Numer. Math. (2013) 124:337–360
DOI 10.1007/s00211-013-0516-x

Numerische
Mathematik

W-methods in optimal control

J. Lang · J. G. Verwer

Received: 14 April 2011 / Revised: 27 October 2012 / Published online: 19 February 2013

Abstract This paper addresses consistency and stability of W-methods up to order three for nonlinear ODE-constrained control problems with possible restrictions on the control. The analysis is based on the transformed adjoint system and the control uniqueness property. These methods can also be applied to large-scale PDE-constrained optimization, since they offer an efficient way to compute gradients of the discrete objective function.

Mathematics Subject Classification (2000) Primary 34H05 · 49J15 · 65L05 · 65L06

1 Introduction

Suppose one is given the nonlinear optimal control problem

$$\text{minimize } C(\mathbf{x}(1)) \tag{1.1}$$
$$\text{subject to } \mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{u}(t) \in U, \quad t \in (0, 1], \tag{1.2}$$
$$\mathbf{x}(0) = \mathbf{x}_0, \tag{1.3}$$

J. Lang (✉)
Department of Mathematics, Center of Smart Interfaces, Graduate School of Computational Engineering, Technische Universität Darmstadt, Dolivostraße 15, 64293 Darmstadt, Germany
e-mail: lang@mathematik.tu-darmstadt.de

J. G. Verwer
Center for Mathematics and Computer Science, P.O. Box 94079, 1090 GB Amsterdam,
The Netherlands
e-mail: Jan.Verwer@cwi.nl

where the state $\mathbf{x}(t) \in \mathbb{R}^d$, the control $\mathbf{u}(t) \in \mathbb{R}^m$, $\mathbf{f} : \mathbb{R}^d \times \mathbb{R}^m \mapsto \mathbb{R}^d$, the objective function $C : \mathbb{R}^d \mapsto \mathbb{R}$, and $U \subset \mathbb{R}^m$ is closed and convex. Assuming sufficient smoothness for $\mathbf{f}$ and C (see, e.g. [4]), there exists associated Lagrange multipliers $\boldsymbol{\psi}^*$ such that the first-order optimality conditions are satisfied at $(\mathbf{x}^*, \boldsymbol{\psi}^*, \mathbf{u}^*)$:

$$\mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad t \in (0, 1], \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{1.4}$$

$$\boldsymbol{\psi}'(t) = -\boldsymbol{\psi}\nabla_x \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad t \in [0, 1), \quad \boldsymbol{\psi}(1) = \nabla C(\mathbf{x}(1)), \tag{1.5}$$

$$-\boldsymbol{\psi}\nabla_u \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \in N_U(\mathbf{u}(t)), \quad t \in [0, 1]. \tag{1.6}$$

Here, $\boldsymbol{\psi}$ is a row vector in $\mathbb{R}^d$, $\nabla_x \mathbf{f}$ and $\nabla_u \mathbf{f}$ are the Jacobian matrices of $\mathbf{f}$ with respect to $\mathbf{x}$ and $\mathbf{u}$, and the normal cone mapping $N_U(\mathbf{u})$ is defined for any $\mathbf{u} \in U$ as follows

$$N_U(\mathbf{u}) = \{\mathbf{w} \in \mathbb{R}^m : \mathbf{w}^T(\mathbf{v} - \mathbf{u}) \le 0 \text{ for all } \mathbf{v} \in U\}. \tag{1.7}$$

In the *first-optimize-then-discretize* approach the system (1.4)–(1.6) is discretized by applying the numerical solver of choice. The focus of this paper is to analyze discrete adjoints which are derived from W-method discretizations of (1.2)–(1.3). They are useful in optimization since they allow the efficient computation of gradients of the discretized objective function, i.e., the numerical function that is being numerically minimized. This approach is known as *first-discretize-then-optimize*.

Hager [4] has studied discrete Runge–Kutta adjoints with strictly positive weights and found that additional order conditions have to be satisfied to achieve order three and higher for optimal control problems, while any first- or second-order Runge–Kutta scheme retains its order. All fourth-order four-stage explicit Runge–Kutta schemes automatically satisfy the order conditions for optimal control. His analysis utilizes a transformed adjoint system and the control uniqueness property, which will be also used in our context of W-methods. It turned out that the consistency analysis of Runge–Kutta schemes coming from the discretization of optimal control problems can be elegantly done in the class of partitioned symplectic Runge–Kutta schemes. Applying the technique of oriented free trees, Bonnans and Laurent-Varin [1] have computed the corresponding order conditions up to order seven by means of an appropriate computer program. The same number of conditions were already given by Murua [12]. A larger class of non-symplectic second-order Runge–Kutta methods has been investigated by Pulova [13]. Reverse mode automatic differentiation on explicit Runge–Kutta methods has been considered by Walther [24], who concluded that the order of the discretization is always preserved by the discrete adjoints. For problems where only the initial conditions are the control variables, consistency properties of discrete adjoint Runge–Kutta and linear multistep methods are presented by Sandu [15,16].

Many practical optimal control problems demand for stiff ODE integrators, especially when the constraints are derived from semi-discretizations of nonlinear time-dependent parabolic PDEs. In this case, the inherent nonlinear coupling of all stage values of a fully implicit Runge–Kutta scheme may become a severe structural disadvantage and computational bottleneck. Linearly implicit methods of Runge–Kutta–Rosenbrock type are much less expensive and have proven successful at the numerical solution of a wide range of stiff and large-scale systems [7,11,14,22].

Among this class of time integrators, W-methods are very popular, since they allow the use of an arbitrary matrix in place of the Jacobian matrix while maintaining the order of accuracy and thus have the potential to significantly reduce the computational costs [7,21]. W-methods fulfill the order conditions for explicit Runge–Kutta methods. This makes them also attractive for (automatic) partitioning strategies, where stiff and nonstiff components are treated in an implicit and explicit way, respectively [2,22].

2 Discrete optimal control problem

We discretize the differential equations (1.2) using an s-stage W-method [21] on a uniform mesh of width $h = 1/N$, where N is a natural number. Let $\mathbf{x}_n$ denote the sequence of approximations to the exact solution values $\mathbf{x}(t_n)$ with $t_n = nh$. Then the discrete optimal control problem reads

$$\text{minimize } \; C(\mathbf{x}_N) \tag{2.1}$$

$$\text{subject to } \; \mathbf{x}_{n+1} = \mathbf{x}_n + \sum_{i=1}^{s} b_i \mathbf{y}_{ni}, \quad \mathbf{x}_0 \text{ given}, \tag{2.2}$$

$$\mathbf{y}_{ni} = h\mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \mathbf{u}_{ni}\right) + hT_n \sum_{j=1}^{i} \gamma_{ij}\mathbf{y}_{nj}, \quad \mathbf{u}_{ni} \in U, \tag{2.3}$$

$$1 \le i \le s, \quad 0 \le n \le N-1. \tag{2.4}$$

The vectors $\mathbf{y}_{ni}$ and $\mathbf{u}_{ni}$ are intermediate state and control variables on the interval $[t_n, t_{n+1}]$. If h is small enough, the $\mathbf{y}_{ni}$ in (2.3) are uniquely determined in the neighbourhood of $(\mathbf{x}^*, \mathbf{u}^*)$. The coefficients b_i, α_{ij}, and γ_{ij} are chosen to obtain a desired order of consistency and A-stability or even L-stability. As usual, all coefficients γ_{ii} are taken constant, $\gamma_{ii} = \gamma$, so that per time step only linear systems with the same matrix $I - h\gamma T_n$ have to be solved. We formally set $\alpha_{ij} = 0,\ j \ge i$, and $\gamma_{ij} = 0,\ j > i$. The matrices T_n are arbitrary and constant within each time step. Thus in the analysis that follows, we will exploit the property that all derivatives of T_n vanish. Note that $T_n = 0$ yields a standard explicit Runge–Kutta method.

The main idea of W-methods is to use the matrix T_n to assure stability of the scheme. An illustrative example are large systems that can be partitioned into a small stiff and a large nonstiff system,

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{z}), \tag{2.5}$$

$$\mathbf{z}' = \mathbf{g}(\mathbf{y}, \mathbf{z}), \tag{2.6}$$

where $\mathbf{y}$ and $\mathbf{z}$ are the stiff and nonstiff components, respectively. Assuming that $\|\partial_{\mathbf{y}}\mathbf{f}\| \gg \|(\partial_{\mathbf{y}}\mathbf{g}, \partial_{\mathbf{z}}\mathbf{g})\|$, we can apply an implicit scheme for $\mathbf{y}$ and an explicit one for $\mathbf{z}$. In this case, an appropriate choice of the matrix T_n is

$$T_n = \begin{pmatrix} T_1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.7}$$

with $T_1 \approx \partial_{\mathbf{y}}\mathbf{f}(\mathbf{y}_n, \mathbf{z}_n)$. Here, $\mathbf{y}_n$ and $\mathbf{z}_n$ are approximate solutions at t_n. Since the order conditions are satisfied for arbitrary T_n, T_1 can be computed by finite differences without loosing accuracy and can often be maintained (together with its decomposition) over several time steps, which in general gives a spectacular reduction of the work necessary to solve the small linear systems of the size of the stiff components $\mathbf{y}$. In a similar way the idea also applies to systems with $\mathbf{f}(\mathbf{y}) = \mathbf{f}_1(\mathbf{y}) + \mathbf{f}_2(\mathbf{y})$, where $\mathbf{f}_1$ represents the stiff and $\mathbf{f}_2$ the nonstiff part. Reaction–diffusion equations with nonstiff reactions are a typical example for this kind of problems. Several applications are given in our numerical illustrations.

Suppose that multipliers $\boldsymbol{\lambda}_{ni}$ are introduced for the intermediate state equations (2.3) and that $\boldsymbol{\psi}_{n+1}$ is the associated (discrete) multiplier for Eq. (2.2). Then the first-order optimality conditions are the following:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \sum_{i=1}^{s} b_i \mathbf{y}_{ni}, \quad \mathbf{x}_0 \text{ given}, \tag{2.8}$$

$$\mathbf{y}_{ni} = h\mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \mathbf{u}_{ni}\right) + hT_n \sum_{j=1}^{i} \gamma_{ij}\mathbf{y}_{nj}, \tag{2.9}$$

$$\boldsymbol{\psi}_n - \boldsymbol{\psi}_{n+1} = h\sum_{i=1}^{s} \boldsymbol{\lambda}_{ni} \nabla_x \mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \mathbf{u}_{ni}\right), \quad \boldsymbol{\psi}_N = \nabla C(\mathbf{x}_N), \tag{2.10}$$

$$\boldsymbol{\lambda}_{ni} = b_i \boldsymbol{\psi}_{n+1} + h\sum_{j=1}^{s} \boldsymbol{\lambda}_{nj} \left(\alpha_{ji} \nabla_x \mathbf{f}\left(\mathbf{x}_n + \sum_{k=1}^{j-1} \alpha_{jk}\mathbf{y}_{nk}, \mathbf{u}_{nj}\right) + \gamma_{ji} T_n\right), \tag{2.11}$$

$$\mathbf{u}_{ni} \in U, \quad -\boldsymbol{\lambda}_{ni} \nabla_u \mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \mathbf{u}_{ni}\right) \in N_U(\mathbf{u}_{ni}), \tag{2.12}$$

$$1 \le i \le s, \quad 0 \le n \le N-1. \tag{2.13}$$

Remember that all dual multipliers are treated as row vectors. In the case that $b_i \neq 0$ for each i, Eqs. (2.10)–(2.11) can be reformulated in terms of new variables $\boldsymbol{\xi}_{ni} = \boldsymbol{\lambda}_{ni}/b_i$, $1 \le i \le s$,

$$\boldsymbol{\psi}_n = \boldsymbol{\psi}_{n+1} + h\sum_{i=1}^{s} b_i \boldsymbol{\xi}_{ni} \nabla_x \mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \mathbf{u}_{ni}\right), \quad \boldsymbol{\psi}_N = \nabla C(\mathbf{x}_N), \tag{2.14}$$

$$\boldsymbol{\xi}_{ni} = \boldsymbol{\psi}_{n+1} + h\sum_{j=1}^{s} \frac{b_j}{b_i} \boldsymbol{\xi}_{nj} \left(\alpha_{ji} \nabla_x \mathbf{f}\left(\mathbf{x}_n + \sum_{k=1}^{j-1} \alpha_{jk}\mathbf{y}_{nk}, \mathbf{u}_{nj}\right) + \gamma_{ji} T_n\right). \tag{2.15}$$

Condition (2.12) is replaced by

$$\mathbf{u}_{ni} \in U, \quad -b_i \boldsymbol{\xi}_{ni} \nabla_u \mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij} \mathbf{y}_{nj}, \mathbf{u}_{ni}\right) \in N_U(\mathbf{u}_{ni}). \tag{2.16}$$

Remark 2.1 A usual way to solve the first-order optimality conditions is to apply a gradient method. Let $\mathbf{u} \in \mathbb{R}^{msN}$ denote the vector of all intermediate control variables $\mathbf{u}_{ni}$. Since $\mathbf{x}_N$ depends on all components of $\mathbf{u}$, we can consider the minimization of the discrete cost function $\hat{C}(\mathbf{u}) = C(\mathbf{x}_N(\mathbf{u}))$. A short calculation shows

$$\nabla_{u_{ni}} \hat{C}(\mathbf{u}) = h b_i \boldsymbol{\xi}_{ni} \nabla_u \mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij} \mathbf{y}_{nj}, \mathbf{u}_{ni}\right). \tag{2.17}$$

Suppose a current iterate of the control variables is given. Using these values, the discrete state equations (2.8)–(2.9) can be solved for $\mathbf{x}_n$ and $\mathbf{y}_{ni}$ by marching forward from $n = 0$ to $n = N - 1$. Then all variables are given to solve the discrete costate equations (2.14)–(2.15) for $\boldsymbol{\psi}_n$ and $\boldsymbol{\xi}_{ni}$ by marching backward from $n=N-1$ to $n=0$. Notice that the special structure of the parameters α_{ji} and γ_{ji} allows a convenient way to successively compute the intermediate values $\boldsymbol{\xi}_{ni}$ for $i = s, s-1, \ldots, 1$, in each time step. Finally, the gradient is computed from (2.17) and the control iterate is updated.

We observe that the transformed adjoint equations (2.14)–(2.15) march backwards in time while the W-method (2.8)–(2.9) marches forwards in time. Following the approach used in [4] to facilitate the consistency analysis, we first reverse the order of time in the discrete adjoint equations. That is, we solve for $\boldsymbol{\psi}_{n+1}$ in (2.14) and substitute in (2.15) to obtain the following forward marching scheme:

$$\boldsymbol{\psi}_{n+1} = \boldsymbol{\psi}_n - h \sum_{i=1}^{s} b_i \boldsymbol{\xi}_{ni} \nabla_x \mathbf{f}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij} \mathbf{y}_{nj}, \mathbf{u}_{ni}\right), \tag{2.18}$$

$$\boldsymbol{\xi}_{ni} = \boldsymbol{\psi}_n - h \sum_{j=1}^{s} \bar{\alpha}_{ij} \boldsymbol{\xi}_{nj} \nabla_x \mathbf{f}\left(\mathbf{x}_n + \sum_{k=1}^{j-1} \alpha_{jk} \mathbf{y}_{nk}, \mathbf{u}_{nj}\right) - h \sum_{j=1}^{s} \bar{\gamma}_{ij} \boldsymbol{\xi}_{nj} T_n. \tag{2.19}$$

with the new coefficients

$$\bar{\alpha}_{ij} = \frac{b_i b_j - b_j \alpha_{ji}}{b_i}, \quad \bar{\gamma}_{ij} = -\frac{b_j \gamma_{ji}}{b_i}. \tag{2.20}$$

Next we will remove the control variables $\mathbf{u}$ by use of the control uniqueness property introduced in [4]. If $(\mathbf{x}, \boldsymbol{\psi})$ is sufficiently close to $(\mathbf{x}^*, \boldsymbol{\psi}^*)$, then under suitable assumptions there exists a locally unique minimizer $\mathbf{u} = \mathbf{u}(\mathbf{x}, \boldsymbol{\psi})$ of the Hamiltonian $\boldsymbol{\psi}\mathbf{f}(\mathbf{x}, \mathbf{u})$ over all $\mathbf{u} \in U$ and we can define functions

$$\boldsymbol{\phi}(\mathbf{x}, \boldsymbol{\psi}) = -\boldsymbol{\psi}\nabla_x \mathbf{f}(\mathbf{x}, \mathbf{u})|_{\mathbf{u}=\mathbf{u}(\mathbf{x},\boldsymbol{\psi})}, \quad \mathbf{g}(\mathbf{x}, \boldsymbol{\psi}) = \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}, \boldsymbol{\psi})). \tag{2.21}$$

We assume that the intermediate control variables have the special form

$$\mathbf{u}_{ni} = \mathbf{u}\left(\mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \boldsymbol{\xi}_{ni}\right), \quad 0 \le n \le N-1, \quad 1 \le i \le s, \tag{2.22}$$

and the weights b_i are strictly positive to assure that the associated controls are minimizers of the Hamiltonian.

Remark 2.2 Theorems 2.1 and 7.2. in [4] show for Runge–Kutta schemes that if $b_i > 0$ for each i then convergence to $(\mathbf{x}^*, \boldsymbol{\psi}^*, \mathbf{u}^*)$ can be achieved for both unconstrained, i.e., $U = \mathbb{R}^m$, and constrained control problems. Strictly positive weights must also be assumed in the context of automatic differentiation [24]. In [17] Runge–Kutta methods with strictly positive summarized weights that correspond to distinct control variables $\mathbf{u}_{ni}$ are studied. We will use this approach to construct a third-order W-method suitable for optimal control.

Introducing intermediate values $\mathbf{x}_{ni}$ for the state, the complete forward marching scheme can be written as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \sum_{i=1}^{s} b_i \mathbf{y}_{ni}, \quad \mathbf{x}_0 \text{ given}, \tag{2.23}$$

$$\boldsymbol{\psi}_{n+1} = \boldsymbol{\psi}_n + h\sum_{i=1}^{s} b_i \boldsymbol{\phi}(\mathbf{x}_{ni}, \boldsymbol{\xi}_{ni}), \quad \boldsymbol{\psi}_N = \nabla C(\mathbf{x}_N), \tag{2.24}$$

$$\mathbf{y}_{ni} = h\mathbf{g}(\mathbf{x}_{ni}, \boldsymbol{\xi}_{ni}) + hT_n \sum_{j=1}^{i} \gamma_{ij}\mathbf{y}_{nj}, \tag{2.25}$$

$$\boldsymbol{\xi}_{ni} = \boldsymbol{\psi}_n + h\sum_{j=1}^{s} \bar{\alpha}_{ij}\boldsymbol{\phi}(\mathbf{x}_{nj}, \boldsymbol{\xi}_{nj}) - h\sum_{j=1}^{s} \bar{\gamma}_{ij}\boldsymbol{\xi}_{nj}T_n, \tag{2.26}$$

$$\mathbf{x}_{ni} = \mathbf{x}_n + \sum_{j=1}^{i-1} \alpha_{ij}\mathbf{y}_{nj}, \tag{2.27}$$

$$1 \le i \le s, \quad 0 \le n \le N-1. \tag{2.28}$$

The key for consistency analysis is the observation that this scheme can be viewed as a discretization of the following two-point boundary-value problem:

$$\mathbf{x}'(t) = \mathbf{g}(\mathbf{x}(t), \boldsymbol{\psi}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{2.29}$$

$$\boldsymbol{\psi}'(t) = \boldsymbol{\phi}(\mathbf{x}(t), \boldsymbol{\psi}(t)), \quad \boldsymbol{\psi}(1) = \nabla C(\mathbf{x}(1)). \tag{2.30}$$

The same problem can be derived by solving (1.6) for $\mathbf{u}$ in terms of $(\mathbf{x}, \boldsymbol{\psi})$ and substituting in (1.4)–(1.5).

In order to make sure that the control approximations have the same order of accuracy as that of the discrete state and costate, we compute discrete controls $\mathbf{u}_n$, obtained by minimization of the Hamiltonian $\boldsymbol{\psi}_n \mathbf{f}(\mathbf{x}_n, \mathbf{u})$. In other words, we solve

$$\mathbf{u}_n \in U, \quad -\boldsymbol{\psi}_n \nabla_u \mathbf{f}(\mathbf{x}_n, \mathbf{u}_n) \in N_U(\mathbf{u}_n), \quad 0 \leq n \leq N, \tag{2.31}$$

for given pairs $(\mathbf{x}_n, \boldsymbol{\psi}_n)$.

Eventually, we would like to emphasize that there are essentially two main hypotheses in the analysis presented so far. The class of considered W-methods has to be restricted to those methods having weights $b_i > 0$ for $i = 1, \ldots, s$. Second, we have to assume sufficient smoothness of the optimal control problem, so that the Hamiltonian has a locally unique minimizer in the control and an equivalent, reduced scheme for state and costate can be established. This is in accordance with the analysis used by Hager [4] for Runge–Kutta discretizations.

3 Order conditions

In this section we shall derive order conditions for the discretization (2.23)–(2.28) to reach order two and three. Since the scheme does not fit into any classical form, we follow the general approach of substituting the continuous solution into the discrete equations, applying Taylor expansions and comparing the error terms with those obtained from the Taylor expansion of the exact solution.

Let $\mathbf{z}$ and $\boldsymbol{\delta}$ denote the following pairs:

$$\mathbf{z} = \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\psi} \end{pmatrix}, \quad \boldsymbol{\delta}(\mathbf{z}) = \begin{pmatrix} \mathbf{g}(\mathbf{z}) \\ \boldsymbol{\phi}(\mathbf{z}) \end{pmatrix}. \tag{3.1}$$

Then the system of differential equations (2.29)–(2.30) has the form $\mathbf{z}'(t) = \boldsymbol{\delta}(\mathbf{z}(t))$. The standard Taylor expansion for $\mathbf{z}(t)$ around $t = t_n$ reads

$$\mathbf{z}(t_{n+1}) = \mathbf{z}(t_n) + \boldsymbol{\delta} h + \frac{1}{2} \nabla_z \boldsymbol{\delta}\boldsymbol{\delta} h^2 + \frac{1}{6}\left(\nabla_z^2 \boldsymbol{\delta}\boldsymbol{\delta}^2 + \nabla_z \boldsymbol{\delta} \nabla_z \boldsymbol{\delta}\boldsymbol{\delta}\right) h^3 + \mathcal{O}(h^4), \tag{3.2}$$

where $\nabla_z \boldsymbol{\delta}$ is the Jacobian matrix of $\boldsymbol{\delta}$ with respect to $\mathbf{z}$ and $\nabla_z^2 \boldsymbol{\delta}$ denotes its Hessian tensor which operates on the pair $\boldsymbol{\delta}^2$ (to give a vector). The function $\boldsymbol{\delta}$ and all its derivatives are evaluated at $\mathbf{z}(t_n)$.

An analogous expansion can be derived for the numerical solution $\mathbf{z}_{n+1} = (\mathbf{x}_{n+1}, \boldsymbol{\psi}_{n+1})$ when the initial values $\mathbf{x}_n$ and $\boldsymbol{\psi}_n$ in (2.23)–(2.28) are replaced by the exact solutions $\mathbf{x}(t_n)$ and $\boldsymbol{\psi}(t_n)$. For given values $\mathbf{x}_n$ and $\boldsymbol{\psi}_n$, the intermediate values $\mathbf{y}_{ni}$, $\boldsymbol{\xi}_{ni}$ and $\mathbf{x}_{ni}$ are functions of the step size h. Substituting $\mathbf{y}_{ni}(h)$ in (2.23) gives

$$\mathbf{z}_{n+1}(h) = \mathbf{z}(t_n) + h\mathbf{G}(\mathbf{y}_{n1}(h), \boldsymbol{\xi}_{n1}(h), \mathbf{x}_{n1}(h), \ldots, \mathbf{y}_{ns}(h), \boldsymbol{\xi}_{ns}(h), \mathbf{x}_{ns}(h)), \tag{3.3}$$

where

$$\mathbf{G}(h) = \sum_{i=1}^{s} b_i \left(\boldsymbol{\delta}(\mathbf{x}_{ni}(h), \boldsymbol{\xi}_{ni}(h)) + \begin{pmatrix} T_n \sum_{j=1}^{i} \gamma_{ij} \mathbf{y}_{nj}(h) \\ 0 \end{pmatrix} \right). \tag{3.4}$$

Combining successive substitution of the intermediate values $\mathbf{y}_{ni}(h)$, $\boldsymbol{\xi}_{ni}(h)$ and $\mathbf{x}_{ni}(h)$ in $\mathbf{G}$ with Taylor expansions around $h = 0$, we have

$$\mathbf{z}_{n+1}(h) = \mathbf{z}(t_n) + \mathbf{C}_1 h + \mathbf{C}_2 h^2 + \mathbf{C}_3 h^3 + \mathcal{O}(h^4), \tag{3.5}$$

where the vector-valued coefficients $\mathbf{C}_i$ depend on the function $\boldsymbol{\delta}$, its first and second derivatives (all evaluated at $\mathbf{z}(t_n)$), the matrix T_n and its transpose, and the coefficients b_i, α_{ij}, and γ_{ij}. We say that the W-method (2.23)–(2.28) for the system (2.29)–(2.30) has the order p if the expansions (3.2) and (3.5) agree through terms of order h^p, i.e., $\mathbf{z}(t_{n+1}) - \mathbf{z}_{n+1}(h) = \mathcal{O}(h^{p+1})$.

Let us define

$$\beta_{ij} = \alpha_{ij} + \gamma_{ij}, \quad \beta_i = \sum_{j=1}^{i-1} \beta_{ij}, \quad c_i = \sum_{j=1}^{i-1} \alpha_{ij}, \tag{3.6}$$

$$\bar{\beta}_{ij} = \bar{\alpha}_{ij} + \bar{\gamma}_{ij}, \quad \bar{\beta}_i = \sum_{j=1}^{s} \bar{\beta}_{ij}, \quad \bar{c}_i = \sum_{j=1}^{s} \bar{\alpha}_{ij}. \tag{3.7}$$

As usual, we formally set $\beta_{ij} = 0$ for all $i \leq j$.

Following straightforward the approach described above to derive the expansion of the local error $\mathbf{z}(t_{n+1}) - \mathbf{z}_{n+1}(h)$, we can state (after a quite lengthy calculation).

Theorem 3.1 *The W-method* (2.23)–(2.28) *has order* $p = 1, 2,$ *or* 3, *if the order conditions of Table* 1 *are satisfied.*

Notice that, except the positivity requirement on the weights b_i, the order conditions $A1$–$A8$ are the usual order conditions associated with a W-method when applied to a system of ordinary differential equations [7]. As a consequence, any classical W-method of order $p = 2$ with strictly positive weights maintains its order for optimal control. Only at order $p = 3$, three new conditions emerge in the control context. Condition $A9$ yields together with $A2$ the additional order condition for Runge–Kutta methods of order $p = 3$ as found in [4]. Clearly, this reflects the fact that with $T_n = 0$ all explicit Runge–Kutta methods are covered. Conditions $A10$ and $A11$ guarantee order $p = 3$ for arbitrary matrices T_n.

4 Stability

Since we aim at handling stiff and even very stiff problems in (1.2), we would like to construct L-stable methods (see [7, Section IV.3], for a discussion). From Remark 2.1, we observe that in practical computations the discrete state and costate equations

Table 1 Order conditions of W-methods for optimal control. The summation is over each index, taking values from 1 to s

Order	Number	Order conditions
1	$A1$	$\sum b_i = 1, b_i > 0, i = 1, \ldots, s$
2	$A2$	$\sum b_i c_i = \frac{1}{2}$
	$A3$	$\sum b_i \beta_i = \frac{1}{2} - \gamma$
3	$A4$	$\sum b_i c_i^2 = \frac{1}{3}$
	$A5$	$\sum b_i \alpha_{ij} c_j = \frac{1}{6}$
	$A6$	$\sum b_i \alpha_{ij} \beta_j = \frac{1}{6} - \frac{\gamma}{2}$
	$A7$	$\sum b_i \beta_{ij} c_j = \frac{1}{6} - \frac{\gamma}{2}$
	$A8$	$\sum b_i \beta_{ij} \beta_j = \frac{1}{6} - \gamma + \gamma^2$
	$A9$	$\sum b_i \bar{c}_i^2 = \frac{1}{3}$
	$A10$	$\sum b_i \beta_i^2 = \frac{1}{3}$
	$A11$	$\sum b_i \bar{\beta}_i^2 = \frac{1}{3}$

are solved one after the other if iterates of the control variables are given. Thus it is reasonable to consider the famous Dahlquist test equation

$$x(t) \in \mathbb{R}^1 : x' = \lambda x, \quad x(0) = x_0, \quad \lambda \in \mathbb{C}, \quad Re(\lambda) < 0, \quad t > 0, \tag{4.1}$$

for stability investigations. As in [22], we follow classical stability concepts for W-methods and set $T_n = \lambda$, which is now a constant. The corresponding adjoint test equation acting backwards in time reads

$$\psi(t) \in \mathbb{R}^1 : \psi' = -\lambda\psi, \quad \psi(0) = \psi_0, \quad \lambda \in \mathbb{C}, \quad Re(\lambda) < 0, \quad t < 0, \tag{4.2}$$

where ψ_0 is given.

Let us introduce the notations

$$\mathbf{b}^T = (b_1, \ldots, b_s), \quad B = (\beta_{ij})_{i,j=1}^s, \quad z = \lambda h, \quad \mathbf{1}^T = (1, \ldots, 1) \in \mathbb{R}^s. \tag{4.3}$$

If we apply method (2.8)–(2.9) to the test equation (4.1) then the numerical solution becomes $x_{n+1} = R_x(z)x_n$ with the stability function

$$R_x(z) = 1 + z\mathbf{b}^T(I - zB)^{-1}\mathbf{1}. \tag{4.4}$$

Properties of such functions are well known from diagonally implicit Runge–Kutta methods (see e.g. [7, Section IV.6]). Applying method (2.10)–(2.11) to the test equation (4.2), we find $\psi_n = R_\psi(z)\psi_{n+1}$ with

$$R_\psi(z) = 1 + z\mathbf{1}^T(I - zB^T)^{-1}\mathbf{b}. \tag{4.5}$$

Since $(I - zB^T)^{-1} = ((I - zB)^{-1})^T$, the stability functions are equal, i.e., $R_x(z) = R_\psi(z)$. Thus it is sufficient to consider $R_x(z)$ defined by the discrete state solver.

Table 2 Regions of γ for L-stability with $a_s = 0$

s	$p = s - 1$	$p = s$
2	$1 - \frac{1}{2}\sqrt{2} \le \gamma \le 1 + \frac{1}{2}\sqrt{2}$	$\gamma = 1 \pm \frac{1}{2}\sqrt{2}$
4	$0.22364780 \le \gamma \le 0.57281606$	$\gamma = 0.57281606$

A W-method with $T_n = \lambda$ and stability function $R_x(z)$ is called A-stable if its stability domain $S = \{z \in \mathbb{C} : |R_x(z)| \le 1\}$ is a subset of the left complex half-plane $\mathbb{C}^- = \{z \in \mathbb{C} : Re(z) \le 0\}$. If in addition $R_x(-\infty) = 0$ then it is called L-stable. For W-methods of order p, $R_x(z)$ is a rational function which satisfies

$$e^z - R_x(z) = C\, z^{p+1} + \mathcal{O}(z^{p+2}) \quad \text{for} \;\; z \to 0, \tag{4.6}$$

where $C \neq 0$ is the error constant. Its form is given by

$$R_x(z) = \frac{P(z)}{(1-\gamma z)^s}, \quad P(z) = \det(I - zB + z\mathbf{1}\mathbf{b}^T), \tag{4.7}$$

where the numerator $P(z)$ is a polynomial of degree s at most. Let $P(z) = \sum_{i=0,\ldots,s} a_i z^i$. In order to have $R_x(-\infty) = 0$ for L-stability, the highest coefficient a_s of the numerator is set to zero, which can be ensured by a proper choice of the matrix B and the vector $\mathbf{b}$. Then, if the method has order $p \ge s - 1$, the remaining coefficients and the error constant in (4.6) are uniquely determined by γ and we have

$$a_i = (-1)^s L_s^{(s-i)}\left(\frac{1}{\gamma}\right)\gamma^i, \quad i = 0, \ldots, s-1, \quad C = (-1)^s L_s\left(\frac{1}{\gamma}\right)\gamma^s. \tag{4.8}$$

Here,

$$L_s(y) = \sum_{j=0}^{s} (-1)^j \binom{s}{j} \frac{y^j}{j!} \tag{4.9}$$

denotes the s-degree Laguerre polynomial and $L_s^{(k)}(y)$ its kth derivative. As a consequence, regions of L-stability and small error constants can now be determined by varying the parameter γ. For an overview of known results, we refer to Table 6.4 in [7]. For later use, we collect the corresponding γ-values for $s = 2, 4$ in Table 2.

Next we will describe a method of order 2, which belongs to a family of already known ROS2-methods, and construct a new method of order 3 for optimal control.

5 Construction of W-methods for optimal control

5.1 Second-order W-method

As stated above, any classical second-order W-methods with strictly positive weights is also suitable for optimal control. Let $s = 2$. Then method (2.23)–(2.28) is second-order consistent for any T_n iff

$$b_1 = 1 - b_2, \quad \gamma_{21} = -\frac{\gamma}{b_2}, \quad c_2 = \alpha_{21} = \frac{1}{2b_2}, \tag{5.1}$$

where $b_2 \in (0, 1)$ and γ are free parameters. We choose $\gamma = 1 - \sqrt{2}/2$ to get an L-stable method with a small error constant and select $b_2 = 1/2$ as proposed in [23] for ROS2.

5.2 Third-order W-method

From a practical point of view, we would like to have an as small as possible stage number s. The method's coefficients have to satisfy the 11 order conditions given in Table 1, besides a few restrictions on the stability parameter γ. Let us start with $s = 3$. In this case, we have 10 parameters to be chosen. Not surprising, there is only a negative result.

Theorem 5.1 *There is no third-order three-stage W-method* (2.23)–(2.28) *which satisfies the order conditions* $A1$–$A11$ *with* $\gamma \neq 0$.

Proof To prove this statement, it is sufficient to consider conditions $A5$–$A8$. They read

$$(A5) \quad b_3 c_2 \alpha_{32} = \frac{1}{6}, \quad (A6) \quad b_3 \alpha_{32} \beta_2 = \frac{1}{6} - \frac{\gamma}{2}, \tag{5.2}$$

$$(A7) \quad b_3 \beta_{32} c_2 = \frac{1}{6} - \frac{\gamma}{2}, \quad (A8) \quad b_3 \beta_{32} \beta_2 = \frac{1}{6} - \gamma + \gamma^2. \tag{5.3}$$

We compute α_{32} from $A5$ and substitute it in $A6$. This gives $\beta_2 = c_2(1 - 3\gamma)$. Then, from A8, we derive a condition for the product $b_3 \beta_{32} c_2$, which can be compared to that given in $A7$. Thus, we find

$$\left(\frac{1}{6} - \frac{\gamma}{2}\right)(1 - 3\gamma) = \frac{1}{6} - \gamma + \gamma^2. \tag{5.4}$$

This relation gives $\gamma = 0$ as unique solution. □

Hence, it is reasonable to look for a third-order W-method with $s = 4$. Now 17 parameters are available to fit all conditions. Our main design criteria are the following: (i) L-stability, i.e., $\gamma \in [0.22364780, 0.57281606]$ and $a_4 = 0$ (highest coefficient of the polynomial $P(z)$ in (4.7)), (ii) small error constant, and (iii) $c_i \in [0, 1]$, which is a desirable property for non-autonomous differential equations.

The condition $b_i > 0, i = 1, \ldots, 4$, appears to be quite restrictive for satisfying all desired criteria. Therefore, we follow the advise given in [17, Remark 4.13] for Runge–Kutta methods, and ensure positivity of the summarized weights that correspond to distinct values of the constants c_i. We set $\alpha_{21} = 0$, which gives $c_1 = c_2 = 0$, and request $b_1 + b_2 > 0$, $b_3 > 0$, and $b_4 > 0$. Since now $\mathbf{x}_{n2} = \mathbf{x}_{n1} = \mathbf{x}_n$, which are defined in (2.27), we also identify $\mathbf{u}_{n1}$ with $\mathbf{u}_{n2}$, yielding the control vector $\mathbf{u}_n = (\mathbf{u}_{n2}, \mathbf{u}_{n3}, \mathbf{u}_{n4})$. As a consequence, the first two relations for $i = 1, 2$, in (2.16) sum up to

Table 3 Coefficients for the L-stable third-order four-stage ROS3WO-method with $b_1+b_2>0$, $b_3>0$ and $b_4>0$

$\gamma = 0.22375933090210537159 0$	
$\alpha_{21} = 0.00000000000000000000$	$\gamma_{21} = 0.62304925695186060083 5$
$\alpha_{31} = 0.69884611483389190730 4$	$\gamma_{31} = -0.21681173383970731447 2$
$\alpha_{32} = -0.01079251169431481814 9$	$\gamma_{32} = -0.12438442037082067800 6$
$\alpha_{41} = -0.87576615372743954771 0$	$\gamma_{41} = 1.08299939965162189152 4$
$\alpha_{42} = -0.28471256637661401286 6$	$\gamma_{42} = 0.47765669465674627348 9$
$\alpha_{43} = 1.71139458518839102011 2$	$\gamma_{43} = -1.14882152187372163994 0$
$b_1 = 0.36190531683406064361 9$	$c_1 = 0.00000000000000000000$
$b_2 = -0.11680340160699614796 6$	$c_2 = 0.00000000000000000000$
$b_3 = 0.61335901969541743705 8$	$c_3 = 0.68805360313957708915 4$
$b_4 = 0.14153906507751806728 9$	$c_4 = 0.55091586508433745953 5$

$$\mathbf{u}_{n2} \in U, \quad -(b_1\boldsymbol{\xi}_{n1} + b_2\boldsymbol{\xi}_{n2})\nabla_u \mathbf{f}(\mathbf{x}_{n2}, \mathbf{u}_{n2}) \in N_U(\mathbf{u}_{n2}). \tag{5.5}$$

Introducing the new variable $\boldsymbol{\eta}_n = (b_1\boldsymbol{\xi}_{n1} + b_2\boldsymbol{\xi}_{n2})/(b_1+b_2)$ being an approximation of the costate $\boldsymbol{\psi}$ at t_n, the condition reads

$$\mathbf{u}_{n2} \in U, \quad -(b_1+b_2)\boldsymbol{\eta}_n \nabla_u \mathbf{f}(\mathbf{x}_{n2}, \mathbf{u}_{n2}) \in N_U(\mathbf{u}_{n2}). \tag{5.6}$$

Since $b_1+b_2>0$, the associated control $\mathbf{u}_{n2}$ is well defined by the control uniqueness property as local minimizer of the Hamiltonian $\boldsymbol{\psi}\mathbf{f}(\mathbf{x},\mathbf{u})$ with $\boldsymbol{\psi} = \boldsymbol{\eta}_n$ and $\mathbf{x} = \mathbf{x}_{n2}$.

Newton's method is applied to find appropriate roots of the system of nonlinear equations. The new W-method constructed along these principles is called ROS3WO, which is an abbreviation for *Ros*enbrock, *W*-method and *o*ptimal control. In Table 3, we give the method defining coefficients with 20-digit accuracy.

6 Numerical illustrations

Numerical results are given for optimal control problems, where the underlying ODE system ranges from linear and nonstiff to nonlinear and very stiff. We study (i) a nonstiff problem with known exact solution [4], (ii) the nonlinear Rayleigh problem [8], (iii) the stiff van der Pol oscillator, and (iv) a nonlinear boundary control problem for the heat equation with control constraints [3,9]. These types of problems are often used in optimal control benchmarking.

To report on numerically observed convergence orders, we perform a least square fit of the errors to a function of the form ch^p. The order thus obtained is denoted by p_{fit}.

6.1 A nonstiff problem

We first study a simple test problem from [4] to illustrate the convergence behaviour of classical explicit and implicit Runge–Kutta–Rosenbrock methods and our newly designed W-methods. Let us consider the following quadratic problem with a linear ODE given as constraint:

$$\text{Minimize } \frac{1}{2}\int_0^1 u(t)^2 + 2x(t)^2\, dt \tag{6.1}$$

$$\text{subject to } x'(t) = \frac{1}{2}x(t) + u(t), \quad t \in (0, 1], \tag{6.2}$$

$$x(0) = 1, \tag{6.3}$$

with the optimal solution

$$x^*(t) = \frac{2e^{3t} + e^3}{e^{3t/2}(2 + e^3)}, \quad u^*(t) = \frac{2(e^{3t} - e^3)}{e^{3t/2}(2 + e^3)}. \tag{6.4}$$

The first-order optimality system reads

$$x'(t) = \frac{1}{2}x(t) + u(t), \quad t \in (0, 1], \quad x(0) = 1, \tag{6.5}$$

$$\psi'(t) = -\frac{1}{2}\psi(t) - 2x(t), \quad t \in [0, 1), \quad \psi(1) = 0, \tag{6.6}$$

$$0 = u(t) + \psi(t). \tag{6.7}$$

That is, we have $u(t) = -\psi(t)$ and therefore the following boundary-value problem:

$$x'(t) = \frac{1}{2}x(t) - \psi(t), \quad x(0) = 1, \tag{6.8}$$

$$\psi'(t) = -\frac{1}{2}\psi(t) - 2x(t), \quad \psi(1) = 0. \tag{6.9}$$

Numerical results for the classical Runge–Kutta-methods RK3a, RK3b, RK4 (for example, see [4] and references therein), the fourth-order Rosenbrock method RODAS [7], ROS2 and ROS3WO are given in Tables 4 and 5. Only RK3a, RK4, ROS2, and ROS3WO fulfill the additional consistency conditions for optimal control and show their full order. In contrast, the control discretization of the explicit RK3b and the implicit RODAS drops down to second-order accuracy. This behaviour is typical for all Runge–Kutta and Rosenbrock methods that violate one of the new conditions.

We also varied the arbitrary matrix T_n, i.e., we used $T_n = 0$ (which yields the embedded explicit method), $T_n = 0.5$ (the exact Jacobian), and $T_n = 1.0$. In all cases, the full order is obtained for the state and control variables.

6.2 The nonlinear unconstrained Rayleigh problem

The following problem is taken from [8]. It describes the behaviour of a so-called tunnel-diode oscillator. The state variable is the electric current $x_1(t)$ at time $t \in [0, T]$

Table 4 Test problem 1: order of L^∞ convergence of the discrete state errors $\mathbf{x}(t_n) - \mathbf{x}_n$, $n = 0, \ldots, N$, for classical Runge–Kutta and Rosenbrock methods, ROS2 and ROS3WO applied to solve (6.8)–(6.9). The exact Jacobian is $T_n = 0.5$

N	10	20	40	80	160	p_{fit}
RK3a	8.82e−5	9.72e−6	1.11e−6	1.32e−7	1.61e−8	3.10
RK3b	7.24e−4	1.73e−4	4.23e−5	1.05e−5	2.60e−6	2.03
RK4	5.98e−6	3.85e−7	2.44e−8	1.54e−9		3.98
RODAS	9.93e−4	2.65e−4	6.82e−5	1.73e−5	4.35e−6	1.96
ROS2, $T_n = 0$	2.96e−3	7.23e−4	1.78e−4	4.42e−5	1.10e−5	2.02
ROS2, $T_n = 0.5$	2.60e−3	6.16e−4	1.50e−4	3.68e−5	9.13e−6	2.04
ROS2, $T_n = 1$	2.38e−3	5.43e−4	1.29e−4	3.15e−5	7.77e−6	2.06
ROS3WO, $T_n = 0$	5.78e−5	8.39e−6	1.12e−6	1.45e−7	1.84e−8	2.91
ROS3WO, $T_n = 0.5$	6.53e−5	8.80e−6	1.14e−6	1.44e−7	1.82e−8	2.95
ROS3WO, $T_n = 1$	1.05e−4	1.29e−5	1.60e−6	1.98e−7	2.47e−8	3.01

Table 5 Test problem 1: order of L^∞ convergence of the discrete control errors $\mathbf{u}(t_n) - \mathbf{u}_n$, $n = 0, \ldots, N$, for classical Runge–Kutta and Rosenbrock methods, ROS2 and ROS3WO applied to solve (6.8)–(6.9). The exact Jacobian is $T_n = 0.5$

N	10	20	40	80	160	p_{fit}
RK3a	2.06e−4	2.78e−5	3.58e−6	4.52e−7	5.68e−8	2.96
RK3b	3.65e−3	9.59e−4	2.46e−4	6.21e−5	1.56e−5	1.97
RK4	2.02e−6	1.37e−7	8.82e−9	5.58e−10		3.94
RODAS	6.08e−3	1.50e−3	3.74e−4	9.32e−5	2.33e−5	2.01
ROS2, $T_n = 0$	2.11e−3	6.09e−4	1.63e−4	4.21e−5	1.07e−5	1.91
ROS2, $T_n = 0.5$	1.90e−3	5.12e−4	1.32e−4	3.37e−5	8.49e−6	1.95
ROS2, $T_n = 1$	1.49e−3	3.75e−4	9.41e−5	2.35e−5	5.89e−6	2.00
ROS3WO, $T_n = 0$	5.00e−5	4.97e−6	5.35e−7	6.14e−8	7.33e−9	3.18
ROS3WO, $T_n = 0.5$	9.18e−5	9.49e−6	1.05e−6	1.23e−7	1.48e−8	3.15
ROS3WO, $T_n = 1$	1.84e−4	1.94e−5	2.20e−6	2.60e−7	3.16e−8	3.12

and the control $u(t)$ is a transformed voltage at the generator. The unconstrained Rayleigh problem is defined as follows:

$$\text{Minimize} \int_0^T u(t)^2 + x_1(t)^2 \, dt \tag{6.10}$$

$$\text{subject to } x_1''(t) = -x_1(t) + x_1'(1.4 - 0.14 x_1'(t)^2) + 4u(t), \quad t \in (0, T], \tag{6.11}$$

$$x_1(0) = x_1'(0) = -5. \tag{6.12}$$

The ODE is of second order and nonlinear. To transform this problem to our setting, we introduce $x_2(t) = x_1'(t)$ and the additional equation $x_3'(t) = u(t)^2 + x_1(t)^2$ with the initial value $x_3(0) = 0$. This gives the new formulation

$$\begin{aligned}
&\text{Minimize } \; x_3(T) && (6.13)\\
&\text{subject to } \; x_1'(t) = x_2(t), && (6.14)\\
&x_2'(t) = -x_1(t) + x_2(1.4 - 0.14x_2(t)^2) + 4u(t), && (6.15)\\
&x_3'(t) = u(t)^2 + x_1(t)^2, \quad t \in (0, T], && (6.16)\\
&x_1(0) = -5, \quad x_2(0) = -5, \quad x_3(0) = 0. && (6.17)
\end{aligned}$$

As final time we set $T = 2.5$.

Computing the gradients of the right hand side in (6.14)–(6.16) with respect to $\mathbf{x}$ and $\mathbf{u}$, the adjoint equations and the condition for the control can be easily derived. We find

$$\begin{aligned}
\psi_1'(t) &= \psi_2(t) - 2x_1(t)\psi_3(t), && (6.18)\\
\psi_2'(t) &= -\psi_1(t) - (1.4 - 0.42x_2(t)^2)\psi_2(t), && (6.19)\\
\psi_3'(t) &= 0, && (6.20)\\
\psi_1(T) &= 0, \quad \psi_2(T) = 0, \quad \psi_3(T) = 1, && (6.21)\\
0 &= 4\psi_2(t) + 2u(t)\psi_3(t), \quad t \in (0, T]. && (6.22)
\end{aligned}$$

We get the trivial solution $\psi_3(t) \equiv 1$. The control is then computed from (6.22), which yields $u(t) = -2\psi_2(t)$. We can separate Eq. (6.16) for $x_3(t)$, which only serves to compute the objective function, from the set of ordinary differential equations and eliminate the control in the first order optimality conditions. This finally gives the following nonlinear boundary value problem in $[0, T]$:

$$\begin{aligned}
x_1'(t) &= x_2(t), && (6.23)\\
x_2'(t) &= -x_1(t) + x_2(1.4 - 0.14x_2(t)^2) - 8\psi_2(t), && (6.24)\\
x_1(0) &= -5, \quad x_2(0) = -5, && (6.25)\\
\psi_1'(t) &= \psi_2(t) - 2x_1(t), && (6.26)\\
\psi_2'(t) &= -\psi_1(t) - (1.4 - 0.42x_2(t)^2)\psi_2(t), && (6.27)\\
\psi_1(T) &= 0, \quad \psi_2(T) = 0. && (6.28)
\end{aligned}$$

To study convergence orders of our W-methods, we computed a reference solution by applying the classical fourth-order RK4 with $N = 320$. In our numerical tests, we chose for T_n the zero matrix, the exact Jacobian and a partitioned matrix that treats the first state variable implicitly and the second one explicitly. More precisely, we used

$$T_{1,n} = 0, \quad T_{2,n} = \begin{pmatrix} 0 & 1 \\ -1 & 1.4 - 0.42x_{2,n}^2 \end{pmatrix}, \quad T_{3,n} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad 0 \le n \le N-1.$$

Table 6 Rayleigh problem: order of L^∞ convergence of the discrete state errors $x_i(t_n)-x_{i,n}$, $i=1,2$, $n=0,\dots,N$, and the discrete control errors $u(t_n)-u_n$, $n=0,\dots,N$, for ROS2 applied to solve (6.23)–(6.28)

N	20	40	80	160	320	p_{fit}
ROS2, $T_n = T_{1,n}$						
1st state variable	2.23e−1	6.28e−2	1.27e−2	2.90e−3	6.98e−4	2.11
2nd state variable	6.59e−1	1.62e−1	3.12e−2	7.08e−3	1.71e−3	2.17
Control variable	2.28e−0	3.46e−1	4.82e−2	1.03e−2	2.46e−3	2.48
ROS2, $T_n = T_{2,n}$						
1st state variable	5.60e−2	3.41e−2	8.99e−3	2.20e−3	5.43e−4	1.73
2nd state variable	3.94e−1	1.50e−1	3.73e−2	9.10e−3	2.25e−3	1.89
Control variable	2.05e−0	4.74e−1	8.89e−2	1.85e−2	4.20e−3	2.25
ROS2, $T_n = T_{3,n}$						
1st state variable	2.19e−1	6.17e−2	1.24e−2	2.82e−3	6.78e−4	2.11
2nd state variable	6.47e−1	1.59e−1	3.06e−2	6.93e−3	1.67e−3	2.17
Control variable	2.27e−0	3.42e−1	4.69e−2	1.01e−2	2.42e−3	2.48

Table 7 Rayleigh problem: order of L^∞ convergence of the discrete state errors $x_i(t_n)-x_{i,n}$, $i=1,2$, $n=0,\dots,N$, and the discrete control errors $u(t_n)-u_n$, $n=0,\dots,N$, for ROS3WO applied to solve (6.23)–(6.28)

N	20	40	80	160	320	p_{fit}
ROS3WO, $T_n = T_{1,n}$						
1st state variable	7.69e−1	2.52e−2	1.13e−3	1.01e−4	1.06e−5	4.02
2nd state variable	4.33e−0	8.35e−2	2.96e−3	2.46e−4	2.54e−5	4.32
Control variable	9.10e−0	4.40e−1	1.63e−2	1.30e−3	1.31e−4	4.06
ROS3WO, $T_n = T_{2,n}$						
1st state variable	1.85e−2	3.03e−3	3.83e−4	4.63e−5	5.46e−6	2.95
2nd state variable	1.54e−2	3.26e−3	4.15e−4	4.82e−5	5.42e−6	2.90
Control variable	4.95e−1	4.86e−2	4.61e−3	4.87e−4	5.45e−5	3.29
ROS3WO, $T_n = T_{3,n}$						
1st state variable	7.76e−1	2.60e−2	1.15e−3	1.01e−4	1.07e−5	4.03
2nd state variable	4.38e−0	8.64e−2	3.04e−3	2.51e−4	2.59e−5	4.32
Control variable	9.10e−0	4.54e−1	1.67e−2	1.33e−3	1.34e−4	4.05

Numerical results for ROS2 and ROS3WO are given in Tables 6 and 7. They clearly show orders close to two and three independently from the choice of the matrix T_n as predicted by the theory. The better order four for ROS3WO in the case of inexact Jacobian matrices results from a relatively huge improvement in the first two refinement steps. The last three values are close to order three.

6.3 The stiff van der Pol oscillator

Our third example is an optimal control problem for the van der Pol oscillator, which is considered in the stiff region. The unconstrained problem reads as follows:

$$\text{Minimize } \int_0^T u(t)^2 + x(t)^2 + x'(t)^2\,dt \tag{6.29}$$

$$\text{subject to } \varepsilon x''(t) - (1 - x(t)^2)x'(t) + x(t) = u(t), \quad t \in (0, T], \tag{6.30}$$

$$x(0) = 0, \quad x'(0) = 2. \tag{6.31}$$

Small positive values of ε give rise to extremely steep profiles in $x(t)$, making the van der Pol equation a challenging test example for any ODE integrator [7]. The control $u(t)$ is used to smooth the solution again. We introduce Lienhard's coordinates $x_2(t) = x(t), x_1(t) = \varepsilon x'(t) + x(t)^3/3 - x(t)$, and the variable $x_3(t)$ through the ordinary differential equation $x_3'(t) = u(t)^2 + x(t)^2 + x'(t)^2$ with initial value $x_3(0) = 0$, to derive the following first order setting:

$$\text{Minimize } x_3(T) \tag{6.32}$$

$$\text{subject to } x_1'(t) = -x_2(t) + u(t), \tag{6.33}$$

$$x_2'(t) = \frac{1}{\varepsilon}\left(x_1(t) + x_2(t) - \frac{x_2(t)^3}{3}\right), \tag{6.34}$$

$$x_3'(t) = \frac{1}{\varepsilon^2}\left(x_1(t)+x_2(t)-\frac{x_2(t)^3}{3}\right)^2 + x_2(t)^2 + u(t)^2, \quad t \in (0, T], \tag{6.35}$$

$$x_1(0) = 2\varepsilon, \quad x_2(0) = 0, \quad x_3(0) = 0. \tag{6.36}$$

We defined $T = 2$ as final time and considered the case $\varepsilon = 0.01$.

Applying the approach described above and eliminating the control and the auxiliary variable $x_3(t)$ and its adjoint, we finally get the following nonlinear boundary value problem in $[0, T]$ for the state and costate variables:

$$x_1'(t) = -x_2(t) - \frac{\psi_1(t)}{2}, \tag{6.37}$$

$$x_2'(t) = \frac{1}{\varepsilon}\left(x_1(t) + x_2(t) - \frac{x_2(t)^3}{3}\right), \tag{6.38}$$

$$x_1(0) = 2\varepsilon, \quad x_2(0) = 0, \tag{6.39}$$

$$\psi_1'(t) = -\frac{1}{\varepsilon}\psi_2(t) - \frac{2}{\varepsilon^2}\left(x_1(t) + x_2(t) - \frac{x_2(t)^3}{3}\right), \tag{6.40}$$

Table 8 Van der Pol oscillator: order of L^∞ convergence of the discrete state errors $x_i(t_n) - x_{i,n}, i = 1, 2, n = 0, \dots, N$, and the discrete control errors $u(t_n) - u_n, n = 0, \dots, N$, for ROS2 applied to solve (6.37)–(6.42)

N	160	320	640	1,280	2,560	p_{fit}
ROS2, $T_n = T_{1,n}$						
1st state variable	6.30e−3	1.59e−3	3.73e−4	8.74e−5	2.03e−5	2.07
2nd state variable	6.24e−3	1.59e−3	3.73e−4	8.79e−5	2.05e−5	2.07
Control variable	4.62e−1	1.06e−1	2.44e−2	5.65e−3	1.31e−3	2.12
ROS2, $T_n = T_{2,n}$						
1st state variable	6.27e−3	1.59e−3	3.70e−4	8.67e−5	2.01e−5	2.08
2nd state variable	6.21e−3	1.58e−3	3.71e−4	8.72e−5	2.03e−5	2.07
Control variable	4.64e−1	1.05e−1	2.42e−2	5.59e−3	1.30e−3	2.12

$$\begin{aligned}\psi_2'(t) &= \psi_1(t) - \frac{1}{\varepsilon}(1 - x_2(t)^2)\psi_2(t)\\ &\quad -\frac{2}{\varepsilon^2}\left(x_1(t) + x_2(t) - \frac{x_2(t)^3}{3}\right)(1 - x_2(t)^2) - 2x_2(t), \qquad (6.41)\\ \psi_1(T) &= 0, \quad \psi_2(T) = 0. \qquad (6.42)\end{aligned}$$

For later use in our convergence study, we note that $u(t) = -0.5\psi_1(t)$. Since the factor ε^{-2} appears in the adjoint equations, this system is even stiffer and hence harder to solve than the original van der Pol equation. Due to the stiffness, an explicit integrator as RK4 works no longer efficiently.

We computed a reference solution by applying ROS3WO with $N = 2{,}560$. To test the robustness with respect to the choice of the matrix T_n, we considered the exact Jacobian and a partitioned matrix that treats the first equation explicitly and the second one implicitly. More precisely, we used

$$T_{1,n} = \begin{pmatrix} 0 & -1 \\ \varepsilon^{-1} & \varepsilon^{-1}(1 - x_{2,n}^2) \end{pmatrix}, \quad T_{2,n} = \begin{pmatrix} 0 & 0 \\ \varepsilon^{-1} & \varepsilon^{-1}(1 - x_{2,n}^2) \end{pmatrix}, \quad 0 \le n \le N - 1.$$

Numerical results for ROS2 and ROS3WO are given in Tables 8 and 9. In accordance to the theory, ROS2 clearly shows orders close to two. The observed order for ROS3WO is slightly better than three independently from the choice of the matrix T_n.

6.4 Nonlinear boundary control for the heat equation

For a practical illustration, we consider the nonlinear boundary control problem

$$\text{minimize } \frac{1}{2}\int_0^1 \left(x(y, T) - \frac{1}{2}(1 - y^2)\right)^2 dy + \frac{\lambda}{2}\int_0^T u(t)^2\, dt \qquad (6.43)$$

Table 9 Van der Pol oscillator: order of L^∞ convergence of the discrete state errors $x_i(t_n) - x_{i,n}, i = 1, 2, n = 0, \dots, N$, and the discrete control errors $u(t_n) - u_n, n = 0, \dots, N$, for ROS3WO applied to solve (6.37)–(6.42)

N	160	320	640	1,280	p_{fit}
ROS3WO, $T_n = T_{1,n}$					
1st state variable	1.47e−2	1.02e−3	1.01e−4	9.27e−6	3.52
2nd state variable	1.46e−2	1.01e−3	1.00e−4	9.17e−6	3.52
Control variable	1.35e−0	9.29e−2	9.08e−3	8.18e−4	3.54
ROS3WO, $T_n = T_{2,n}$					
1st state variable	1.48e−2	1.02e−3	1.01e−4	9.31e−6	3.53
2nd state variable	1.48e−2	1.02e−3	1.01e−4	9.20e−6	3.53
Control variable	1.36e−0	9.26e−2	9.06e−3	8.18e−4	3.54

subject to the heat equation with nonlinear boundary conditions of Stefan–Boltzmann type

$$\begin{aligned} \partial_t x(y,t) - \partial_{yy} x(y,t) &= 0, \qquad (y,t) \in (0,1) \times (0,T], && (6.44)\\ \partial_y x(0,t) &= 0, \qquad t \in (0,T], && (6.45)\\ \partial_y x(1,t) + x(1,t) + x^4(1,t) &= u(t), \quad t \in (0,T], && (6.46)\\ x(y,0) &= 0, \qquad y \in [0,1], && (6.47) \end{aligned}$$

and the box constraints for the control,

$$-0.5 \le u(t) \le 0.5, \quad \text{for almost all } \ t \in [0,T]. \tag{6.48}$$

We considered this problem for final time $T = 1.58$ and regularization parameter $\lambda = 0.1$ as stated in [9] (see also [3] for theoretical aspects). Standard second order finite differences on an equidistant mesh $y_i = i\triangle y, i = 0, \dots, M$, with $\triangle y = 1/M$ and M being a natural number, are used to discretize the nonlinear heat equation in space, which gives approximations $x_{i+1}(t) \approx x(y_i, t), i = 0, \dots, M$. Approximating the spatial integral of the objective function by the linear interpolating spline associated with the spatial mesh, and introducing an additional component $x_{M+2}(t)$ to transform the remaining control term, we get the following optimal control problem:

$$\begin{aligned} \text{Minimize } \ C(\mathbf{x}(T)) &= \frac{1}{2}(\mathbf{x}(T) - \mathbf{x}_y)^T M_y (\mathbf{x}(T) - \mathbf{x}_y) + x_{M+2}(T) && (6.49)\\ \text{subject to } \ \mathbf{x}'(t) &= A_y \mathbf{x}(t) + G_y(\mathbf{x}(t), u(t)), \quad t \in (0,T], && (6.50)\\ \mathbf{x}(0) &= 0, && (6.51) \end{aligned}$$

where $\mathbf{x}_y = \frac{1}{2}(1-y_0^2, \ldots, 1-y_M^2, 0)^T$ and

$$M_y = \frac{\triangle y}{6}\begin{pmatrix} 2 & 1 & & & & \\ 1 & 4 & 1 & & & \\ & & \ddots & & & \\ & & 1 & 4 & 1 & \\ & & & 1 & 2 & \\ & & & & & 0 \end{pmatrix},$$

$$A_y = \frac{1}{(\triangle y)^2}\begin{pmatrix} -2 & 2 & & & & \\ 1 & -2 & 1 & & & \\ & & \ddots & & & \\ & & 1 & -2 & 1 & \\ & & & 2 & -2 & \\ & & & & & 0 \end{pmatrix},$$

as well as

$$(G_y)_i = \begin{cases} 0, & i = 1, \ldots, M, \\ \frac{2}{\triangle y}(u(t) - x_{M+1} - x_{M+1}^4), & i = M+1, \\ \frac{\lambda}{2}u(t)^2, & i = M+2. \end{cases}$$

The dimension of the ODE system is $d = M+2$. We set $M = 400$ to keep spatial discretization errors small with respect to the overall error.

We discretized the optimal control problem using the methods ROS2, ROS3WO and GRK4A [10]. The latter method is a classical four-stage fourth-order Rosenbrock solver with strictly positive weights suitable for stiff equations, but it does not fulfill the additional order conditions for optimal control. The exact Jacobian was used for the matrix T_n, i.e.,

$$T_n = A_y + \text{diag}\left(0, \ldots, 0, -\frac{2}{\triangle y}(1 + 4x_{M+1,n}^3), 0\right), \quad 0 \le n \le N-1.$$

The discrete first-order optimality system (2.8)–(2.13) was solved using the source code for ASA_CG, Version 1.3, based on CG_DESCENT [6]. ASA_CG is an active set algorithm for solving bound constrained optimization problems [5]. We checked the results with those obtained by the DONLP2 software package [18]. In DONLP2, a sequential quadratic programming with an active set strategy and only equality constrained subproblems is implemented [19,20]. Both, ASA_CG and DONLP2, gave similar results for a gradient tolerance 1.0e−11.

To apply the optimization routines, we have to provide the value and the gradient of the reduced objective function $\hat{C}(\mathbf{u}) = C(\mathbf{x}_N(\mathbf{u}))$ and the control constraints. Given a vector $\mathbf{u}$, the final state vector $\mathbf{x}_N(\mathbf{u})$ is derived from the discrete state equations (2.8)–(2.9) by marching forward from $n = 0$ to $n = N-1$. Within each time step, the stage variables $\mathbf{y}_{ni}$, $i = 1, \ldots, s$, can be computed one after another by solving linear systems with one and the same (tridiagonal) matrix $I - h\gamma T_n$. Then all variables are given to solve the discrete costate equations (2.10)–(2.11) for $\boldsymbol{\psi}_n$ and $\boldsymbol{\lambda}_{ni}$ by marching backward from $n = N - 1$ to $n = 0$. Again, the intermediate values $\boldsymbol{\lambda}_{ni}$, $i = s, \ldots, 1$, are successively computable by solving a sequence of linear systems with the matrix $I - h\gamma T_n^T$ within each time step. The gradient of the objective function for ROS2 and GRK4A is determined by the following expressions:

$$\nabla_{u_{ni}} \hat{C}(\mathbf{u}) = h\boldsymbol{\lambda}_{ni} \nabla_u \mathbf{f}(\mathbf{x}_{ni}, \mathbf{u}_{ni}), \quad i = 1, \ldots, s. \tag{6.52}$$

For ROS3WO we have with the control vector $\mathbf{u}_n = (\mathbf{u}_{n2}, \mathbf{u}_{n3}, \mathbf{u}_{n4})$,

$$\nabla_{u_{n2}} \hat{C}(\mathbf{u}) = h(\boldsymbol{\lambda}_{n1} + \boldsymbol{\lambda}_{n2}) \nabla_u \mathbf{f}(\mathbf{x}_{n2}, \mathbf{u}_{n2}), \tag{6.53}$$

$$\nabla_{u_{ni}} \hat{C}(\mathbf{u}) = h\boldsymbol{\lambda}_{ni} \nabla_u \mathbf{f}(\mathbf{x}_{ni}, \mathbf{u}_{ni}), \quad i = 3, 4. \tag{6.54}$$

Here $\mathbf{f}$ is the right hand side in the ODE system (6.50).

For comparison purposes, we computed a reference solution with the exact Jacobian for $N = 800$, from which we derived the reference value for the objective function, $C_{\text{ref}} = 0.02319494$. All methods converge to this value. The corresponding optimal control is plotted in Fig. 1.

The gradient of the reduced version of the objective function in (6.49) can be computed from $\nabla_u \hat{C}(u) = -2\psi_{N+1}/\triangle y - \psi_{N+2}\lambda u$. Since $\psi_{N+2} \equiv 1$, the optimal control satisfies the projection relation

$$u(t) = \mathbb{P}_{[-0.5,0.5]}\{-2\psi_{N+1}(t)/(\lambda \triangle y\}. \tag{6.55}$$

The piecewise linear continuous approximation of $-2\psi_{M+1}(t)/(\lambda\triangle y)$ using numerical approximations of the values $\psi_{M+1}(t_n)$ at the time points, is also shown in Fig. 1. Outside the active region of the control constraints, it fits the numerical approximation of the control very well.

Numerical results for the time integrators tested are given in Fig. 2. ROS3WO converges to the reference solution faster than the other methods. ROS2 performs also remarkably well and even much better than GRK4A. We have also tested various approximations T_n of the Jacobian. Although the absolute values are worse, convergence is maintained for ROS2 and ROS3WO. Not surprisingly, the Rosenbrock solver GRK4A gives unsatisfactory results in this case due to its loss of consistency.

7 Summary and main conclusions

We have developed and discussed W-methods of linearly implicit structure for the numerical approximation of optimal control problems within the *first-discretize-then-optimize* approach. Following the concept of transformed adjoint equations, which was introduced in [4] for Runge–Kutta methods, we analyzed the approximation order and derived novel order conditions that have to be satisfied by the coefficients of the W-method so that the Taylor expansions of the continuous and discrete state and costate solutions match to order three. On the basis of this analysis, two main conclusions can

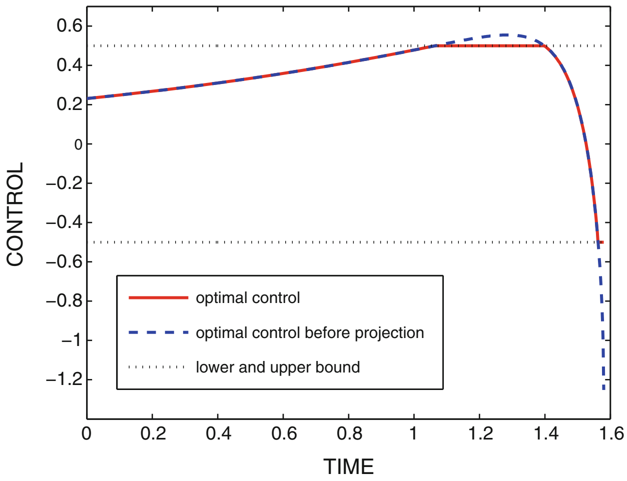

Fig. 1 Nonlinear heat equation: reference optimal control computed with 401 equidistant spatial points and 800 uniform time steps, and piecewise linear continuous approximation of $-2\psi_{M+1}(t)/(\lambda\triangle y)$ using numerical approximations of the values $\psi_{M+1}(t_n)$ at the time points. The control constraints are active in two regions

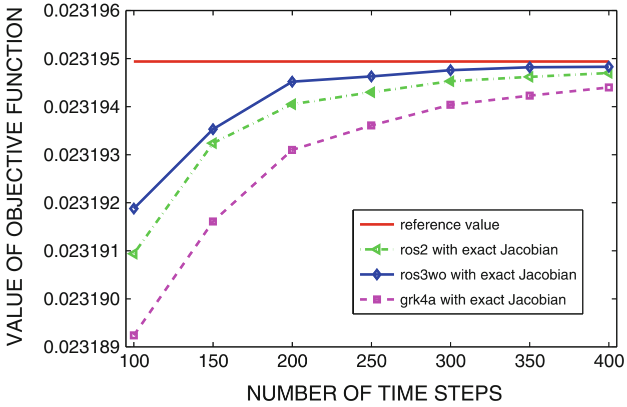

Fig. 2 Nonlinear heat equation: comparison of different time integrators, which are applied to solve (6.49)–(6.51). The methods tested are ROS2, GRK4A and ROS3WO. Exact Jacobian is used. Values of the discrete objective function for different numbers of time steps, $N = 100, 150, \ldots, 400$, are shown. The reference value is $C_{\text{ref}} = 0.02319494$

be drawn: (i) any classical W-method of second order with strictly positive weights maintains its order for optimal control. (ii) For order three, three additional order conditions have to be fulfilled. These conditions include the one already found in [4] for Runge–Kutta methods. There is no implicit third-order three-stage W-method suitable for optimal control.

As base integrators for comparisons, we have taken an L-stable two-stage W-method of second order from the ROS2 family [23] and have constructed a novel L-stable four-stage W-method ROS3WO of third-order. Both methods and other selected Runge–Kutta and Rosenbrock methods were applied to four example problems, ranging from linear and nonstiff to nonlinear and stiff. A semi-discretized nonlinear heat equation was considered to demonstrate the use of the developed W-methods in numerical optimization techniques that require the gradient of the discrete objective functional. From our numerical experience, we have come to two main conclusions. (i) All methods tested show their theoretical orders when they are applied to solve the two-point boundary-value problem (2.29)–(2.30), which is derived from the first-order optimality system. The W-methods are remarkably robust with respect to varying approximations of the Jacobian matrix. This allows for partitioning to treat stiff and nonstiff components more efficiently in the linear algebra. One even could set the Jacobian equal to zero and mimic an explicit method without loosing the order. (ii) Most notable for the W-methods is their structural advantage when they are applied within a gradient approach to solve state and costate equations separately. Only a sequence of linear equations with one and the same system matrix has to be solved to compute the stages values. We expect that this property will become even more important for the numerical solution of large scale PDE-constrained optimal control problems.

Acknowledgments J. Lang gratefully acknowledge the support of the DFG Priority Program 1253 entitled Optimization with Partial Differential Equations. We also want to thank Peter Spellucci for making his code DONLP2 available to us and for assisting us to find suitable coefficients for our ROS3WO method. I thank Jan for his long-time warm friendship and inspiring scientific collaboration. He will be greatly missed.

References

1. Bonnans, J.F., Laurent-Varin, J.: Computation of order conditions for symplectic partitioned Runge–Kutta schemes with application to optimal control. Numer. Math. **103**, 1–10 (2006)
2. Büttner, M., Schmitt, B.A., Weiner, R.: W-methods with automatic partitioning by Krylov techniques for large stiff systems. SIAM J. Numer. Anal. **32**, 260–284 (1995)
3. Dhamo, V., Tröltzsch, F.: Some aspects of reachability for parabolic boundary control problems with control constraints. Comput. Optim. Appl. **50**, 75–110 (2011)
4. Hager, W.W.: Runge–Kutta methods in optimal control and the transformed adjoint system. Numer. Math. **87**, 247–282 (2000)
5. Hager, W.W., Zhang, H.: A new active set algorithm for box constrained optimization. SIAM J. Optim. **17**, 526–557 (2006)
6. Hager, W.W., Zhang, H.: Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent. ACM Trans. Math. Softw. **32**, 113–137 (2006)
7. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Equations, 2nd revised edn. Springer, Berlin (1996)
8. Jacobson, D.H., Mayne, D.Q.: Differential Dynamic Programming. American Elsevier Publishing, New York (1970)
9. Kammann, E.: Modellreduktion und Fehlerabschätzung bei parabolischen Optimalsteuerungsproblemen. Diploma thesis, Department of Mathematics, Technische Universität Berlin (2010)
10. Kaps, P., Rentrop, P.: Generalized Runge–Kutta methods of order four with stepsize control for stiff ordinary differential equations. Numer. Math. **33**, 55–68 (1979)
11. Lang, J.: Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems. Theory, Algorithm and Applications. Lecture Notes in Computational Science and Engineering, vol. 16. Springer, Berlin (2000)
12. Murua, A.: On order conditions for partitioned symplectic methods. SIAM J. Numer. Anal. **34**, 2204–2211 (1997)
13. Pulova, N.V.: Runge–Kutta Schemes in Control Constrained Optimal Control. Lecture Notes in Computer Science, LNCS, vol. 4818, pp. 358–365 (2008)

14. Rosenbrock, H.H.: Some general implicit processes for the numerical solution of differential equations. Comput. J. **5**, 329–331 (1963)
15. Sandu, A.: On the Properties of Runge–Kutta Discrete Adjoints. Lecture Notes in Computer Science, LNCS, vol. 3394, pp. 550–557 (2006)
16. Sandu, A.: On consistency properties of discrete adjoint linear multistep methods. Report TR-07-40, Computer Science Department, Virginia Polytechnical Institute and State University (2007)
17. Schwartz, A., Polak, E.: Consistent approximations for optimal control problems based on Runge–Kutta integration. SIAM J. Control Optim. **34**, 1235–1269 (1996)
18. Spellucci, P.: Donlp2-intv-dyn users guide. Version November 18, 2009
19. Spellucci, P.: A new technique for inconsistent QP problems in the SQP method. Math. Methods Oper. Res. **47**, 355–400 (1998)
20. Spellucci, P.: An SQP method for general nonlinear programs using only equality constrained subproblems. Math. Program. **82**, 413–448 (1998)
21. Steihaug, T., Wolfbrandt, A.: An attempt to avoid exact Jacobian and nonlinear equations in the numerical solution of stiff ordinary differential equations. Math. Comput. **33**, 521–534 (1979)
22. Strehmel, K., Weiner, R.: Linear-implizite Runge–Kutta-Methoden und ihre Anwendungen. Teubner-Texte zur Mathematik, Bd. 127, Teubner (1992)
23. Verwer, J.G., Spee, E.J., Blom, J.G., Hundsdorfer, W.H.: A second order Rosenbrock method applied to photochemical dispersion problems. SIAM J. Sci. Comput. **20**, 1456–1480 (1999)
24. Walther, A.: Automatic differentiation of explicit Runge–Kutta methods for optimal control. Comput. Optim. Appl. **36**, 83–108 (2007)