INEXACT RESTORATION FOR RUNGE–KUTTA DISCRETIZATION
OF OPTIMAL CONTROL PROBLEMS∗

C. YALÇIN KAYA†

Abstract. A numerical method is presented for Runge–Kutta discretization of unconstrained
optimal control problems. First, general Runge–Kutta discretization is carried out to obtain a finite-
dimensional approximation of the continuous-time optimal control problem. Then a recent optimization
technique, the inexact restoration (IR) method, due to Martínez and coworkers [E. G. Birgin
111 (2001), pp. 39–58], is applied to the discretized problem to find an approximate solution. It
is proved that, for optimal control problems, a key sufficiency condition for convergence of the IR
method is readily satisfied. Under reasonable assumptions, the IR method for optimal control prob-
lems is shown to converge to a solution of the discretized problem. Convergence of a solution of
the discretized problem to a solution of the continuous-time problem is also shown. It turns out
that optimality phase equations of the IR method emanate from an associated Hamiltonian system,
and so general Runge–Kutta discretization induces a symplectic partitioned Runge–Kutta scheme.
A computational algorithm is described, and numerical experiments are made to demonstrate the
working of the method for optimal control of the van der Pol system, employing the three-stage
(order 6) Gauss–Legendre discretization.

Key words. optimal control, inexact restoration, Runge–Kutta discretization, symplectic par-
titioned Runge–Kutta scheme, Lagrange multiplier update, costate update, van der Pol system

AMS subject classifications. 49K15, 49M05, 49M25, 65K10, 65L06

DOI. 10.1137/090766668

1. Introduction. It is common practice to discretize a continuous-time optimal
control problem in order to perform computations and to find at least an approximate
solution. One typically discretizes (i.e., partitions) the time horizon of the optimal
control problem and computes optimal values of the states and controls at the discrete
time points by applying a standard finite-dimensional optimization method. In this
approach, three important issues need to be considered, namely,

(i) selection of the discretization scheme,
(ii) convergence of the finite-dimensional optimization technique employed, and
(iii) convergence to a solution of the continuous-time problem as one takes a finer
mesh in the discretization scheme.

Runge–Kutta schemes are widely used as a general discretization tool for optimal
control problems involving ordinary differential equations (ODEs) [3, 6, 8, 13, 19, 20,
21, 44]. In fact, many well-known approximation methods such as Euler, midpoint, and
trapezoidal rules are special cases of the general Runge–Kutta schemes. Alternative
approximation schemes existing in the optimal control literature involve fitting cubic
splines to state and control variables between each of two consecutive discretization
points [43] and fitting constants, linear functions, or splines to control variables alone

∗Received by the editors July 30, 2009; accepted for publication (in revised form) July 1, 2010;
published electronically August 31, 2010. This work was supported by TÜBİTAK (Turkish Scientific
and Technical Research Council, grant 0138-11/04), during the author’s visit to the Institute of
Applied Mathematics, Middle East Technical University, Turkey, between 1 August and 15 October
2007.
†School of Mathematics and Statistics, University of South Australia, Mawson Lakes, S.A. 5095
Australia (yalcin.kaya@unisa.edu.au).

http://www.siam.org/journals/sinum/48-4/76666.html

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
However, with these alternative approximation methods, it is often not possible to achieve high orders of accuracy in the solution, without experiencing an explosion in the number of variables, unless the control is of a special type, such as bang-bang control, as in [25, 26, 33]. We choose general (implicit) Runge–Kutta schemes not only because they allow high orders of accuracy with a moderate number of variables but also because they are capable of dealing with the so-called stiff ODEs and/or preserving the structure of ODEs [21]. This addresses issue (i).

Convergence of the solution of a discretized optimal control problem to a solution of the continuous-time problem has been studied extensively in the literature (see [9, 10, 11, 14, 15, 19, 20, 22, 29, 34, 35, 36, 37, 41] and the references therein). Hager [20] presents a convergence result for general Runge–Kutta discretization of unconstrained optimal control problems. We will use that result, which addresses issue (iii).

Inexact restoration (IR) methods were introduced by Martínez and coworkers in [4, 30, 31] for solving finite-dimensional constrained optimization problems. Each iteration of an IR method consists of two phases. In the feasibility phase, feasibility of the current iterate is improved, and in the optimality phase, the cost (or some merit function) is reduced in a region that approximates the constraint set. IR methods offer freedom to choose different algorithms in both phases so that the problem structure can be exploited and especially large problems can be solved using appropriate subalgorithms. An automatic procedure in IR methods ensures that the degree of infeasibility decreases as the process advances and that feasibility is ultimately achieved. An IR method making special use of the structure of bilevel optimization problems was studied in [2], and a recent application of an IR method where the structure of the specific problem is exploited has appeared in [17]. Another IR method was used as a generalization of the spectral projected gradient method in [18]. As a recent algorithmic development, a novel general scheme incorporating a simple line search for IR methods has appeared in [16].

Birgin and Martínez [4] carried out a local convergence analysis of a general IR method for finite-dimensional optimization problems subject to equality constraints and bounds on the variables. They also provided extensive numerical comparisons with a well-known general-purpose optimization software. They demonstrated that the IR method they use is, overall, considerably more robust.

When discretized and written down as a mathematical programming problem, an unconstrained optimal control problem is transformed into a constrained optimization problem, where the constraints are given by the dynamical equations of the control system. In this paper we implement the IR method due to Birgin and Martínez [4] for the Runge–Kutta discretization of the optimal control problem and so address issue (ii).

Kaya and Martínez have applied in [24], and in the expanded version [23] of [24], the local convergence theory of IR given in [4] to Euler discretization of optimal control problems. They showed that a key sufficient condition for local convergence is satisfied readily, under the basic assumption that the initial conditions of the current and “more feasible” trajectories are the same. They also observed in numerical experiments that the mesh independence principle [1, 12, 42] was obeyed by the IR method, i.e., the behavior of the method remains unchanged as one takes a finer discretization mesh.

In this paper, we continue the line of study started in [24] by considering implicit Runge–Kutta schemes, which are much more general than Euler discretization and have the specific advantages pointed in addressing issue (i) above. In Lemma 2 we prove an analogue of the result given in Lemma 1 of [24], which states that when feasibility is improved, another key IR condition for convergence is automatically
satisfied. The result in Lemma 2 is stronger than (or different from) that in [24], not only because we deal with the more general Runge–Kutta schemes, but also we do not assume that the initial conditions for the current and more feasible trajectories are the same. This relaxation provides more flexibility for inexactness, namely, inexactness of the initial condition is allowed. We provide the main result in Theorem 2 that the IR method is locally convergent for the discretized optimal control problem under the new general scheme. Convergence to the continuous-time solution with finer mesh is facilitated by a result of Hager in [20]. The work of Dontchev and Hager [11] was utilized for obtaining the analogous result in [24].

It has been observed that Runge–Kutta discretization of an optimal control problem induces optimality conditions constituting a symplectic partitioned Runge–Kutta (SPRK) scheme written in terms of the Hamiltonian function of the optimal control problem (see [6, 20, 21]). The SPRK schemes are known to have desirable properties as integrators [21] and therefore are much preferred to use. We show in Proposition 1 that the optimality phase of IR is associated with a Hamiltonian function and that the optimality conditions for this phase constitute an SPRK scheme. This is a new result compared to the earlier study in [24].

Another novelty of this paper over [24] is in its implementation part. While feasibility and optimality phase computations in [24] were carried out exactly for simplicity, we carry out computations in this paper with truly inexact phases. As a result we convey observations under rather largely inexact feasibility and optimality phases. We illustrate the advantage of using both small and large tolerances of inexactness clearly. Furthermore we do our computations using a high-order scheme, which requires a more comprehensive coding of the theory and algorithm. Finally we even illustrate numerically the convergence result given in [20] using two high-order Runge–Kutta schemes for completeness of the numerics.

The paper is organized as follows. In section 2, the problem is stated, and the main assumptions are listed. In section 3, we describe the general Runge–Kutta discretization of the optimal control problem and its Lagrangian formulation, and we cite a convergence theorem from [20]. In section 4, application of the local IR method to the discretized optimal control problem is described, further assumptions are stated, and the main convergence results are provided in Theorem 2 and Corollary 2. In section 5, we describe a numerical implementation of the method by means of an algorithm, and we illustrate and discuss the method through a test problem. We provide comprehensive comparisons with Newton’s method, and we carry out computations under several levels of inexactness of the feasibility and optimality phases separately.

2. Optimal control problem. We consider the optimal control problem

\[
(P) \begin{cases} 
\text{minimize} & \Phi(x(t_f)) + \int_{t_0}^{t_f} f_0(x(t), u(t)) \, dt \\
\text{subject to} & \dot{x}(t) = f(x(t), u(t)), \; x(t_0) = x^0,
\end{cases}
\]

where the state variable \( x(t) \in \mathbb{R}^n, \dot{x} = dx/dt, \) the control variable \( u(t) \in \mathbb{R}^m, \) time \( t \in [t_0, t_f] \) with fixed \( t_0 \) and \( t_f, \) and the functions \( \Phi : \mathbb{R}^n \to \mathbb{R}, f_0 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \) and \( f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n. \) The initial state is specified as \( x^0. \)

The Hamiltonian function associated with problem (P) is defined by

\[
H(x, u, \lambda) = f_0(x, u) + \lambda^T f(x, u),
\]

where \( \lambda(t) \in \mathbb{R}^n \) is the costate variable.
Before listing a set of assumptions so as to cite a result from [20] on convergence of a local solution of the discretized problem to that of the continuous problem, let, as in
[20], \( L^\alpha(t_0, t_f; \mathbb{R}^n) \) denote the Lebesgue space of measurable functions \( x : [t_0, t_f] \rightarrow \mathbb{R}^n \) with \( \|x(\cdot)\|^\alpha \) integrable, equipped with the norm
\[
\|x\|_{L^\alpha} = \left[ \int_{t_0}^{t_f} \|x(t)\|^\alpha \, dt \right]^{1/\alpha},
\]
where \( \| \cdot \| \) is the Euclidean norm. The case \( \alpha = \infty \) corresponds to the space of essentially bounded, measurable functions equipped with the essential supremum norm. By \( W^{m,\alpha}(t_0, t_f; \mathbb{R}^n) \) we denote the Sobolev space consisting of functions \( x : [t_0, t_f] \rightarrow \mathbb{R}^n \) whose \( j \)th derivative lies in \( L^\alpha \) for all \( 0 \leq j \leq m \) with the norm
\[
\|x\|_{W^{m,\alpha}} = \sum_{j=0}^{m} \left\| \frac{d^j x}{dt^j} \right\|_{L^\alpha}.
\]
We pose the following assumptions for which the setting is given in some detail by Hager [20]. A similar setting is given for the Euler discretization of optimal control problems by Dontchev and Hager [11] and Kaya and Martínez [24].

Fix integer \( \kappa \geq 2 \).

**A1** Problem (P) has a local solution \((x^*, u^*)\) which lies in \( W^{\kappa,\infty} \times W^{\kappa-1,\infty} \).

**A2** In a neighborhood of \((x^*, u^*)\), the first \( \kappa \) derivatives of \( \Phi \), \( f_0 \), and \( f \) in their arguments are Lipschitz continuous.

**A3** There exists a nontrivial costate \( \lambda^* \in W^{\kappa,\infty} \) associated with problem (P) for which the following first-order necessary conditions of optimality (maximum principle) are satisfied at \((x^*, u^*, \lambda^*)\):

\[
\begin{align*}
\dot{x} &= f(x, u), \quad x(t_0) = x^0, \\
-\dot{\lambda}^T &= \frac{\partial H}{\partial x}(x, u, \lambda) = \frac{\partial f_0}{\partial x}(x, u) + \lambda^T \frac{\partial f}{\partial x}(x, u), \\
\lambda(t_f) &= \frac{\partial \Phi}{\partial x}(x(t_f)), \\
0 &= \frac{\partial H}{\partial u}(x, u, \lambda) = \frac{\partial f_0}{\partial u}(x, u) + \lambda^T \frac{\partial f}{\partial u}(x, u).
\end{align*}
\]

We say that \((x^*, u^*, \lambda^*)\) is a critical triplet of problem (P) if it satisfies (2)–(4).

We assume the Legendre condition \( \frac{\partial^2 H}{\partial u^2}(x^*, u^*, \lambda^*) > 0 \) holds so that we can solve the control \( u \) from (4) as \( u = u(x, \lambda) \). Then the differential algebraic equations (2)–(4) with the given end conditions constitute a two-point boundary value problem.

One more assumption, the so-called coercivity, is posed below, as in [20]. Let \( w^* = (x^*, u^*, \lambda^*) \), and define
\[
A^* = \frac{\partial f}{\partial x}(x^*, u^*), \quad B^* = \frac{\partial f}{\partial u}(x^*, u^*), \\
Q^* = \frac{\partial^2 H}{\partial x^2}(w^*), \quad M^* = \frac{\partial^2 H}{\partial x \partial u}(w^*), \quad R^* = \frac{\partial^2 H}{\partial u^2}(w^*).
\]

Let \( B(x, u) = \frac{1}{2} \int_{t_0}^{t_f} \left[ x^T(t) Q^* x(t) + u^T(t) R^* u(t) + 2 x^T(t) M^* u(t) \right] \, dt \).
There exists a constant $\alpha > 0$ such that
\[ B(x, u) \geq \alpha \|u\|_{L^2}^2 \quad \text{for all } (x, u) \in M, \]
where
\[ M = \{(x, u) : x \in W^{1,2}, u \in L^2, \; \dot{x} = A^*x + B^*u = 0, \; x(0) = 0 \}. \]

3. Runge–Kutta discretization of the optimal control problem. We subdivide the time horizon $[t_0, t_f]$ into $N$ pieces with the subdivision points $t_k$, $k = 0, 1, \ldots, N$, such that
\[ t_0 < t_1 < t_2 < \cdots < t_N = t_f. \]
Define the partition
\[ \pi := \{t_0, t_1, \ldots, t_N\}. \]
Without loss of generality, we take the partition points equidistant, namely, that $h := t_{k+1} - t_k$, $k = 0, 1, 2, \ldots, N - 1$. Now we consider the following one-step finite-difference approximation, the so-called Runge–Kutta scheme, of the system dynamics in problem $(P)$:
\[
\begin{align*}
    x_{k+1} &= x_k + h \sum_{i=1}^{s} b_i f(x_{ki}, u_{ki}), \\
    x_{ki} &= x_k + h \sum_{j=1}^{s} a_{ij} f(x_{kj}, u_{kj}), \quad i = 1, \ldots, s,
\end{align*}
\]
for $k = 0, 1, \ldots, N - 1$, where $b_i$ and $a_{ij}$ are constants and $s$ is referred to as the number of stages of the Runge–Kutta scheme. Here $x_k$ and $u_k$ are approximations of $x(t_k)$ and $u(t_k)$, respectively, but $x_0 = x^0$, which is specified. Further note that $x_{ki}$ and $u_{ki}$, referred to as stage vectors, are approximations of $x(t_{ki})$ and $u(t_{ki})$, respectively, where $t_{ki} := t_k + c_i h$ with constants $c_i$, $i = 1, \ldots, s$. Note that the usual choice in practice is that $c_i \in [0, 1]$, although $c_i$ can be any real constants.

We approximate the integral
\[ \int_{t_0}^{t_f} f_0(x(t), u(t)) \, dt \]
by
\[ h \sum_{k=0}^{N-1} \sum_{i=1}^{s} b_i f_0(x_{ki}, u_{ki}), \]
where $x_{ki}$, $i = 1, \ldots, s$, satisfy (6).

Now the Runge–Kutta discretization of problem $(P)$ can be stated as
\[ (DP1) \]
\[
\begin{align*}
    \text{minimize} & \quad \Phi(x_N) + h \sum_{k=0}^{N-1} \sum_{i=1}^{s} b_i f_0(x_{ki}, u_{ki}) \\
    \text{subject to} & \quad 0 = x_k - x_{k+1} + h \sum_{i=1}^{s} b_i f(x_{ki}, u_{ki}), \\
    & \quad 0 = x_k + h \sum_{j=1}^{s} a_{ij} f(x_{kj}, u_{kj}) - x_{ki}, \\
    & \quad 0 = x_0 - x_0, \quad i = 1, \ldots, s
\end{align*}
\]
for $k = 0, 1, 2, \ldots, N - 1$. Let
\[ F_{ki} := f(x_{ki}, u_{ki}), \]
and so write
\[
\begin{align*}
    f_0(x_{ki}, u_{ki}) &= f_0 \left( x_k + h \sum_{j=1}^{s} a_{ij} F_{kj}, u_{ki} \right), \\
    f(x_{ki}, u_{ki}) &= f \left( x_k + h \sum_{j=1}^{s} a_{ij} F_{kj}, u_{ki} \right),
\end{align*}
\]
where appropriate. Bonnans and Laurent-Varin [6] rewrite (6) in problem (DP1) in some equivalent form giving rise to problem (DP2) stated below. Problem (DP2) makes the derivation of the first-order necessary conditions of optimality in Hager [20] relatively simpler.

\[
\begin{array}{l}
\begin{aligned}
\text{(DP2)} & \quad \begin{cases}
    \text{minimize} & \Phi(x_N) + h \sum_{k=0}^{N-1} \sum_{i=1}^{s} b_i f_0 \left( x_k + h \sum_{j=1}^{s} a_{ij} F_{kj}, u_{ki} \right) \\
    \text{subject to} & 0 = x_k - x_{k+1} + h \sum_{i=1}^{s} b_i F_{ki}, \\
                     & 0 = f \left( x_k + h \sum_{j=1}^{s} a_{ij} F_{kj}, u_{ki} \right) - F_{ki}, \\
                     & 0 = x^0 - x_0, \quad i = 1, \ldots, s.
\end{cases}
\end{aligned}
\end{array}
\]

We denote the vector of discretized state variables by
\[ x_\pi := (x_0^T, x_1^T, x_2^T, \ldots, x_N^T)^T \in \mathbb{R}^{n(N+1)} \]
and call $x_\pi$ the (state) trajectory of the discretized system. Similarly, let
\[ u_\pi := (u_0^T, u_1^T, u_2^T, \ldots, u_{N-1}^T)^T \in \mathbb{R}^{mN}. \]

We define another partition for the state intermediate states and controls, $\sigma$, which is finer than $\pi$ for $s > 1$, by
\[ \sigma := \{ \{t_{0i}\}_{i=1}^{s}, \ldots, \{t_{(N-1)i}\}_{i=1}^{s} \}. \]

With this partition, we denote
\[ x_\sigma := (x_{01}^T, x_{02}^T, \ldots, x_{(N-1)s}^T)^T \in \mathbb{R}^{nNs}, \quad u_\sigma := (u_{01}^T, u_{02}^T, \ldots, u_{(N-1)s})^T \in \mathbb{R}^{mNs}, \]
representing the state intermediate state and control values of the whole Runge–Kutta discretization. Similarly,
\[ F_\sigma := (F_{01}^T, F_{02}^T, \ldots, F_{(N-1)s})^T \in \mathbb{R}^{nNs}. \]
The Lagrangian for problem (DP2) is given by

\[
L(x_i, F_\sigma, u_\sigma, \lambda_\pi, \rho_\sigma) = \Phi(x_N) + h \sum_{k=0}^{N-1} \sum_{i=1}^{s} b_i f_0 \left( x_k + h \sum_{j=1}^{s} a_{ij} F_{kj}, u_{ki} \right) \\
+ \lambda_0^T (x_0 - x_0) + \sum_{k=0}^{N-1} \lambda_{k+1}^T \left( x_k - x_{k+1} + h \sum_{i=1}^{s} b_i F_{ki} \right) \\
+ \sum_{k=0}^{N-1} \sum_{i=1}^{s} \rho_{ki}^T \left[ f \left( x_k + h \sum_{j=1}^{s} a_{ij} F_{kj}, u_{ki} \right) - F_{ki} \right],
\]

(8)

where

\[
\lambda_\pi := (\lambda_0^T, \lambda_1^T, \ldots, \lambda_N^T)^T \in \mathbb{R}^{N+1} \quad \text{and} \quad \rho_\sigma := (\rho_{01}^T, \rho_{02}^T, \ldots, \rho_{(N-1)s}^T)^T \in \mathbb{R}^{Ns}
\]

are the Lagrange multiplier vectors.

First-order necessary conditions for problem (DP2) (or equivalently for problem DP1) can be obtained through

\[
\nabla L(x_i, F_\sigma, u_\sigma, \lambda_\pi, \rho_\sigma) = 0.
\]

One gets, through differentiation of \(L\) with respect to \(x_k\) (see (9)–(10)), \(F_{ki}\) (see (11)), and \(u_{ki}\) (see (12)):

(9) \[ \frac{\partial \Phi}{\partial x}(x_N) - \lambda_N^T = 0 \]

(10) \[ h \sum_{i=1}^{s} b_i \frac{\partial f_0}{\partial x}(x_{ki}, u_{ki}) + \lambda_{k+1}^T - \lambda_k^T + \sum_{i=1}^{s} \rho_{ki}^T \frac{\partial f}{\partial x}(x_{ki}, u_{ki}) = 0, \]

(11) \[ h^2 \sum_{j=1}^{s} b_j a_{ij} \frac{\partial f_0}{\partial x}(x_{kj}, u_{kj}) + h b_i \lambda_{k+1}^T \\
+ h \sum_{j=1}^{s} a_{ij} \rho_{kj}^T \frac{\partial f}{\partial x}(x_{kj}, u_{kj}) - \rho_{ki}^T = 0, \]

(12) \[ h b_i \frac{\partial f_0}{\partial u}(x_{ki}, u_{ki}) + \rho_{ki}^T \frac{\partial f}{\partial u}(x_{ki}, u_{ki}) = 0 \]

for \(k = 0, 1, \ldots, N - 1, i = 1, \ldots, s\).

Let

\[ \lambda_{ki} := \frac{1}{h b_i} \rho_{ki}. \]

Note that for this definition (or transformation) to be valid, one needs only \(b_i \neq 0\) for every \(i = 1, \ldots, s\). However, for convergence to the continuous solution as stated in Theorem 1, the following assumption is posed:

(A5) \[ b_i > 0 \text{ for every } i = 1, \ldots, s. \]

By eliminating \(\rho_{ki}\) from (10)–(11) and by using the definition of the Hamiltonian in (1), the following first-order necessary optimality conditions are obtained. It should be observed that if one assumes that the constraint \(x_0 = x^0\) is satisfied, then the
multiplier \( \lambda_0 \) becomes redundant.

\[
\begin{align*}
(13) & \quad x_{k+1} = x_k + h \sum_{i=1}^s b_i f(x_{ki}, u_{ki}), \\
(14) & \quad x_{ki} = x_k + h \sum_{j=1}^s a_{ij} f(x_{kj}, u_{kj}), \quad x_0 = x^0, \\
(15) & \quad \lambda^T_{k+1} = \lambda^T_k - h \sum_{i=1}^s b_i H_x(x_{ki}, u_{ki}, \lambda_{ki}), \\
(16) & \quad \lambda^T_k = \lambda^T_k - h \sum_{j=1}^s \tilde{a}_{ij} H_x(x_{kj}, u_{kj}, \lambda_{kj}), \quad \lambda^T_N = \frac{\partial \Phi}{\partial x}(x_N).
\end{align*}
\]

for \( k = 0, 1, \ldots, N - 1, i = 1, \ldots, s \), where \( H_x = \partial H/\partial x, H_u = \partial H/\partial u \), and

\[
(18) \quad \tilde{a}_{ij} = b_j - \frac{b_j}{b_i} a_{ji}.
\]

Recall from [24] that for the case of Euler discretization, \( s = 1, b_1 = 1, a_{11} = 0 \), and \( \tilde{a}_{ij} = 1 \).

By (A3), \( H_{uu}(x, u, \lambda) > 0 \) so that one can locally solve (17) for \( u_{ki} \) in terms of \( x_{ki} \) and \( \lambda_{ki} \), namely, that

\[
u_{ki} = \varphi(x_{ki}, \lambda_{ki})
\]

and similarly

\[
u_k = \varphi(x_k, \lambda_k),
\]

\( \varphi \) sufficiently smooth. After substitution of \( u_{ki} = \varphi(x_{ki}, \lambda_{ki}) \), (13)–(16) constitute a so-called SPRK scheme due to the relationship (18) [6, 21].

A Runge–Kutta scheme of order \( \kappa \) is said to be of order \( \kappa \) for optimal control if it yields discretization of an optimal control problem of order \( \kappa \) (of accuracy) in the state, control, and costate variables. In order for a Runge–Kutta scheme to be of order \( \kappa \) for optimal control, certain algebraic conditions on the coefficients of the Runge–Kutta scheme have to be met. Hager derives and lists such conditions for up to order 4 in [20] and Bonnans and Laurent-Varin for up to order 5 in [6] and up to order 6 in [5]. The Runge–Kutta schemes of various orders listed in Appendix A all satisfy the conditions presented in [5].

We say that \( (x^*_n, u^*_n, \lambda^*_n) \) is a critical triplet of problem (DP1) if it satisfies (13)–(17).

**Theorem 1** (see Hager [20]). If (A1)–(A5) hold and the Runge–Kutta scheme is of order \( \kappa \) for optimal control, then for all sufficiently small \( h \) there exists a strict local minimizer \( (x^*_n, u^*_n) \) of problem (DP1) and an associated Lagrange multiplier \( \lambda^*_n \) such that

\[
(19) \quad \max_{0 \leq k \leq N} \|x^*_k - x^*(t_k)\| + \max_{0 \leq k \leq N} \|\lambda^*_k - \lambda^*(t_k)\| + \max_{0 \leq k \leq N} \|u^*_k - u^*(t_k)\| \leq c h^{\kappa},
\]

where \( u^*_k = \varphi(x^*_k, \lambda^*_k) \) is a local minimizer of the Hamiltonian (1) corresponding to \( x = x^*_k \) and \( \lambda = \lambda^*_k \) and \( c \) is a constant independent of \( h \).

**Corollary 1.** If (A1)–(A5) hold, then, as \( h \to 0 \), a local solution of problem (DP1) converges to a local solution of problem (P).
4. IR for Runge–Kutta discretization of optimal control problems. In this section, we will give a formulation of the IR method for solving Runge–Kutta discretization of optimal control problems and present convergence results.

Recall that an IR iteration consists of two phases. In the feasibility phase, given the current iterate \((x_\pi, x_\sigma, u_\sigma) \in \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nN_s} \times \mathbb{R}^{mN_s}\), one finds a “more feasible” point \((y_\pi, y_\sigma, u_\sigma) \in \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nN_s} \times \mathbb{R}^{mN_s}\). In the optimality phase, one obtains a “more optimal” point \((\hat{x}_\pi, \hat{z}_\sigma, v_\sigma) \in \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nN_s} \times \mathbb{R}^{mN_s}\) in the tangent plane passing through \((y_\pi, y_\sigma, u_\sigma)\).

Form the constraint function from (5) and (6) by

\[
\ell_k(x_k, x_{k+1}, \{x_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s) := \begin{pmatrix}
x_k - x_{k+1} + h \sum_{i=1}^s b_i f(x_{ki}, u_{ki}) \\
x_k - x_i + h \sum_{j=1}^s a_{ij} f(x_{kj}, u_{kj})
\end{pmatrix}
\]

for \(k = 0, 1, \ldots, N - 1, i = 1, \ldots, s\), and let

\[
\ell(x_\pi, x_\sigma, u_\sigma) := (\ell_0^T(x_0, \{x_0\}_{i=1}^s, \{u_0\}), \ell_1^T(x_1, \{x_1\}_{i=1}^s, \{u_1\}), \ldots, \ell_{N-1}^T(x_{N-1}, \{x_{N-1}\}_{i=1}^s, \{u_{N-1}\}))^T, \quad (x^0 - x_0)^T.
\]

Although one might consider fixing \(x_0 = x^0\) in the feasibility phase as was done earlier with the Euler discretization in [24], inclusion of the initial condition (constraint) function \((x^0 - x_0)\) in the definition of \(\ell\) above is convenient because this allows inexactness at the initial time point like the other time points.

The tangent plane in which the optimality phase is performed is formed by \((\hat{z}_\pi, \hat{z}_\sigma, v_\sigma)\) which solves

\[
\nabla \ell(y_\pi, y_\sigma, u_\sigma) (\hat{z}_\pi - y_\pi, \hat{z}_\sigma - y_\sigma, v_\sigma - u_\sigma) = 0.
\]

The tangent plane can equivalently be expressed by using the linearization of \(\ell_k(\hat{z}_k, \hat{z}_{k+1}, \{\hat{z}_{ki}\}_{i=1}^s, \{\hat{v}_{ki}\}_{i=1}^s)\) from (20) at \((y_k, y_{k+1}, \{y_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s)\):

\[
(\hat{z}_k - y_k) - (\hat{z}_{k+1} - y_{k+1}) + h \sum_{i=1}^s b_i [A_{ki} (\hat{z}_k - y_k) + B_{ki} (\hat{v}_k - u_k)] = 0
\]

with

\[
(\hat{z}_k - y_k) - (\hat{z}_k - y_k) + h \sum_{j=1}^s a_{ij} [A_{kj} (\hat{z}_k - y_k) + B_{kj} (\hat{u}_k - u_k)] = 0
\]

and

\[
z_0 - y_0 = 0,
\]

where we have used the shorthand notation

\[
A_{ki} := \frac{\partial f}{\partial x}(y_{ki}, u_{ki}) \quad \text{and} \quad B_{ki} := \frac{\partial f}{\partial u}(y_{ki}, u_{ki})
\]

for \(k = 0, 1, \ldots, N - 1, i = 1, \ldots, s\). Note that the matrices \(A_{ki}\) and \(B_{ki}\) vary with the time discretization indices \(k\) and \(i\).
For notational convenience, let
\[
\tilde{f}(z_{ki}, v_{ki}) := A_{ki} (z_{ki} - y_{ki}) + B_{ki} (v_{ki} - u_{ki}) .
\]
For brevity and simplicity in calculations, we will also use
\[
\tilde{F}_{ki} := \tilde{f}(z_{ki}, v_{ki}) .
\]
Let
\[
\tilde{F}_{\pi} := \left( \tilde{F}_{01}^{T}, \tilde{F}_{02}^{T}, \ldots, \tilde{F}_{(N-1)s}^{T} \right)^{T} \in \mathbb{R}^{nNs} .
\]
In the optimality phase of the IR method, we minimize the Lagrangian
\[
\tilde{L}(z_{\pi}, F_{\sigma}, \tilde{F}_{\sigma}, v_{\sigma}, \mu_{\pi}, \gamma_{\sigma}) := L(z_{\pi}, F_{\sigma}, v_{\sigma}, \mu_{\pi}, \gamma_{\sigma}) + (\mu_{0}^{T} - \lambda_{0}^{T}) (z_{0} - y_{0})
\]
\[
+ \sum_{k=0}^{N-1} \left( (\mu_{k+1}^{T} - \lambda_{k+1}^{T}) \left[ (z_{k} - y_{k}) - (z_{k+1} - y_{k+1}) + h \sum_{i=1}^{s} b_{i} \tilde{F}_{ki} \right] \right)
\]
\[
+ \sum_{k=0}^{N-1} \sum_{i=1}^{s} (\gamma_{ki}^{T} - \rho_{ki}^{T}) \left[ \tilde{f}(y_{ki} + z_{k} - y_{k} + h \sum_{j=1}^{s} a_{ij} \tilde{F}_{kj}, v_{ki}) - \tilde{F}_{ki} \right]
\]
\[
= \Phi(x_{N}) + h \sum_{k=0}^{N-1} \sum_{i=1}^{s} b_{i} f_{0} \left( z_{k} + h \sum_{j=1}^{s} a_{ij} \tilde{F}_{kj}, v_{ki} \right)
\]
\[
+ \lambda_{0}^{T} (x^{0} - z_{0}) + \sum_{k=0}^{N-1} \lambda_{k+1}^{T} \left( z_{k} - z_{k+1} + h \sum_{i=1}^{s} b_{i} F_{ki} \right)
\]
\[
+ \sum_{k=0}^{N-1} \sum_{i=1}^{s} \rho_{ki}^{T} \left[ f \left( z_{k} + h \sum_{j=1}^{s} a_{ij} F_{kj}, v_{ki} \right) - F_{ki} \right]
\]
\[
+ (\mu_{0}^{T} - \lambda_{0}^{T}) (z_{0} - y_{0})
\]
\[
+ \sum_{k=0}^{N-1} (\mu_{k+1}^{T} - \lambda_{k+1}^{T}) \left[ (z_{k} - y_{k}) - (z_{k+1} - y_{k+1}) + h \sum_{i=1}^{s} b_{i} \tilde{F}_{ki} \right]
\]
\[
+ \sum_{k=0}^{N-1} \sum_{i=1}^{s} (\gamma_{ki}^{T} - \rho_{ki}^{T}) \left[ \tilde{f}(y_{ki} + z_{k} - y_{k} + h \sum_{j=1}^{s} a_{ij} \tilde{F}_{kj}, v_{ki}) - \tilde{F}_{ki} \right],
\]
where \( \mu_{\pi} := (\mu_{0}, \ldots, \mu_{N}) \in \mathbb{R}^{n(N+1)} \) and \( \gamma_{\sigma} := (\gamma_{01}^{T}, \gamma_{02}^{T}, \ldots, \gamma_{(N-1)s}^{T}) \in \mathbb{R}^{nNs} \), and
(\mu_{k} - \lambda_{k}) and (\gamma_{ki} - \rho_{ki}) are the Lagrange multipliers employed for the linear constraints from (22) with (24) and (23), respectively. The first-order necessary conditions of optimality in this case are given by
\[
\nabla \tilde{L}(z_{\pi}, F_{\sigma}, \tilde{F}_{\sigma}, v_{\sigma}, \mu_{\pi}, \gamma_{\sigma}) = 0 .
\]
One gets, through differentiation of \( L \) with respect to \( z_k \) (see (26)–(27)), \( F_{ki} \), \( \tilde{F}_{ki} \) (see (28)–(29)), and \( v_{ki} \) (see (30)):

\[
\begin{align*}
\frac{\partial \Phi}{\partial x}(z_N) - \mu_T^N &= 0, \\
h \sum_{i=1}^s b_i \frac{\partial f_0}{\partial x}(z_{ki}, v_{ki}) + \sum_{i=1}^s \rho_{ki} \frac{\partial f}{\partial x}(z_{ki}, v_{ki}) + \mu_{k+1}^T - \mu_k^T \\
+ \sum_{i=1}^s (\gamma_{ki}^T - \rho_{ki}^T) \frac{\partial f}{\partial x}(z_{ki}, v_{ki}) &= 0, \\
h^2 \sum_{j=1}^s b_j a_{ji} \frac{\partial f_0}{\partial x}(z_{kj}, v_{kj}) + h b_i \lambda_{k+1}^T \\
+ h \sum_{j=1}^s a_{ji} \rho_{kj}^T \frac{\partial f}{\partial x}(z_{kj}, v_{kj}) - \rho_{ki}^T &= 0, \\
h b_i (\mu_{k+1}^T - \lambda_{k+1}^T) + h \sum_{j=1}^s a_{ji} (\gamma_{kj}^T - \rho_{kj}^T) A_{kj} - (\gamma_{ki}^T - \rho_{ki}^T) &= 0, \\
h b_i \frac{\partial f_0}{\partial u}(z_{ki}, v_{ki}) + \rho_{ki}^T \frac{\partial f}{\partial u}(z_{ki}, v_{ki}) + (\gamma_{ki}^T - \rho_{ki}^T) B_{ki} &= 0
\end{align*}
\]

for \( k = 0, 1, \ldots, N - 1, i = 1, \ldots, s \). Let

\[
\lambda_{ki} = \frac{1}{h b_i} \rho_{ki} \quad \text{and} \quad \mu_{ki} = \frac{1}{h b_i} \gamma_{ki}.
\]

By eliminating \( \rho_{ki} \) and \( \gamma_{ki} \) from (27)–(30) and by carrying out manipulations, we obtain the following first-order necessary optimality conditions. Note that if one assumes that the constraint \( z_0 = x^0 \) is satisfied, then the multiplier \( (\mu_0 - \lambda_0) \) becomes redundant.

\[
\begin{align*}
\mu_{k+1}^T &= \mu_k^T - h \sum_{i=1}^s b_i \left[ \frac{\partial f_0}{\partial x}(z_{ki}, v_{ki}) + \lambda_{k+1}^T \frac{\partial f}{\partial x}(z_{ki}, v_{ki}) + (\mu_{ki}^T - \lambda_{k+1}^T) A_{ki} \right], \\
\mu_{ki}^T &= \mu_k^T - h \sum_{j=1}^s \left( b_j - \frac{b_j}{b_i} a_{ji} \right) \left[ \frac{\partial f_0}{\partial x}(z_{kj}, v_{kj}) + \lambda_{kj}^T \frac{\partial f}{\partial x}(z_{kj}, v_{kj}) + (\mu_{kj}^T - \lambda_{kj}^T) A_{kj} \right], \\
0 &= \frac{\partial f_0}{\partial u}(z_{ki}, v_{ki}) + \lambda_{ki}^T \frac{\partial f}{\partial u}(z_{ki}, v_{ki}) + (\mu_{ki}^T - \lambda_{ki}^T) B_{ki}, \\
\mu_N^T &= \frac{\partial \Phi}{\partial x}(z_N)
\end{align*}
\]

for \( k = 0, 1, \ldots, N - 1, i = 1, \ldots, s \).

**Proposition 1.** The IR equations (22)–(24) and (31)–(34) constitute the symplectic partitioned Runge–Kutta discretization of a Hamiltonian system.
Proof. Define the optimality phase Hamiltonian function

\begin{equation}
\tilde{H}(\tilde{z}, \tilde{v}, \mu, t) := f_0(z, v) + \lambda^T(t) f(z, v) + (\mu^T - \lambda^T(t)) (A(t) \tilde{z} + B(t) \tilde{v}),
\end{equation}

where \( \tilde{z} = z - y(t), \tilde{v} = v - u(t), \ A(t) = \partial f / \partial x(y(t), u(t)), \) and \( B(t) = \partial f / \partial u(y(t), u(t)). \) Here \( y(t) \) and \( u(t) \) are specified functions of time such that their discretization with respect to the partitions \( \pi \) and \( \sigma \) gives the discrete values \( y_k = y(t_k), \ y_{ki} = y(t_k + c_i h), \ u_k = u(t_k), \) and \( u_{ki} = u(t_k + c_i h). \) The first-order necessary conditions of optimality (31)–(34), along with the linearized constraint equations (22)–(24), can now be written in terms of the Hamiltonian in (35) as follows:

\begin{align}
\tilde{z}_{k+1} &= \tilde{z}_k + h \sum_{i=1}^{s} b_i \tilde{H}_\mu(\tilde{z}_{ki}, \tilde{v}_{ki}, \mu_{ki}, t_{ki}), \\
\tilde{z}_k &= \tilde{z}_k + h \sum_{j=1}^{s} a_{ij} \tilde{H}_\mu(\tilde{z}_{kj}, \tilde{v}_{kj}, \mu_{kj}, t_{kj}), \quad \tilde{z}_0 = 0, \\
\mu_{k+1}^T &= \mu_k^T - h \sum_{i=1}^{s} b_i \tilde{H}_z(\tilde{z}_{ki}, \tilde{v}_{ki}, \mu_{ki}, t_{ki}), \\
\mu_{ki}^T &= \mu_{ki}^T - h \sum_{j=1}^{s} a_{ij} \tilde{H}_z(\tilde{z}_{kj}, \tilde{v}_{kj}, \mu_{kj}, t_{kj}), \quad \mu_N^T = \frac{\partial \Phi}{\partial x}(z_N), \\
0 &= \tilde{H}_v(\tilde{z}_{ki}, \tilde{v}_{ki}, \mu_{ki}, t_{ki})
\end{align}

for \( k = 0, 1, \ldots, N - 1, \) and \( t_{ki} = t_k + c_i h. \) In the above equations, \( \tilde{z}_k = z_k - y_k, \ \tilde{v}_k = v_k - u_k, \ \tilde{H}_\mu = \partial \tilde{H} / \partial \mu, \ \tilde{H}_z = \partial \tilde{H} / \partial \tilde{z}, \) and \( \tilde{H}_v = \partial \tilde{H} / \partial \tilde{v}. \) We have also used the fact that \( \tilde{H}_z = H_z. \)

In what follows we define certain vectors used in the sufficient conditions for local convergence of the IR method presented in [4]. Norms of these vectors represent some measure of optimality in the optimality phase. Let

\[
G_k(\{x_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s, \lambda_k, \lambda_{k+1}, \{\lambda_{ki}\}_{i=1}^s) :=
\begin{pmatrix}
\lambda_k^T - h \sum_{i=1}^s b_i H_x(x_{ki}, u_{ki}, \lambda_{ki}) - \lambda_{k+1}^T \\
\lambda_k^T - h \sum_{j=1}^s a_{ij} H_x(x_{kj}, u_{kj}, \lambda_{kj}) - \lambda_{ki}^T (i = 1, \ldots, s) \\
\lambda_N^T - \frac{\partial \Phi}{\partial x}(x_N) \\
H_u(x_{ki}, u_{ki}, \lambda_{ki}) (i = 1, \ldots, s)
\end{pmatrix}
\]

for \( k = 0, 1, \ldots, N - 1. \) Next we define (omitting the arguments of \( G_k \) for clarity in appearance) the column vector function

\[
G(x_{\sigma}, u_{\sigma}, \lambda_{\pi}, \lambda_{\sigma}) := (G_0, G_1, \ldots, G_{N-1})^T.
\]
Also let
\[
\bar{G}_k(\{z_{ki}\}_{i=1}^{s}, \{v_{ki}\}_{i=1}^{s}, \mu_k, \mu_{k+1}, \{\mu_{ki}\}_{i=1}^{s}) := \left(\mu_k^T - h \sum_{i=1}^{s} b_i \mathbf{H}(\bar{z}_{ki}, \bar{v}_{ki}, \mu_k) - \mu_{k+1}^T, \mu_k^T - h \sum_{j=1}^{s} \bar{a}_{ij} \mathbf{H}(\bar{z}_{kj}, \bar{v}_{kj}, \mu_k) - \mu_{k+1}^T (i = 1, \ldots, s), \mathbf{H}_u(\bar{z}_{ki}, \bar{v}_{ki}, \mu_k) (i = 1, \ldots, s)\right)
\]

with \(\bar{z}_k = z_k - y_k\) and \(\bar{v}_k = v_k - u_k\) for \(k = 0, 1, \ldots, N - 1\). Next we define the column vector function
\[
\bar{G}(\bar{z}, \bar{v}, \mu, \lambda) := (\bar{G}_0, \bar{G}_1, \ldots, \bar{G}_{N-1})^T.
\]

An IR iteration starting from \((x_\pi, x_\sigma, u_\pi, \lambda_\pi, \lambda_\sigma) \in \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nNs} \times \mathbb{R}^{mNs} \times \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nNs}\) can be completed (or is well defined) if one can compute \((\bar{y}_\pi, \bar{y}_\sigma) \in \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nNs}, (\bar{z}_\pi, \bar{z}_\sigma, \bar{v}_\pi, \bar{v}_\lambda, \bar{\lambda}_\pi, \bar{\lambda}_\sigma) \in \mathbb{R}^{n(N+1)} \times \mathbb{R}^{nNs} \times \mathbb{R}^{mNs} \times \mathbb{R}^{mN} \times \mathbb{R}^{mNs}\) such that it satisfies the following conditions (see Birgin and Martínez [4, Conditions (6)–(10)], where these conditions are given for a standard finite-dimensional optimization problem):

\[
\|\ell(y_\pi, y_\sigma)\| \leq \theta \|\ell(x_\pi, x_\sigma, u_\sigma)\|, \tag{41}
\]

\[
\|y_\pi - x_\pi\| + \|y_\sigma - x_\sigma\| \leq K_1 \|\ell(x_\pi, x_\sigma, u_\sigma)\|, \tag{42}
\]

\[
\|\nabla\ell(y_\pi, y_\sigma, u_\sigma)(z_\pi - y_\pi, z_\sigma - y_\sigma, v_\pi - u_\sigma)\| \leq K_2 \|G(x_\pi, x_\sigma, u_\sigma, \lambda_\pi, \lambda_\sigma)\|^2, \tag{43}
\]

\[
\|\bar{G}(\bar{z}, \bar{v}, \mu, \lambda)\| \leq \eta \|G(x_\pi, u_\sigma, \lambda_\pi, \lambda_\sigma)\|, \tag{44}
\]

\[
\|z_\pi - y_\pi\| + \|z_\sigma - y_\sigma\| + \|v_\pi - u_\pi\| + \|v_\sigma - u_\sigma\| + \|\mu_\pi - \lambda_\pi\| + \|\mu_\sigma - \lambda_\sigma\|
\leq K_3 \|G(x_\pi, u_\sigma, \lambda_\pi, \lambda_\sigma)\|, \tag{45}
\]

where \(\theta, \eta \in [0, 1), K_1, K_3 > 0, K_2 \geq 0, \) and \(\| \cdot \|\) is any norm in the relevant finite-dimensional space.

**Lemma 1.** Suppose (A2) holds. Then, for fixed \(t_k\) and small enough \(h\), in a neighborhood of the continuous-time solution \((x^*(t_k), u^*(t_k))\) there exists \(K > 0\) such that

\[
\|f(y_{ki}, u_{ki}) - f(x_{ki}, u_{ki})\| \leq K\|y_k - x_k\|. \tag{46}
\]

**Proof.** Under (A2), there exists \(\overline{K} > 0\) such that for all \(x, y, u,\) and \(t,\)

\[
\|f(y(t), u(t)) - f(x(t), u(t))\| \leq \overline{K}\|y(t) - x(t)\|. \tag{46}
\]
By the state uniqueness property given by Hager in [20], for \((x_k, u_{ki}), i = 1, \ldots, s\), which belongs to a neighborhood of \((x^*(t_k), u^*(t_k))\), and for small enough \(h\), \(x_{ki}\) in (6) can be determined uniquely in \(x_k\) and \(\{u_{kj}\}_{j=1}^s\), i.e.,

\[
x_{ki} = \Psi(x_k, \{u_{ki}\}_{j=1}^s).
\]

This fact is furnished by the smoothness assumption stated in (A2) and the implicit function theorem. Note furthermore that \(\Psi\) is \(\kappa\)-times continuously differentiable in \(x_k\) and \(u_{kj}\), \(j = 1, \ldots, s\). Therefore, there exists \(\bar{K} > 0\) such that, for all \(x, y, \{u_{kj}\}_{j=1}^s\),

\[
\|\Psi(y, \{u_{kj}\}_{j=1}^s) - \Psi(x, \{u_{kj}\}_{j=1}^s)\| \leq \bar{K}\|y - x\|.
\]

Consequently, by (46)–(48), for \((x_{ki}, u_{ki})\) and \((y_{ki}, u_{ki})\) sufficiently close to \((x^*(t_k), u^*(t_k))\), one has

\[
\|f(y_{ki}, u_{ki}) - f(x_{ki}, u_{ki})\| \leq \bar{K}\|y_{ki} - x_{ki}\|
\]

\[
= \bar{K}\|\Psi(y_k, \{u_{kj}\}_{j=1}^s) - \Psi(x_k, \{u_{kj}\}_{j=1}^s)\|
\]

\[
\leq \bar{K}\|y_k - x_k\|
\]

where \(K = \bar{K}\bar{K}\).

**Lemma 2.** Suppose that (A2) and (A5) hold. If (41) is satisfied, so is (42), for fixed \(t_k\) and small enough \(h\), in a neighborhood of the continuous-time solution \((x^*(t_k), u^*(t_k))\).

**Proof.** Suppose (41) is satisfied. Then we have

\[
\|\ell(y, y_\sigma, u_\sigma)\| \leq \theta\|\ell(x, x_\sigma, u_\sigma)\| \leq \|\ell(x, x_\sigma, u_\sigma)\|.
\]

Denote by \(\ell_1^k\) and \(\ell_2^k\) the first and second components of \(\ell_k\) in (20), respectively. Then we can rewrite the first equation in (20) as

\[
x_{k+1} = x_k + h\sum_{i=1}^s b_i f(x_{ki}, u_{ki}) - \ell_1^k(x_k, x_{k+1}, \{x_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s)
\]

and, similarly

\[
y_{k+1} = y_k + h\sum_{i=1}^s b_i f(y_{ki}, u_{ki}) - \ell_1^k(y_k, y_{k+1}, \{y_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s)
\]

for \(k = 0, 1, \ldots, N - 1\). Now

\[
y_{k+1} - x_{k+1} = y_k - x_k + h\sum_{i=1}^s b_i (f(y_{ki}, u_{ki}) - f(x_{ki}, u_{ki}))
\]

\[
- \ell_1^k(y_k, y_{k+1}, \{y_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s) + \ell_1^k(x_k, x_{k+1}, \{x_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s).
\]

For notational convenience, define

\[
r_k := \ell_1^k(y_k, y_{k+1}, \{y_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s) - \ell_1^k(x_k, x_{k+1}, \{x_{ki}\}_{i=1}^s, \{u_{ki}\}_{i=1}^s)
\]

for \(k = 0, 1, \ldots, N - 1\),

\[
r_N := (x^0 - y_0) - (x^0 - x_0) = -(y_0 - x_0),
\]
and $r_\pi := (r_0, \ldots, r_N)$. Without loss of generality, in what follows we use the sup-norm, for example, $\|r_\pi\|_\infty = \sup_i \|r_k\|$, where $\|\cdot\|$ denotes the 1-norm in $\mathbb{R}^n$, or in another finite-dimensional space, appropriately. Note that $\|r_\pi\|_\infty = \|\ell(y_\pi, y_\sigma, u_\sigma) - \ell(x_\pi, x_\sigma, u_\sigma)\|_\infty$. Now, for $k = 0, 1, \ldots, N - 1$,

$$\|y_{k+1} - x_{k+1}\| \leq \|y_k - x_k\| + h \sum_{i=1}^{s} b_i \|f(y_{ki}, u_{ki}) - f(x_{ki}, u_{ki})\| + \|r_k\|,$$

where we also used the assumption that $b_i > 0$ for every $i$. Thus, by Lemma 1 and the order condition $\sum_{i=1}^{s} b_i = 1$ (for any $s \geq 1$),

$$\|y_{k+1} - x_{k+1}\| \leq (1 + h s K) \|y_k - x_k\| + \|r_k\|. \tag{50}$$

Inequality (50) implies that

$$\|y_1 - x_1\| \leq (1 + h s K) \|r_N\| + \|r_0\|. \tag{51}$$

Furthermore, the inequalities (50), for $k = 1$, and (51) yield

$$\|y_2 - x_2\| \leq (1 + h s K)^2 \|r_N\| + (1 + h s K) \|r_0\| + \|r_1\|. \tag{52}$$

Proceeding inductively, we prove that, for all $k = 0, 1, \ldots, N - 1$,

$$\|y_k - x_k\| \leq (1 + h s K)^k \|r_N\| + (1 + h s K)^k - 1 \|r_0\| + \cdots + (1 + h s K) \|r_{k-1}\| + \|r_k\|$$

$$\leq (1 + h s K)^N \|r_0\| + \cdots + \|r_N\|$$

$$\leq (1 + h s K)^N \|r_\pi\|_\infty.$$ 

Similarly, we can rewrite the second equation in (20) as

$$x_{ki} = x_k + h \sum_{j=1}^{s} a_{ij} f(x_{kj}, u_{kj}) - \xi_k^2(x_k, x_{k+1}, \{x_{kj}\}_{j=1}^{s}, \{u_{kj}\}_{j=1}^{s})$$

and

$$y_{ki} = y_k + h \sum_{j=1}^{s} a_{ij} f(y_{kj}, u_{kj}) - \xi_k^2(y_k, y_{k+1}, \{y_{kj}\}_{j=1}^{s}, \{u_{kj}\}_{j=1}^{s})$$

for $k = 0, 1, \ldots, N - 1$, $i = 1, \ldots, s$. Define

$$r_{ki} := \ell_k^2(y_k, y_{k+1}, \{y_{ki}\}_{i=1}^{s}, \{u_{ki}\}_{i=1}^{s}) - \ell_k^2(x_k, x_{k+1}, \{x_{ki}\}_{i=1}^{s}, \{u_{ki}\}_{i=1}^{s})$$

for $k = 0, 1, \ldots, N - 1$ and $r_\sigma := (r_0, \ldots, r_{(N-1)s})$. Let

$$\overline{\pi} = \max_{i,j} a_{ij}.$$ 

Now

$$\|y_{ki} - x_{ki}\| \leq \|y_k - x_k\| + h \sum_{j=1}^{s} a_{ij} \|f(y_{kj}, u_{kj}) - f(x_{kj}, u_{kj})\| + \|r_{ki}\|$$

$$\leq \|y_k - x_k\| + h s K \overline{\pi} \|y_k - x_k\| + \|r_{ki}\|$$

$$\leq (N + 1 + \overline{\pi}) (1 + h s K)^N \|r_\pi\|_\infty + \|r_\sigma\|_\infty, \tag{54}$$
where in the second inequality, Lemma 1 and (53), and in the last, (52) and \( \| r_{ki} \| \leq \| r_{\sigma} \|_{\infty} \) have been used. Now (52) and (54) can be combined to give

\[
\| y_k - x_k \| + \| y_{ki} - x_{ki} \| \leq 2 (N + 1 + \overline{\pi}) (1 + h s K)^N \| r_{\sigma} \|_{\infty} + \| r_{\sigma} \|_{\infty}.
\]

This is valid for any \( k \) and \( i \), so, by (49),

\[
\| y_{\pi} - x_{\pi} \|_{\infty} + \| y_{\sigma} - x_{\sigma} \|_{\infty} = \sup_{0 \leq j \leq N-1} \| y_j - x_j \| + \sup_{0 \leq m \leq N-1} \| y_{jm} - x_{jm} \|
\leq 2 (N + 1 + \overline{\pi}) (1 + h s K)^N \| r_{\pi} \|_{\infty} + \| r_{\sigma} \|_{\infty}
\leq 2 (N + 1 + \overline{\pi}) (1 + h s K)^N (\| r_{\pi} \|_{\infty} + \| r_{\sigma} \|_{\infty})
\leq 2 (N + 1 + \overline{\pi}) (1 + h s K)^N
\cdot (\| \ell(y_{\pi}, y_{\sigma}, u_{\sigma}) + \ell(x_{\pi}, x_{\sigma}, u_{\sigma}) \|_{\infty})
\leq 4 (N + 1 + \overline{\pi}) (1 + h s K)^N \| \ell(x_{\pi}, x_{\sigma}, u_{\sigma}) \|_{\infty} \tag{55}
\]

Remark 1. (55) implies that \( K_1 \) depends on the Lipschitz constant \( K \), the number of subdivisions \( N \), the number of Runge–Kutta stages \( s \), and the maximum of \( |a_{ij}|, \overline{\pi} \), namely, that we have \( K_1 \geq 4(N + 1 + \overline{\pi}) (1 + h s K)^N \). We observe that \( (1 + h s K)^N = (1 + (t_f - t_0) K/N)^N \) is increasing in \( N \) and that, as \( h \to 0 \), \( (1 + h s K)^N \to e^{(t_f - t_0) s K} \). So (42) mentioned in the conclusion of Lemma 2 is satisfied with \( K_1 \geq 4(N + 1 + \overline{\pi}) e^{(t_f - t_0) s K} \) for any \( N \).

Lemma 3. Suppose (A2) holds. Then the first-order partial derivative \( \nabla \ell \) is Lipschitz continuous in \((x_{\pi}, x_{\sigma}, u_{\sigma})\), for fixed \( t_k \) and small enough \( h \), in a neighborhood of the continuous-time solution \((x^*(t_k), u^*(t_k))\), \( k = 0, 1, \ldots, N - 1 \).

Proof. The proof is similar to that of Lemma 1. It follows from the Lipschitz continuity of the first derivative of \( f \) given by (A2) and the state uniqueness property given by [20].

Theorem 2. Suppose that (A2) and (A5) hold and that (41) and (43)–(45) are satisfied. Then for all \((x_k, u_{ki})\) sufficiently close to the continuous-time solution \((x^*(t_k), u^*(t_k))\) and all sufficiently small \( h \), the sequence of IR iterates \( \{x_\pi^{(j)}, x_{\sigma}^{(j)}, u_{\sigma}^{(j)}, \lambda_\pi^{(j)}, \lambda_{\sigma}^{(j)}\} \) is locally convergent to the critical quintet \( \{x^*, x_{\sigma}^*, u_{\sigma}^*, \lambda_{\pi}^*, \lambda_{\omega}^*\} \) of problem (DP1). Furthermore, if \( \theta = \eta = 0 \), convergence of the iterates is r-quadric.

Proof. By the hypotheses and Lemma 2, (41)–(45) are satisfied. Also by Lemma 3, the remaining hypotheses of Theorem 3 in [4] hold, yielding the first conclusion. The second conclusion is then provided by Theorem 5 in [4].

Corollary 2. Suppose (A1)–(A6) hold and (41) and (43)–(45) are satisfied. As \( h \to 0 \), the solution \( \{x_\pi^*, \varphi(x_{\pi}^*, \lambda_{\pi}^*), \lambda_{\pi}^*\} \) found by the IR method locally tends to the critical triplet \( (x^*, u^*, \lambda^*) \).

Proof. The statement is furnished by Theorem 2 and Corollary 1.

5. Numerical implementation. Convergence theory of IR for Runge–Kutta discretization of optimal control problems presented in this paper holds locally, namely, we assume that the initial guess with which we start the IR iterations is “close enough” to a solution. It is well known that, in a small enough neighborhood of an isolated solution, the update direction of Newton’s method, which we refer to as a Newton step, is a “descent direction” along which improvement is guaranteed. In our implementation of the IR method for optimal control problems, we use as many steps of Newton’s method as necessary in the feasibility and optimality phases until (41) and (44) are satisfied for fixed values of \( \theta \) and \( \eta \), respectively.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
The Newton step for the feasibility phase (with fixed $u_\sigma$) is given by
\begin{equation}
\delta F(x_\pi, x_\sigma) = -\nabla_{(x_\pi, x_\sigma)} \ell(x_\pi, x_\sigma, u_\sigma) \ell^{-1}(x_\pi, x_\sigma, u_\sigma),
\end{equation}
where $\nabla_{(x_\pi, x_\sigma)} \ell$ stands for the Jacobian of $\ell$ with respect to $(x_\pi, x_\sigma)$. Recall that the control $u_\sigma$ is not updated in the feasibility phase. Note that, by (A2), $f$ is bounded, and so, for small enough $h$, the matrix $\nabla_{(x_\pi, x_\sigma)} \ell(x_\pi, x_\sigma, u_\sigma)$ is invertible.

Before expressing the Newton step for the optimality phase, let us define
\begin{equation}
g_1(z_\pi, z_\sigma, v_\sigma) := (z_\pi^T - x_\sigma^T, (\nabla \ell(y_\pi, y_\sigma, u_\sigma)(z_\pi - y_\pi, z_\sigma - y_\sigma, v_\sigma - u_\sigma))^T)
\end{equation}
and
\begin{equation}
g(z_\pi, z_\sigma, \mu_\pi, \mu_\sigma, v_\sigma) = (g_1^T(z_\pi, z_\sigma, v_\sigma), G^T(x_\sigma, u_\sigma, \lambda_\pi, \lambda_\sigma))^T.
\end{equation}

The Newton step for the optimality phase is then given by
\begin{equation}
\delta O(z_\pi, z_\sigma, \mu_\pi, \mu_\sigma, v_\sigma) = -[\nabla g(z_\pi, z_\sigma, \mu_\pi, \mu_\sigma, v_\sigma)]^{-1}g(z_\pi, z_\sigma, \mu_\pi, \mu_\sigma, v_\sigma),
\end{equation}
where $\nabla g$ stands for the Jacobian of $g$ with respect to all of its arguments.

We describe an implementation of the proposed IR method for Runge–Kutta discretization of optimal control problems in the so-called IR algorithm below. In the algorithm, (42)–(43) and (45) are not checked with specified constants $K_1$, $K_2$, and $K_3$. One should recall that (42) is satisfied automatically for some finite $K_1$ by Lemma 2. (43) and (45) are verified to hold with finite (but unspecified) $K_2$ and $K_3$, respectively, in the numerical experiments presented in the following section. In an alternative implementation, one may as well choose to specify the values of the constants $K_1$, $K_2$, and $K_3$.

It is worth pointing to differences of the algorithm here and the implementation in [24]. In [24], where only Euler discretization is considered, the inexactness parameters are set to zero, namely, that $\theta = \eta = 0$. In other words, the feasibility and optimality phases are carried out exactly. In the present implementation, however, apart from the much more general Runge–Kutta discretization, we allow inexact computations in both phases by specifying nonzero values of $\theta$ and $\eta$. Furthermore, in [24], $K_2 = 0$, while we don’t have such a restriction here either.

IR Algorithm

Step 0 (initialization) Choose the inexactness parameters, $\theta, \eta \in [0, 1]$, set the (small enough) stopping criterion tolerances, $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 > 0$, and the IR iteration index, $k = 0$. Choose the initial guesses $x^{(0)}_\pi$, $x^{(0)}_\sigma$, $\lambda^{(0)}_\pi$, $\lambda^{(0)}_\sigma$, and $u^{(0)}_\sigma$.

Step k.1 (feasibility phase) Set $j = 0$ and $\tilde{x}^{(0)} = (x^{(k)}_\pi, x^{(k)}_\sigma, u^{(k)}_\sigma)$.

(a) Update $\tilde{x}^{(j)}$ using the Newton step $\delta F$ given in (56):
\begin{equation}
\tilde{x}^{(j+1)} = \tilde{x}^{(j)} + \delta F(\tilde{x}^{(j)})).
\end{equation}

(b) Set $(y^{(k)}_\pi, y^{(k)}_\sigma, u^{(k)}_\sigma) = \tilde{x}^{(j+1)}$. If (41) is satisfied, then continue with Step k.2; otherwise set $j = j + 1$, and go to Step k.1(a).

Step k.2 (optimality phase) Set $j = 0$ and $\tilde{w}^{(0)} = (y^{(k)}_\pi, y^{(k)}_\sigma, \lambda^{(k)}_\pi, \lambda^{(k)}_\sigma, u^{(k)}_\sigma)$.

(a) Update $\tilde{w}^{(j)}$ using the Newton step $\delta O$ given in (57):
\begin{equation}
\tilde{w}^{(j+1)} = \tilde{w}^{(j)} + \delta O(\tilde{w}^{(j)}).
\end{equation}
(b) Set \( (z^{(k)}_\pi, z^{(k)}_\sigma, \mu^{(k)}_\pi, \mu^{(k)}_\sigma, \nu^{(k)}_\sigma) = \tilde{w}^{(j+1)} \). If (44) is satisfied, then continue with Step k.3; otherwise set \( j = j + 1 \), and go to Step k.2(a).

**Step k.3 (stopping criterion)**  If

\[
\|\ell(x^{(k)}_\pi, x^{(k)}_\sigma, u^{(k)}_\sigma)\| \leq \varepsilon_1, \quad \|g(z^{(k)}_\pi, z^{(k)}_\sigma, \mu^{(k)}_\pi, \mu^{(k)}_\sigma, \nu^{(k)}_\sigma)\| \leq \varepsilon_2,
\]

\[
\|\delta F(y^{(k)}_\pi, y^{(k)}_\sigma)\|_\infty + \|\delta O(z^{(k)}_\pi, z^{(k)}_\sigma, \mu^{(k)}_\pi, \mu^{(k)}_\sigma, \nu^{(k)}_\sigma)\|_\infty \leq \varepsilon_3,
\]

\[
\|\left( \left( z^{(k)}_\pi, z^{(k)}_\sigma, \mu^{(k)}_\pi, \mu^{(k)}_\sigma, \nu^{(k)}_\sigma \right) - \left( x^{(k)}_\pi, x^{(k)}_\sigma, \lambda^{(k)}_\pi, \lambda^{(k)}_\sigma, u^{(k)}_\sigma \right) \right) \|_\infty \leq \varepsilon_4,
\]

then STOP. Otherwise, update the current iterate,

\[
(x^{(k+1)}_\pi, x^{(k+1)}_\sigma, \lambda^{(k+1)}_\pi, \lambda^{(k+1)}_\sigma, u^{(k+1)}_\sigma) = (z^{(k)}_\pi, z^{(k)}_\sigma, \mu^{(k)}_\pi, \mu^{(k)}_\sigma, \nu^{(k)}_\sigma),
\]

set \( k = k + 1 \), and go to Step k.1.

### 5.1. Example application.

The van der Pol system has been used as a test problem in various optimal control studies [26, 24, 27, 32]. We consider the van der Pol system with unbounded control,

\[
\begin{align*}
\dot{x}_1(t) &= x_2(t), \\
\dot{x}_2(t) &= -x_1(t) - (x_1^2(t) - 1)x_2(t) + u(t),
\end{align*}
\]

where \( x(0) = (-2, 4) \), with the aim of minimizing the quadratic cost

\[
\frac{1}{2} \int_0^1 (x_1^2(t) + x_2^2(t) + u^2(t))dt.
\]

We consider the three-stage Gauss–Legendre discretization (an approximation of order 6) for this problem. For comparison purposes, we implement both the IR method (using the IR algorithm) and Newton’s method (as applied to (9)–(12)).

We carry out IR iterations with four different sets of values of \( \theta \) and \( \eta \), satisfying the IR sufficiency conditions (41) and (44), respectively. Namely, we consider the following cases:

- (I) \( \theta = 0.001, \eta = 0.001 \);
- (II) \( \theta = 0.3, \eta = 0.001 \) (inexact feasibility phase);
- (III) \( \theta = 0.001, \eta = 0.3 \) (inexact optimality phase);
- (IV) \( \theta = 0.3, \eta = 0.3 \) (inexact feasibility and optimality phases).

Note that the choice of relatively small values for \( \theta \) and \( \eta \) in case I constitutes an almost exact case. We emphasize inexactness in either or both of the feasibility and optimality phases in cases II–IV.

In Table 1, the IR iterations with case (I) are summarized, along with Newton iterations, for comparison. In Tables 2–4, the IR iterations with cases (II)–(IV) are summarized. In each of Tables 1–4, we list the following values. Note that the optimal solution appearing in (60) has been found numerically, correct to 16 decimal places.
Table 1

Case I—IR ($\theta = 0.001$, $\eta = 0.001$) and Newton iterations for the three-stage Gauss–Legendre discretization (order 6) with $N = 100$.

| $k$ | $K_1^{(e)}$ | $K_2^{(e)}$ | $K_3^{(e)}$ | $n_F$ | $n_O$ | $d^{(e)}$ | $\frac{d^{(e)}}{(d^{(e-1)})^2}$ | cost | $\|\ell^{(e)}\|_\infty$ | $d^{(e)}$ | $\frac{d^{(e)}}{(d^{(e-1)})^2}$ | cost | $\|\ell^{(e)}\|_\infty$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.73</td>
<td>0.75</td>
<td>912</td>
<td></td>
<td></td>
<td>$1.2 \times 10^4$</td>
<td>2.50000000000000</td>
<td>2.0 $\times 10^9$</td>
<td>1.2 $\times 10^4$</td>
<td>2.50000000000000</td>
<td>2.0 $\times 10^9$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>51.6</td>
<td>0.00</td>
<td>421</td>
<td>1</td>
<td>4</td>
<td>$4.4 \times 10^7$</td>
<td>6.11798120801632</td>
<td>3.2 $\times 10^{-2}$</td>
<td>5.7 $\times 10^9$</td>
<td>0.0392</td>
<td>8.5511370751078</td>
<td>8.9 $\times 10^{-2}$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>34.9</td>
<td>0.00</td>
<td>572</td>
<td>2</td>
<td>2</td>
<td>$1.7 \times 10^{-1}$</td>
<td>0.0086</td>
<td>5.55436494433536</td>
<td>2.6 $\times 10^{-1}$</td>
<td>8.3 $\times 10^{-1}$</td>
<td>0.0256</td>
<td>5.60342654181445</td>
<td>2.2 $\times 10^{-1}$</td>
</tr>
<tr>
<td>3</td>
<td>20.7</td>
<td>0.00</td>
<td>395</td>
<td>1</td>
<td>1</td>
<td>$1.8 \times 10^{-3}$</td>
<td>0.0062</td>
<td>5.54457146327757</td>
<td>5.5 $\times 10^{-1}$</td>
<td>1.4 $\times 10^{-1}$</td>
<td>0.0201</td>
<td>5.54447010131239</td>
<td>1.1 $\times 10^{-1}$</td>
</tr>
<tr>
<td>4</td>
<td>19.0</td>
<td>0.45</td>
<td>402</td>
<td>1</td>
<td>1</td>
<td>$1.2 \times 10^{-10}$</td>
<td>0.0088</td>
<td>5.54456410155263</td>
<td>3.5 $\times 10^{-12}$</td>
<td>2.5 $\times 10^{-12}$</td>
<td>0.0127</td>
<td>5.5445641474568</td>
<td>1.7 $\times 10^{-12}$</td>
</tr>
<tr>
<td>5</td>
<td>11.9</td>
<td>0.99</td>
<td>375</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5.54456410154876</td>
<td>0</td>
<td>7.6 $\times 10^{-14}$</td>
<td>0.0126</td>
<td>5.54456410154877</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0126</td>
<td>5.54456410154877</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


**Table 2**

<table>
<thead>
<tr>
<th>$k$</th>
<th>$K_1^{(k)}$</th>
<th>$K_2^{(k)}$</th>
<th>$K_3^{(k)}$</th>
<th>$n_F$</th>
<th>$n_O$</th>
<th>$d^{(k)}$</th>
<th>$d^{(k)}_{(\ell(k)-1)}$</th>
<th>cost</th>
<th>$|\ell^{(k)}|_\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.72</td>
<td>0.03</td>
<td>544</td>
<td>1</td>
<td></td>
<td>$1.2 \times 10^4$</td>
<td>2.50000000000000</td>
<td>2.0 $\times 10^{-9}$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>27.7</td>
<td>0.00</td>
<td>450</td>
<td>2</td>
<td>1</td>
<td>$8.9 \times 10^9$</td>
<td>0.0614</td>
<td>7.69882257104077</td>
<td>2.3 $\times 10^{-1}$</td>
</tr>
<tr>
<td>2</td>
<td>14.4</td>
<td>0.00</td>
<td>671</td>
<td>1</td>
<td>2</td>
<td>$7.9 \times 10^{-1}$</td>
<td>0.0100</td>
<td>5.66231962934820</td>
<td>4.7 $\times 10^{-2}$</td>
</tr>
<tr>
<td>3</td>
<td>16.0</td>
<td>0.00</td>
<td>216</td>
<td>2</td>
<td>1</td>
<td>$2.7 \times 10^{-3}$</td>
<td>0.0043</td>
<td>5.54469909627560</td>
<td>1.4 $\times 10^{-4}$</td>
</tr>
<tr>
<td>4</td>
<td>19.1</td>
<td>0.47</td>
<td>851</td>
<td>1</td>
<td>1</td>
<td>$8.3 \times 10^{-8}$</td>
<td>0.0111</td>
<td>5.54456410520833</td>
<td>1.2 $\times 10^{-9}$</td>
</tr>
<tr>
<td>5</td>
<td>3.00</td>
<td>0.01</td>
<td>18.1</td>
<td>1</td>
<td>1</td>
<td>$0$</td>
<td>5.54456410154876</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Case III—IR iterations ($\theta = 0.3$, $\eta = 0.001$) for the three-stage Gauss–Legendre discretization (order 6) with $N = 100$.**

**Table 3**

<table>
<thead>
<tr>
<th>$k$</th>
<th>$K_1^{(k)}$</th>
<th>$K_2^{(k)}$</th>
<th>$K_3^{(k)}$</th>
<th>$n_F$</th>
<th>$n_O$</th>
<th>$d^{(k)}$</th>
<th>$d^{(k)}_{(\ell(k)-1)}$</th>
<th>cost</th>
<th>$|\ell^{(k)}|_\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.73</td>
<td>0.75</td>
<td>912</td>
<td>1</td>
<td></td>
<td>$1.2 \times 10^4$</td>
<td>2.50000000000000</td>
<td>2.0 $\times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>51.0</td>
<td>0.00</td>
<td>439</td>
<td>4</td>
<td>1</td>
<td>$3.4 \times 10^9$</td>
<td>0.0304</td>
<td>6.11798120801632</td>
<td>3.2 $\times 10^{-7}$</td>
</tr>
<tr>
<td>2</td>
<td>44.9</td>
<td>0.00</td>
<td>926</td>
<td>2</td>
<td>1</td>
<td>$3.1 \times 10^{-1}$</td>
<td>0.0101</td>
<td>5.55381249426666</td>
<td>2.8 $\times 10^{-8}$</td>
</tr>
<tr>
<td>3</td>
<td>25.0</td>
<td>0.00</td>
<td>614</td>
<td>1</td>
<td>1</td>
<td>$1.1 \times 10^{-3}$</td>
<td>0.0112</td>
<td>5.54459673099600</td>
<td>2.0 $\times 10^{-9}$</td>
</tr>
<tr>
<td>4</td>
<td>24.4</td>
<td>0.47</td>
<td>552</td>
<td>1</td>
<td>1</td>
<td>$9.4 \times 10^{-5}$</td>
<td>0.0077</td>
<td>5.54456410182752</td>
<td>1.7 $\times 10^{-10}$</td>
</tr>
<tr>
<td>5</td>
<td>3.00</td>
<td>0.01</td>
<td>25.1</td>
<td>1</td>
<td>1</td>
<td>$0$</td>
<td>5.54456410154876</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Case IV—IR iterations ($\theta = 0.3$, $\eta = 0.3$) for the three-stage Gauss–Legendre discretization (order 6) with $N = 100$.**

**Table 4**

<table>
<thead>
<tr>
<th>$k$</th>
<th>$K_1^{(k)}$</th>
<th>$K_2^{(k)}$</th>
<th>$K_3^{(k)}$</th>
<th>$n_F$</th>
<th>$n_O$</th>
<th>$d^{(k)}$</th>
<th>$d^{(k)}_{(\ell(k)-1)}$</th>
<th>cost</th>
<th>$|\ell^{(k)}|_\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.7</td>
<td>0.63</td>
<td>544</td>
<td>1</td>
<td></td>
<td>$1.2 \times 10^4$</td>
<td>2.50000000000000</td>
<td>2.0 $\times 10^{-7}$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>27.7</td>
<td>0.00</td>
<td>452</td>
<td>2</td>
<td>1</td>
<td>$8.9 \times 10^9$</td>
<td>0.0614</td>
<td>7.69882257104077</td>
<td>2.3 $\times 10^{-1}$</td>
</tr>
<tr>
<td>2</td>
<td>14.8</td>
<td>0.00</td>
<td>771</td>
<td>1</td>
<td>1</td>
<td>$1.2 \times 10^{-1}$</td>
<td>0.0101</td>
<td>5.60982983921010</td>
<td>4.8 $\times 10^{-2}$</td>
</tr>
<tr>
<td>3</td>
<td>20.5</td>
<td>0.00</td>
<td>547</td>
<td>1</td>
<td>1</td>
<td>$1.1 \times 10^{-2}$</td>
<td>0.0084</td>
<td>5.34489638391042</td>
<td>2.8 $\times 10^{-3}$</td>
</tr>
<tr>
<td>4</td>
<td>23.8</td>
<td>0.45</td>
<td>544</td>
<td>1</td>
<td>1</td>
<td>$8.6 \times 10^{-5}$</td>
<td>0.0068</td>
<td>5.34456412621439</td>
<td>1.6 $\times 10^{-5}$</td>
</tr>
<tr>
<td>5</td>
<td>9.00</td>
<td>0.49</td>
<td>80.0</td>
<td>1</td>
<td>1</td>
<td>$0$</td>
<td>5.34456410154876</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\quad & d^{(k)} := \|z_\pi - x^\pi\| + \|z_\sigma - x^\sigma\| + \|v_\pi - u^\pi\| + \|v_\sigma - u^\sigma\| + \|\mu_\pi - \lambda^\pi\| + \|\mu_\sigma - \lambda^\sigma\|, \\
(60) & K_1^{(k)} := \frac{\|y_\pi - x_\pi\| + \|y_\sigma - x_\sigma\|}{\|\ell(x_\pi, x_\sigma, y_\pi, y_\sigma)\|}, \\
(61) & K_2^{(k)} := \frac{\|\nabla \ell(y_\pi, y_\sigma, u_\pi, u_\sigma)\|}{\|G(y_\pi, u_\pi, \lambda_\pi, \lambda^\pi)\|}, \\
(62) & K_3^{(k)} := \frac{\|z_\pi - y_\pi\| + \|z_\sigma - y_\sigma\| + \|v_\pi - u_\pi\| + \|v_\sigma - u_\sigma\|}{\|\mu_\pi - \lambda^\pi\| + \|\mu_\sigma - \lambda^\sigma\|} / \|G(y_\pi, u_\pi, \lambda_\pi, \lambda^\pi)\|, \\
(63) & \end{align*}
\]
where the superscript $k$ has been omitted from the terms on the right-hand side for clarity. The norm $\| \cdot \|$ stands for $\| \cdot \|_{\infty}$. Note that we list the distance $d^{(k)}$ for both IR and Newton iterations. We also tabulate the cost, as the value given in (7), and the constraint violation, $\| \ell^{(k)} \|_{\infty}$, at each IR and Newton iteration. Note that the maxima of the sequences $K_1^{(k)}$, $K_2^{(k)}$, and $K_3^{(k)}$ throughout all of the iterations constitute roughly lower bounds for the parameters $K_1$, $K_2$, and $K_3$, respectively.

As the initial iterate (or initial guess), we have chosen, for both IR and Newton iterations,

\[
x_i^{(0)} = (-1, 2),
\]
\[
x_{ij}^{(0)} = (-1, 2),
\]
\[
\lambda_i^{(0)} = (0, 0),
\]
\[
\lambda_{ij}^{(0)} = (0, 0),
\]
\[
u_{ij}^{(0)} = 0,
\]

where $i = 0, 1, 2, \ldots, N$ and $j = 1, 2, \ldots, s$.

The optimal cost found in the iterations is correct to 14 decimal places. Table 1 demonstrates that IR and Newton methods take about the same number of steps to reach the same accuracy in the solution. We note from Tables 2–4 that the same solution is found in the same number of iterations with inexact computation of the feasibility and optimality phases of the IR method.

In Tables 1–4, $n_F$ and $n_O$ denote the number of times a Newton step has been performed in each of the feasibility and optimality phases, respectively (Steps $k.1(a)$ and $k.2(a)$ of the IR algorithm). In the almost exact case (I), $\Sigma n_F = 9$ and $\Sigma n_O = 6$ in the overall five iterations. In the inexact cases (II)–(IV), these numbers slightly decrease. In case (IV), where the degree of inexactness is the highest, only one Newton step is performed in each of the feasibility and optimality phases. One should also note that in case (IV), the solution is obtained in the same number of iterations as in the other cases. So, from a computational savings point of view, it is fair to say that case (IV) is the most desirable.

It is interesting to note that the constraint violations in case (IV) of the IR iterations are similar to those in the Newton iterations.

It is also interesting to note that very near the solution, $n_F = 1$ and $n_O = 1$, even for rather small values of the inexactness parameters $\theta$ and $\eta$, for example, in case (I).

When a number is too small to report, it is listed in the tables as 0.

The tabulated values of the quantity $d^{(k)}/(d^{(k-1)})^2$ in each of the tables corroborate the fact that the convergence rate of the IR method (even with relatively large $\theta$ and $\eta$) is at least quadratic, like that of Newton’s method. Kaya and Martínez [23] prove that the IR method with exact feasibility and optimality phases is the same as a projected Newton method in the case when the problem is quadratic (with a quadratic objective to minimize which is subject to quadratic equality constraints). It is interesting to note that Tapia and Whitley [39] establish superquadratic convergence of order $1 + \sqrt{2}$ for the projected Newton method applied to the eigenvalue problem of symmetric matrices, which is an equality constrained quadratic problem. Their discovery does not hold, however, for the eigenvalue problem of nonsymmetric matrices. Although the numerical example that we are experimenting with in this paper is quite more complicated than what Tapia and Whitley looked at, the “more than quadratic” appearing convergence we get warrants a further study of the IR
method and its convergence rate for some special types of nonlinear programming problems.

Going back to the choice of the Runge–Kutta scheme, one could have considered using the four-stage Lobatto IIIA instead of the three-stage Gauss–Legendre (see Appendix A) because both are symplectic and yield the same order of accuracy, although the latter involves fewer variables than the former. Note further that, when applied to the state equations (in the feasibility phase), the row of zeroes in the coefficient tableau of the Lobatto IIIA scheme, in particular the matrix \( A \) formed by \( a_{ij} \), means that one can obtain one of the coordinates explicitly in terms of the others. This would certainly simplify an implementation of Newton’s method and would make the Newton’s method applicable at probably an equal difficulty level as that of the three-stage Gauss–Legendre scheme. However, in the optimality phase, the same advantage is lost in the costate part of the equations because of the symplectic partitioned structure of the optimality equations: the discretization of the costate equations implied by optimality gives a coefficient tableau with a matrix \( \hat{A} \) (formed by \( \hat{a}_{ij} \)) which has no row of zeroes even if the matrix \( A \) (formed by \( a_{ij} \)) has a row of zeroes, namely, the Runge–Kutta scheme for the costate equations in this case would all be implicit. For this reason, the three-stage Gauss–Legendre stands out as the preferable scheme to use over the four-stage Lobatto IIIA. The same comparison can be carried out between the pairs, two-stage Gauss–Legendre and three-stage Lobatto IIIA, and Störmer–Verlet and midpoint schemes.

In our implementation of the Runge–Kutta discretization, we have prepared a computer code which does not exploit the situations where there is a row of zeroes in the coefficient tableau of the discretization scheme. In a more efficient implementation, taking into account the particular structure of a coefficient tableau may speed up the computations.

Next, using the van der Pol example, we numerically verify Theorem 1 in Tables 5 and 6 for the three-stage Gauss–Legendre and four-stage Lobatto IIIA discretization with \( \kappa = 6 \). We define

\[
c_N := \left( \max_{0 \leq k \leq N} \|x_k^* - x^*(t_k)\| + \max_{0 \leq k \leq N} \|\lambda_k^* - \lambda^*(t_k)\| + \max_{0 \leq k \leq N} \|u_k^* - u^*(t_k)\| \right) / h^\kappa.
\]

The tabulated values of \( c_N \) in both Tables 5 and 6 suggest that for each discretization, there exists a constant \( c > 0 \) (an upper bound of the sequence \( \{c_N\} \) such that the discretization error in the solutions for the states, costates, and control is bounded by \( c h^\kappa \).

It is well known that Newton’s method obeys the so-called mesh independence principle [1, 12]; in other words, its behavior does not change with the mesh size, i.e., with \( N \). In [24], it was observed through numerical experiments that IR, as applied to Euler discretization, obeys the mesh independence principle. We have observed in the numerical experiments we have had for the more general Runge–Kutta discretiza-

### Table 5

<table>
<thead>
<tr>
<th>( N )</th>
<th>( \max_{1 \leq k \leq N} |x_k^* - x^*(t_k)| )</th>
<th>( \max_{1 \leq k \leq N} |\lambda_k^* - \lambda^*(t_k)| )</th>
<th>( \max_{1 \leq k \leq N} |u_k^* - u^*(t_k)| )</th>
<th>( c_N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>( 3.9 \times 10^{-7} )</td>
<td>( 2.5 \times 10^{-6} )</td>
<td>( 1.2 \times 10^{-6} )</td>
<td>4.14</td>
</tr>
<tr>
<td>20</td>
<td>( 7.0 \times 10^{-7} )</td>
<td>( 4.7 \times 10^{-6} )</td>
<td>( 2.1 \times 10^{-6} )</td>
<td>4.86</td>
</tr>
<tr>
<td>40</td>
<td>( 1.1 \times 10^{-6} )</td>
<td>( 7.8 \times 10^{-6} )</td>
<td>( 3.5 \times 10^{-6} )</td>
<td>5.08</td>
</tr>
<tr>
<td>100</td>
<td>( 4.6 \times 10^{-6} )</td>
<td>( 3.2 \times 10^{-5} )</td>
<td>( 1.4 \times 10^{-5} )</td>
<td>5.14</td>
</tr>
</tbody>
</table>
tion of optimal control problems that IR still seems to obey the mesh independence principle.

6. Conclusion. The IR approach to optimal control problems can be extended to problems where one has path inequality constraints involving the states and controls. Convergence theory of the Euler discretization of the optimal control problem to the solution of the original (continuous) optimal control problem as the mesh size is made finer is furnished by Malanowski, Büskens, and Maurer for mixed state control inequality constraints in [29] and by Dontchev and Hager for pure state inequality constraints in [11]. In [29], an additional condition on the derivative of the constraint with respect to the control is assumed to hold, which does not hold for pure state constrained problems. So the results in both [11] and [29] can be taken into account, depending on the type of inequality constraints.

In this paper, we restricted the implementation of our approach to a simple illustrative, albeit test bed, example. The method otherwise is applicable to a wide range of problems, formulated as in (P), from various disciplines including engineering, economics, and biomedicine. With possible extension to constrained problems as mentioned in the previous paragraph, applicability of the method would be enlarged.

The more general class of problems to which one can further extend the IR approach is the class which involves PDEs instead of, or in addition to, ODEs. Because the IR method can exploit the structure of the problem, it would be particularly suited to tackling and studying such problems.

Appendix A. Some common Runge–Kutta schemes. The coefficients of the Runge–Kutta discretization (5)–(6) are usually displayed in the literature as follows:

\[
\begin{array}{cccccccc}
  c_1 & a_{11} & \cdots & a_{1s} \\
  \vdots & \vdots & & \vdots \\
  c_s & a_{s1} & \cdots & a_{ss} \\
  b_1 & \cdots & b_s
\end{array}
\]

The Runge–Kutta discretization both in (13)–(16) and (36)–(40) constitute symplectic partitioned Runge–Kutta schemes. We can employ any Runge–Kutta scheme satisfying the order conditions for optimal control in [5, 6, 20] and the condition \(b_i > 0\) (a requirement for convergence in Theorem 1). The resulting discretization for optimal control would be symplectic, which is a very desirable property for a discretization.

The following discretization schemes [21] are examples that satisfy the order conditions for optimal control and can thus be utilized.

- Explicit and implicit Euler (order 1):

\[
\begin{array}{c|c|c|c|c|c}
  0 & 0 & 1 & 1 & 1 & 1 \\
  1 & 1 & 1 & 1 & 1 & 1 \\
\end{array}
\]
• Midpoint (order 2):

\[
\begin{array}{c|cc}
1/2 & 1/2 & 1 \\
\end{array}
\]

• Implicit trapezoidal or Störmer–Verlet (order 2):

\[
\begin{array}{c|ccc}
0 & 0 & 0 & 0 \\
1 & 1/2 & 1/2 & 1/2 \\
\end{array}
\]

• Gauss–Legendre (order 4):

\[
\begin{array}{c|ccccc}
1/2 - \sqrt{3}/6 & 1/4 & 1/4 - \sqrt{3}/6 & 1/2 & 1/2 \\
1/2 + \sqrt{3}/6 & 1/4 + \sqrt{3}/6 & 1/4 & 1/2 & 1/2 \\
\end{array}
\]

• Lobatto IIIA (order 4):

\[
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
1/2 & 5/24 & 1/3 & -1/24 & 1/2 \\
1 & 1/6 & 2/3 & 1/6 & 1 \\
\end{array}
\]

• Gauss–Legendre (order 6):

\[
\begin{array}{c|cccccc}
1/2 - \sqrt{15}/10 & 5/36 & 2/9 - \sqrt{15}/15 & 5/36 & \sqrt{15}/30 & 1/2 & 1/2 \\
1/2 & 5/36 + \sqrt{15}/24 & 2/9 & 5/36 & \sqrt{15}/24 & 1/2 & 1/2 \\
1/2 + \sqrt{15}/10 & 5/36 + \sqrt{15}/30 & 2/9 + \sqrt{15}/15 & 5/36 & 1/2 & 1/2 \\
\end{array}
\]

• Lobatto IIIA (order 6):

\[
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
5 - \sqrt{5}/10 & 11 + \sqrt{5}/120 & 25 - \sqrt{5}/120 & 25 - 13\sqrt{5}/120 & -1 + \sqrt{5}/120 \\
5 + \sqrt{5}/10 & 11 - \sqrt{5}/120 & 25 + 13\sqrt{5}/120 & 25 + \sqrt{5}/120 & -1 - \sqrt{5}/120 \\
1 & 1/12 & 5/12 & 5/12 & 1/12 \\
\end{array}
\]

Acknowledgments. The author offers warm thanks to Bülent Karasözen for useful discussions on Runge–Kutta methods. He is also indebted to José Mario Martínez for motivating discussions on IR methods. Two anonymous reviewers are gratefully acknowledged for their constructive comments and suggestions which improved the paper.
REFERENCES


