efficient optimal control package". In: *Optimization Techniques*. Ed. by J. Stoer. Vol. 7. Lecture Notes in Control and Information Sciences. Berlin, Heidelberg: Springer, 1978, pp. 158–168. doi: 10.1007/BFb0006520.

[82] S. Sass and A. Mitsos. "Optimal Operation of Dynamic (Energy) Systems: When are Quasi-Steady Models Adequate?" In: *Computers & Chemical Engineering* 124 (2019), pp. 133–139. doi: 10.1016/j.compchemeng.2019.02.011.

[83] S. Sass, A. Mitsos, D. Bongartz, I. H. Bell, N. I. Nikolov, and A. Tsoukalas. "A branch-and-bound algorithm with growing datasets for large-scale parameter estimation". In: *European Journal of Operational Research* 316.1 (2024), pp. 36–45. doi: 10.1016/j.ejor.2024.02.020.

[84] P. Schäfer, A. M. Schweidtmann, P. H. Lenz, H. M. Markgraf, and A. Mitsos. "Wavelet-based grid-adaptation for nonlinear scheduling subject to time-variable electricity prices". In: *Computers & Chemical Engineering* 132 (2020), p. 106598. doi: 10.1016/j.compchemeng.2019.106598.

[85] L. F. Shampine and M. W. Reichelt. "The MATLAB ODE Suite". In: *SIAM Journal on Scientific Computing: A publication of the Society for Industrial and Applied Mathematics* 18.1 (1997), pp. 1–22. doi: 10.1137/S1064827594276424.

[86] A. B. Singer and P. I. Barton. "Global Solution of Optimization Problems with Parameter-Embedded Linear Dynamic Systems". In: *Journal of Optimization Theory and Applications* 121.3 (2004), pp. 613–646. doi: 10.1023/B:JOTA.0000037606.79050.a7.

[87] H. Sirisena. "Computation of optimal controls using a piecewise polynomial parameterization". In: *IEEE Transactions on Automatic Control* 18.4 (1973), pp. 409–411. doi: 10.1109/TAC.1973.1100329.

[88] B. Srinivasan. "A guide to the Michaelis-Menten equation: steady state and beyond". In: *The FEBS journal* 289.20 (2022), pp. 6086–6098. doi: 10.1111/febs.16124.

[89] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. 3 ed. Vol. 12. New York, NY: Springer New York, 2002. doi: 10.1007/978-0-387-21738-3.

[90] D. Tabak and B. C. Kuo. "Application of mathematical programming in the design of optimal control systems". In: *International Journal of Control* 10.5 (1969), pp. 545–552. doi: 10.1080/00207176908905855.

[91] V. S. Vassiliadis, R. W. Sargent, and C. C. Pantelides. "Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints". In: *Industrial & Engineering Chemistry Research* 33.9 (1994), pp. 2111–2122. doi: 10.1021/ie00033a014.

[92] J. V. Villadsen and W. E. Stewart. "Solution of boundary-value problems by orthogonal collocation". In: *Chemical Engineering Science* 22.11 (1967), pp. 1483–1501. doi: 10.1016/0009-2509(67)80074-5.

[93] P. Voll, S. Kirschbaum, and A. Bardow. "Evaluation of Quasi-Stationary Simulation for the Analysis of Industrial Energy Systems". In: *ecos 2010 : 23rd International Conference on Efficiency, Cost, Optimization, Simulation & Environmental Impact of Energy System, June 14-20, 2010, Lausanne, Switzerland ; book of papers*. 2010, pp. 34–41.

[94] W. E. Williamson. "Use of polynomial approximations to calculate suboptimal controls". In: *AIAA Journal* 9.11 (1971), pp. 2271–2273. doi: 10.2514/3.6499.

[95] K. Wright. "Some relationships between implicit Runge-Kutta, collocation and Lanczos τ methods, and their stability properties". In: *BIT* 10.2 (1970), pp. 217–227. doi: 10.1007/BF01936868.

[96] M. Zennaro. "Natural Continuous Extensions of Runge-Kutta Methods". In: *Mathematics of Computation* 46.173 (1986), p. 119. doi: 10.2307/2008218.

[97] Q. Zhang and I. E. Grossmann. "Enterprise-wide optimization for industrial demand side management: Fundamentals, advances, and perspectives". In: *Chemical Engineering Research and Design* 116 (2016), pp. 114–131. doi: 10.1016/j.cherd.2016.10.006.

[98] Q. Zhang, A. Sundaramoorthy, I. E. Grossmann, and J. M. Pinto. "A discrete-time scheduling model for continuous power-intensive process networks with various power contracts". In: *Computers & Chemical Engineering* 84 (2016), pp. 382–393. doi: 10.1016/j.compchemeng.2015.09.019.

[99] X. Zhang, G. Hug, and I. Harjunkoski. "Cost-Effective Scheduling of Steel Plants With Flexible EAFs". In: *IEEE Transactions on Smart Grid* 8.1 (2017), pp. 239–249. doi: 10.1109/TSG.2016.2575000.

[100] D. Zhou, K. Zhou, L. Zhu, J. Zhao, Z. Xu, Z. Shao, and X. Chen. "Optimal scheduling of multiple sets of air separation units with frequent load-change operation". In: *Separation and Purification Technology* 172 (2017), pp. 178–191. doi: 10.1016/j.seppur.2016.08.009.