THE EULER APPROXIMATION
IN STATE CONSTRAINED OPTIMAL CONTROL

A. L. DONTCHEV AND WILLIAM W. HAGER

ABSTRACT. We analyze the Euler approximation to a state constrained control problem. We show that if the active constraints satisfy an independence condition and the Lagrangian satisfies a coercivity condition, then locally there exists a solution to the Euler discretization, and the error is bounded by a constant times the mesh size. The proof couples recent stability results for state constrained control problems with results established here on discrete-time regularity. The analysis utilizes mappings of the discrete variables into continuous spaces where classical finite element estimates can be invoked.

1. INTRODUCTION

Discrete approximations to optimal control problems have been analyzed since the 1960s. The first work dealt with the convergence of the optimal value or an optimal control for the discrete problem to the continuous solution (see, e.g., [5], [7]–[13], and [33]). A survey of some of the earlier work is given by Polak in [34]. More recent results on convergence, based on consistent approximations and modern variational techniques, are contained in [35], [36], and [39]. For a survey of work in this area, see [16].

In this paper, we are concerned not only with convergence, but also with convergence rate. That is, for the Euler discretization of a state constrained control problem, we estimate the distance between a solution to the continuous problem and a solution to the discrete problem as a function of the mesh size. This estimate represents the first analysis for a discretization of a fairly general state constrained control problem. When the constraint qualification of [28] holds and the Lagrangian satisfies a local coercivity condition, we show that for a sufficiently fine mesh, the Euler discretization has a solution and corresponding Lagrange multipliers which are at distance $O(h)$ from a continuous solution/multiplier pair. Here distance is measured in the L^2 metric for the control and the constraint multiplier, and in the H^1 metric for the state and adjoint variables. By an embedding result, the error is $O(h^{2/3})$ in the L^∞ norm.

We now give a brief survey of earlier work on convergence rates for discrete approximations in optimal control. In [2] Bosarge and Johnson studied dual finite element approximations for unconstrained linear/quadratic problems obtaining error estimates of order $O(h^k)$ in the L^2 norm, where h is the mesh size and k is the

degree of the approximating piecewise polynomials. In [3] similar estimates were obtained for Ritz-Galerkin approximations of unconstrained nonlinear problems. In the series of papers [25], [27], and [30], Hager analyzed dual finite element approximations to convex constrained control problems (linear differential equation, convex cost function, convex control and state constraints) obtaining an $O(h)$ estimate in L^2 for piecewise linear splines, and an $O(h^{3/2})$ estimate for piecewise quadratics.

In the first paper [26] to consider the usual Range-Kutta and multistep integration schemes, Hager focused on unconstrained optimal control problems and determined the relationship between the continuous dual variables and the Kuhn-Tucker multipliers associated with the discrete problem. It was observed that an order k integration scheme for the differential equation did not always lead to an order k discrete approximation. In fact, for some integration schemes, the discrete approximation did not converge to a solution of the continuous problem as the mesh was refined; for related work following these results see [24]. In [14] (see also [15, Chapter 4]) Dontchev analyzed Euler's approximation to a constrained convex control problem obtaining an error estimate of order $O(h)$ in the L^2 norm, where h is the size of the uniform mesh. In [18] we analyzed nonlinear optimal control problems with *control constraints*, obtaining an $O(h)$ estimate in L^∞ for the error in the Euler discretization. Most recently, in [29] the convergence rate is determined for general Runge-Kutta discretizations of control constrained optimal control problems. These conditions on the coefficients in the Runge-Kutta scheme determine whether the discrete (approximating) solution is second-, third-, or fourth-order accurate. In [29] it is assumed that the coefficients in the final stage of the Runge-Kutta scheme are all positive, while in [21] this positivity requirement is removed for second-order Runge-Kutta schemes by imposing additional conditions on the coefficients.

In [17] Dontchev obtained an estimate for the distance from a solution to the discrete problem to a solution of the continuous problem by making assumptions for the discrete solutions rather than for the continuous solution. In [32] Malanowski, Büskens, and Maurer obtained error estimates for a nonlinear problem with mixed control and state constraints. In their analysis, it is assumed that the derivative of the constraint with respect to the control satisfies a surjectivity condition which does not hold for pure state constrained problems. In [38] Veliov examined a Runge-Kutta discretization of a nonlinear optimal control problem with control constraints obtaining higher-order estimates for the sets of feasible controls and for the optimal value.

Our approach in this paper for the analysis of state constrained control problems is that presented in [18]. Loosely speaking, we show that the solution of the linearized first-order optimality conditions for the discrete control problem is stable under perturbation, and that the linear operator is sufficiently close to the nonlinear operator. These two results combine to give the error estimate. In carrying out the analysis, many technicalities arise. For example, the coercivity condition for the Lagrangian is naturally posed is L^2; however, the cost function does not have derivatives in L^2. This forces us to work in a nonlinear space of functions that are Lipschitz continuous with derivatives bounded by some fixed number. In this nonlinear setting, L^2 convergence implies L^∞ convergence. In order to show that the analysis can be carried out in this nonlinear space, we need to establish a discrete regularity result. That is, if the linearized discrete problem is perturbed, then discrete derivatives of the solution can be bounded by discrete derivatives of

the perturbation. This regularity result is the discrete analogue of the continuous regularity result given in [28].

To analyze the difference between the nonlinear first-order conditions and their linearization, we transform from discrete variables to functions continuous in time using various interpolation operators. This allows us to perform the analysis in continuous time, and to use finite element techniques to analyze the continuous expressions. Also, embeddings associated with continuous spaces can be used to deduce, through interpolation, corresponding embeddings in the discrete setting.

To briefly summarize the paper, Section 2 formulates the state constrained problem and its discrete approximation, and presents the main error estimate. This error estimate is based on an abstract existence theorem given in Section 3. In Section 4 we summarize the various finite element estimates and embeddings that are used in the analysis. Sections 5 through 8 show that each of the assumptions of the abstract theorem are satisfied, while Section 9 pulls together the analysis and proves the error estimate for the Euler discretization. A numerical example is given in Section 10.

2. The problem and its discretization

We consider the following optimal control problem:

(1)
$$\text{minimize } \int_0^1 \varphi(x(t), u(t))dt$$

$$\text{subject to } \quad \dot{x}(t) = f(x(t), u(t)) \text{ a.e. } t \in [0, 1], \quad x(0) = x^0,$$

$$g(x(t)) \leq 0 \text{ for all } t \in [0, 1], \quad u \in L^\infty, \quad x \in W^{1,\infty},$$

where the state $x(t) \in \mathbf{R}^n$, $\dot{x} \equiv \frac{d}{dt}x$, the control $u(t) \in \mathbf{R}^m$, the functions $\varphi : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n$, and $g : \mathbf{R}^n \to \mathbf{R}^k$.

Throughout the paper, $L^\alpha(J; \mathbf{R}^n)$ denotes the usual Lebesgue space of measurable functions $x : J \to \mathbf{R}^n$ with $|x(\cdot)|^\alpha$ integrable, equipped with its standard norm

$$\|x\|_{L^\alpha} = (\int_J |x(t)|^\alpha dt)^{1/\alpha},$$

where $|\cdot|$ is the Euclidean norm. Of course, $\alpha = \infty$ corresponds to the space of essentially bounded, measurable functions equipped with the essential supremum norm. Further, $W^{m,\alpha}(J; \mathbf{R}^n)$ is the Sobolev space consisting of vector-valued functions $x : J \to \mathbf{R}^n$ whose j-th derivative lies in L^α for all $0 \leq j \leq m$ with the norm

$$\|x\|_{W^{m,\alpha}} = \sum_{j=0}^m \|x^{(j)}\|_{L^\alpha}.$$

When either the domain J or the range $\mathbf{R}^n$ is clear from context, it is omitted. We let H^m denote the space $W^{m,2}$, and Lip denote $W^{1,\infty}$, the space of Lipschitz continuous functions. Throughout, c is a generic constant, that has different values in different equations, and which is independent of time and the mesh spacing in the approximating problem. The transpose of a matrix A is A^T, and $\mathbf{B}_a(x)$ is the closed ball centered at x with radius a. Given a vector $y \in \mathbf{R}^m$ and a set $\mathcal{A} \subset \{1, 2, \ldots, m\}$, $y_{\mathcal{A}}$ denotes the subvector consisting of components associated with indices in $\mathcal{A}$. And if $Y \in \mathbf{R}^{m \times n}$, then $Y_{\mathcal{A}}$ is the submatrix consisting of rows associated with indices in $\mathcal{A}$. The complement of the set $\mathcal{A}$ is $\mathcal{A}^c$.

We now present the assumptions that are employed in our analysis of the Euler discretization of (1). The first assumption is related to the regularity of the solution and the problem functions, and the solution of the associated optimality system.

Smoothness. *Problem* (*1*) *has a local solution* (x^*, u^*) *which lies in* $W^{2,\infty} \times W^{1,\infty}$. *There exists an open set* $\Omega \subset \mathbf{R}^n \times \mathbf{R}^m$ *and* $\rho > 0$ *such that* $\mathbf{B}_\rho(x^*(t), u^*(t)) \subset \Omega$ *for every* $t \in [0, 1]$, *and the first two derivatives of* φ *and* f, *and the first three derivatives of* g *are Lipschitz continuous in* Ω. *Finally, there exist associated Lagrange multipliers* $\psi^* \in W^{2,\infty}$ *and* $\nu^* \in W^{1,\infty}$ *for which the following form of the first-order optimality conditions (minimum principle) is satisfied at* $(x^*, \psi^*, u^*, \nu^*)$:

(2) $$\dot{x} = f(x, u), \quad x(0) = x^0,$$

(3) $$\dot{\psi} = -\nabla_x H(x, \psi, u, \nu), \quad \psi(1) = 0,$$

(4) $$0 = \nabla_u H(x, \psi, u, \nu),$$

(5) $$g(x) \in \mathcal{N}(\nu), \quad \nu(1) \le 0, \quad \dot{\nu} \ge 0.$$

Here and elsewhere, multipliers such as ψ and ν are treated as row vectors, H is the Hamiltonian defined by

$$H(x, \psi, u, \nu) = \varphi(x, u) + \psi f(x, u) - \nu \nabla g(x) f(x, u),$$

and the set-valued map $\mathcal{N}$ is understood in the following way: given a nondecreasing Lipschitz continuous function ν, a continuous function y lies in $\mathcal{N}(\nu)$ if and only if

(6) $$y(t) \le 0 \quad \text{and} \quad \dot{\nu}(t)y(t) = 0 \text{ for a.e. } t \in [0, 1], \quad \text{and} \quad \nu(1)y(1) = 0.$$

In the terminology of [31], the form of the minimum principle we employ is the "indirect adjoining approach with continuous adjoint function." Typically, the multiplier ν, associated with the state constraint, and the derivative of ψ have bounded variation. In our statement of the minimum principle above, we are assuming some additional regularity so that ν and $\dot{\psi}$ are not only of bounded variation, but Lipschitz continuous. As shown in [28] for a linear-convex problem (also see [20]), the assumed regularity of both the solution and the Lagrange multipliers is a consequence of the Uniform Independence and Coercivity conditions introduced below.

Note that problem (1) is posed in L^∞ and the elements of L^∞ are equivalence classes of functions. By the Smoothness assumption, there exists a solution to the control problem in the equivalence class associated with u^* such that the optimality conditions (2), (3), and (4) are satisfied everywhere in $[0, 1]$.

Let A, B, and K be the matrices defined by

$$A = \nabla_x f(x^*, u^*), \quad B = \nabla_u f(x^*, u^*), \quad \text{and } K = \nabla g(x^*).$$

Let $\mathcal{A}(t)$ be the set of indices of the active constraints at $x^*(t)$:

$$\mathcal{A}(t) = \big\{ j \in \{1, 2, \cdots, k\} : g_j(x^*(t)) = 0 \big\}.$$

Our next assumption relates to the stability of the state constraint (see [19]).

Independence at $\mathcal{A}$. *The set* $\mathcal{A}(0)$ *is empty and there exists a scalar* $\beta > 0$ *such that*

$$\Big| \sum_{j \in \mathcal{A}(t)} v_j K_j(t) B(t) \Big| \ge \beta |v_{\mathcal{A}(t)}|$$

for each $t \in [0, 1]$ *where* $\mathcal{A}(t) \ne \emptyset$ *and for each choice of* v.

Defining

$$Q^* = \nabla_{xx} H(w^*), \quad M^* = \nabla_{xu} H(w^*), \quad \text{and } R^* = \nabla_{uu} H(w^*),$$

where $w^* = (x^*, \psi^*, u^*, \nu^*)$, let $\mathcal{B}^*$ be the quadratic form defined by

$$\mathcal{B}^*(x, u) = \frac{1}{2} \int_0^1 x(t)^\mathsf{T} Q^*(t) x(t) + u(t)^\mathsf{T} R^*(t) u(t) + 2x(t)^\mathsf{T} M^*(t) u(t) dt.$$

Our third assumption is a growth condition.

Coercivity. *There exists a constant $\alpha > 0$ such that*

$$\mathcal{B}^*(x, u) \geq \alpha \|u\|_{L^2}^2 \quad \text{for all } (x, u) \in \mathcal{M}^*,$$

where

$$\mathcal{M}^* = \{(x, u) : x \in H^1, u \in L^2, \ \dot{x} - Ax - Bu = 0, \ x(0) = 0\}.$$

Coercivity is a strong form of a second-order sufficient optimality condition in the sense that when combined with Independence, it implies not only optimality, but also Lipschitzian dependence of the solution and the multipliers with respect to parameters (see [19]). For recent work on second-order sufficient conditions, see [23] and [40].

We now introduce the Euler discretization of (1). If N is a natural number and $h = 1/N$, we consider the following discrete problem:

(7)   minimize $\displaystyle\sum_{i=0}^{N-1} \varphi(x_i, u_i)$

subject to $\quad x_i' = f(x_i, u_i), \quad x_0 = x^0, \quad g(x_i) \leq 0, \quad 0 \leq i \leq N - 1.$

Here the prime is shorthand notation for the forward difference

$$x_i' = \frac{x_{i+1} - x_i}{h}.$$

Throughout the analysis, we let t_i stand for ih, and we use the subscript i to denote the time level i. Hence, $x_i \in \mathbf{R}^n$ is the discrete analogue of $x(t_i)$, while x_j denotes the j-th component of the vector $x \in \mathbf{R}^n$. Although the final constraint $g(x_N) \leq 0$ is not imposed in the discrete problem (7), there are no significant changes in the analysis if this final constraint is included.

The first-order necessary optimality conditions associated with (7), often called the Karush-Kuhn-Tucker conditions in this finite dimensional context, can be written (see [1])

(8) $x_i' = f(x_i, u_i), \quad x_0 = x^0,$

(9) $p_{i-}' = -\nabla_x \mathcal{H}(x_i, u_i, p_i) - \mu_i \nabla g(x_i), \quad p_{N-1} = 0,$

(10) $0 = \nabla_u \mathcal{H}(x_i, u_i, p_i),$

(11) $g(x_i) \in \mathcal{N}_{\mathbf{R}_+^k}(\mu_i), \quad \mu_i \geq 0,$

where $0 \leq i \leq N - 1$. Here the multipliers p_i and μ_i are row vectors, the discrete Hamiltonian is defined by

$$\mathcal{H}(x, u, p) = \varphi(x, u) + pf(x, u),$$

p_{i-}' is shorthand for the backward difference

$$p_{i-}' = \frac{p_i - p_{i-1}}{h},$$

and $\mathcal{N}_{\mathbf{R}_+^k}(\mu_i)$ denotes the normal cone to the positive orthant $\mathbf{R}_+^k$ at the point $\mu_i \geq 0$:

$$y \in \mathcal{N}_{\mathbf{R}_+^k}(\mu_i) \quad \text{if and only if} \quad y \leq 0 \text{ and } \mu_i y = 0.$$

In formulating the necessary conditions, p_i is the multiplier associated with the constraint $f(x_i, u_i) - x_i' = 0$. The multiplier p_{-1} can be identified with an artificial constraint $(x_0 - x^0)/h = 0$, and μ_i is the multiplier for the constraint $g(x_i) \leq 0$.

In order to relate the continuous first-order conditions (2)–(5) to the discrete conditions (8)–(11), we introduce transformed dual variables

$$(12) \qquad \nu_i = -h \sum_{l=i}^{N} \mu_l \quad \text{and} \quad \psi_i = p_i + \nu_{i+1} \nabla g(x_{i+1}), \quad \text{where} \quad \mu_N = 0.$$

The last equation is the definition of μ_N since the constraint $g(x_N) \leq 0$ is not imposed in discrete problem (7), and we take the associated multiplier to be zero. Observe that $\nu_N = 0$ and the variable μ in (12) is the (discrete) derivative of ν: $\mu = \nu'$. After making the substitutions $\mu_i = \nu_i'$ and $p_i = \psi_i - \nu_{i+1} \nabla g(x_{i+1})$, the optimality system takes the following form:

$$(13) \qquad\qquad x_i' = f(x_i, u_i), \quad x_0 = x^0,$$

$$(14) \qquad\qquad \psi_{i-}' = -\nabla_x \mathcal{H}(x_i, u_i, \psi_i) + \mathcal{P}_i, \quad \psi_{N-1} = 0,$$

$$(15) \qquad\qquad 0 = \nabla_u \mathcal{H}(x_i, u_i, \psi_i) - \nu_{i+1} \nabla g(x_{i+1}) \nabla_u f(x_i, u_i),$$

$$(16) \qquad\qquad g(x_i) \in \mathcal{N}_{\mathbf{R}_+^k}(\nu_i'),$$

$0 \leq i \leq N - 1$, where

$$(17) \qquad\qquad \mathcal{P}_i = \nu_{i+1} \nabla g(x_i)' + \nu_{i+1} \nabla g(x_{i+1}) \nabla_x f(x_i, u_i).$$

In order to analyze the discrete problem (7), we need to introduce discrete analogues of various continuous spaces and norms. In particular, for a sequence $z_0, z_1, \dots, z_N$ whose i-th element is a vector $z_i \in \mathbf{R}^n$, the discrete analogues of the L^2, L^∞, and H^1 norms are the following:

$$\|z\|_{L^2} = \sqrt{\sum_{i=0}^{N} h|z_i|^2}, \quad \|z\|_{L^\infty} = \sup_{0 \leq i \leq N} |z_i|, \quad \text{and} \quad \|z\|_{H^1} = \sqrt{\|z\|_{L^2}^2 + \|z'\|_{L^2}^2},$$

where z' is the sequence whose i-th element is the forward difference $(z_{i+1} - z_i)/h$. Estimates are obtained for the discrete state sequence x_i and multiplier sequence ν_i where i ranges from zero to N and for the control sequence u_i and multiplier sequences p_i and ψ_i where i ranges from zero to $N - 1$. When taking the norm of any of these sequences, we assume that the index range is chosen appropriately.

Our main result is the following estimate for the error in the discrete approximation. In stating this result, our convention is that when both a discrete and a continuous variable appear in an expression, then the continuous variable is treated as a discrete variable whose components are the continuous variable evaluated at the mesh points, the t_i. That is, if u^h is a discrete variable and u^* is continuous, then $u^h - u^*$ is discrete with $(u^h - u^*)_i = u_i^h - u^*(t_i)$. Also, we say that a discrete variable u is Lipschitz continuous in (discrete) time with Lipschitz constant ξ if $|u_i'| \leq \xi$ for each i.

Theorem 2.1. *If Smoothness, Independence at $\mathcal{A}$, and Coercivity hold, then for all sufficiently small h, there exists a local solution (x^h, u^h) of the discrete optimal control problem (γ) and associated Lagrange multipliers (ψ^h, ν^h) such that*

$$(18) \qquad \|x^h - x^*\|_{H^1} + \|u^h - u^*\|_{L^2} + \|\psi^h - \psi^*\|_{H^1} + \|\nu^h - \nu^*\|_{L^2} \le ch,$$

and

(19)

$$\|x^h - x^*\|_{W^{1,\infty}} + \|u^h - u^*\|_{L^\infty} + \|\psi^h - \psi^*\|_{W^{1,\infty}} + \|\nu^h - \nu^*\|_{L^\infty} \le ch^{2/3}.$$

Moreover, $(x^h)'$, u^h, $(\psi^h)'$, and ν^h are Lipschitz continuous in time with a Lipschitz constant independent of h.

3. ABSTRACT SETTING

Our proof of Theorem 2.1 is based on the following abstract result which is related to [19, Lemma 2.1].

Theorem 3.1. *Let $\mathcal{X}$ be a complete metric space with metric ρ, let $\mathcal{Y}$ be a linear normed space with norm $\|\cdot\|$, and let Π be a subset of $\mathcal{Y}$. Suppose that $\mathcal{T} : \mathcal{X} \mapsto \mathcal{Y}$, $\mathcal{L} : \mathcal{X} \mapsto \mathcal{Y}$, and $\mathcal{F} : \mathcal{X} \mapsto 2^{\mathcal{Y}}$, and that for some given $w^* \in \mathcal{X}$, $\delta^* \in \mathcal{Y}$, and scalars ε, λ, $r > 0$, we have*

(P1) $\mathcal{T}(w^*) + \delta^* \in \mathcal{F}(w^*)$ *and* $(\mathcal{T} - \mathcal{L})(w^*) + \delta^* \in \Pi$.
(P2) $\|(\mathcal{T} - \mathcal{L})(w_1) - (\mathcal{T} - \mathcal{L})(w_2)\| \le \varepsilon \rho(w_1, w_2)$ *for all* $w_1, w_2 \in \mathbf{B}_r(w^*)$.
(P3) *The map $(\mathcal{F} - \mathcal{L})^{-1}$ is single-valued and Lipschitz continuous in Π with Lipschitz constant λ.*
(P4) $\mathcal{T} - \mathcal{L}$ *maps* $\mathbf{B}_r(w^*)$ *into* Π.

If $\varepsilon\lambda < 1$ and $r \ge \lambda\|\delta^*\|/(1 - \varepsilon\lambda)$, then there exists a unique $w \in \mathbf{B}_r(w^*)$ such that $\mathcal{T}(w) \in \mathcal{F}(w)$. Moreover, we have the estimate

$$(20) \qquad\qquad\qquad \rho(w, w^*) \le \frac{\lambda}{1 - \lambda\varepsilon}\|\delta^*\|.$$

Proof. Let us define $\Phi(w) = (\mathcal{F} - \mathcal{L})^{-1}(\mathcal{T}(w) - \mathcal{L}(w))$. By (P2), (P3), and (P4), we have for all $w_1, w_2 \in \mathbf{B}_r(w^*)$,

$$
\begin{aligned}
\rho(\Phi(w_1), \Phi(w_2)) &= \rho((\mathcal{F} - \mathcal{L})^{-1}(\mathcal{T} - \mathcal{L})(w_1), (\mathcal{F} - \mathcal{L})^{-1}(\mathcal{T} - \mathcal{L})(w_2)) \\
&\le \lambda\|(\mathcal{T} - \mathcal{L})(w_1) - (\mathcal{T} - \mathcal{L})(w_2)\| \\
&\le \lambda\varepsilon\rho(w_1, w_2).
\end{aligned}
$$

Since $\lambda\varepsilon < 1$, Φ is a contraction on $\mathbf{B}_r(w^*)$. Utilizing the fact that $(\mathcal{T} - \mathcal{L})(w^*) + \delta^*$ is contained in Π by (P1) and that $(\mathcal{F} - \mathcal{L})^{-1}$ is single-valued on Π by (P3), we have

$$w^* = (\mathcal{F} - \mathcal{L})^{-1}[(\mathcal{T} - \mathcal{L})(w^*) + \delta^*].$$

It follows from (P2) and (P3) that

$$
\begin{aligned}
\rho(\Phi(w), w^*) &= \rho((\mathcal{F} - \mathcal{L})^{-1}[(\mathcal{T} - \mathcal{L})(w)], (\mathcal{F} - \mathcal{L})^{-1}[(\mathcal{T} - \mathcal{L})(w^*) + \delta^*]), \\
&\le \lambda\|(\mathcal{T} - \mathcal{L})(w) - (\mathcal{T} - \mathcal{L})(w^*) + \delta^*\| \\
(21) \qquad &\le \lambda(\varepsilon\rho(w, w^*) + \|\delta^*\|) \\
&\le \lambda(\varepsilon r + \|\delta^*\|)
\end{aligned}
$$

for all $w \in \mathbf{B}_r(w^*)$. The condition $\lambda\|\delta^*\|/(1 - \varepsilon\lambda) \le r$ implies that $\lambda(\varepsilon r + \|\delta^*\|) \le r$, and hence, $\rho(\Phi(w), w^*) \le r$. Since Φ maps $\mathbf{B}_r(w^*)$ into itself and Φ is a contraction on $\mathbf{B}_r(w^*)$, it follows from the contraction mapping principle that there is a unique

fixed point $w \in \mathbf{B}_r(w^*)$. Since $\rho(\Phi(w), w^*) = \rho(w, w^*)$ for this fixed point, (21) gives (20). □

We apply Theorem 3.1 to the first-order conditions (13)–(16). We show that when h is sufficiently small, the assumptions of the theorem are satisfied with constants independent of h. In applying Theorem 3.1, we need to work in spaces of (discrete) Lipschitz continuous functions. For the space whose elements are sequences of the form $z_0, z_1, \ldots, z_N$, the i-th element being a vector $z_i \in \mathbf{R}^n$, we write $z \in \mathrm{Lip}_\xi$ if $\|z'\|_{L^\infty} \le \xi$. Similarly, if z'' denotes the centered second-order divided difference sequence defined by

$$z_i'' = \frac{z_{i+1} - 2z_i + z_{i-1}}{h^2},$$

then $z \in \mathrm{Lip}_\xi^1$ if $\|z''\|_{L^\infty} \le \xi$. For the control problem, the space $\mathcal{X}$ of Theorem 3.1 consists of 4-tuples whose components are vector sequences,

$$w = (x, \psi, u, \nu),$$

where

(22) $x, \psi \in \mathrm{Lip}_\xi^1$ (with the H^1 metric), $x_0 = x^0,$ $\psi_{N-1} = 0,$

(23) $u, \nu \in \mathrm{Lip}_\xi$ (with the L^2 metric), $\nu' \ge 0.$

An appropriate value for ξ is chosen later in Lemma 8.1. Since $\mathcal{X}$ depends on the choice of ξ, we often write $\mathcal{X}_\xi$ to denote this dependence.

The mappings $\mathcal{T}$ and $\mathcal{F}$ of Theorem 3.1 are selected in the following way:

$$\mathcal{T}(w)_i = \begin{pmatrix} x_i' - f(x_i, u_i) \\ \psi_{i-}' + \nabla_x \mathcal{H}(x_i, u_i, \psi_i) - \mathcal{P}_i \\ \nabla_u \mathcal{H}(x_i, \psi_i, u_i) - \nu_{i+1} \nabla g(x_{i+1}) \nabla_u f(x_i, u_i) \\ g(x_i) \end{pmatrix},$$

$$\mathcal{F}(w)_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathcal{N}_{\mathbf{R}_+^k}(\nu_i') \end{pmatrix}.$$

The space $\mathcal{Y}$, associated with the four components of $\mathcal{T}$, is a space of 4-tuples of finite sequences equipped with the norm of $(L^2)^3 \times H^1$. The reference point w^* of Theorem 3.1 is the sequence with elements

$$w_i^* = (x_i^*, u_i^*, \psi_i^*, \nu_i^*),$$

where $x_i^* = x^*(t_i)$, $u_i^* = u^*(t_i)$, $\psi_i^* = \psi^*(t_i)$, and $\nu_i^* = \nu^*(t_i)$. The operator $\mathcal{L}$ of Theorem 3.1 is the derivative of $\mathcal{T}$ evaluated at w^*: $\mathcal{L} = \nabla \mathcal{T}(w^*)$. The residual is defined by $\delta^* = -\mathcal{T}(w^*) + (0, 0, 0, \Delta)^\mathsf{T}$ where Δ is defined in the following way:

$$(\Delta_i)_j = \begin{cases} g_j(x_i^*) \text{ if } g_j(x^*(t)) < 0 \text{ for all } t \in (t_i, t_{i+1}), \\ 0 \text{ otherwise,} \end{cases}$$

$j = 1, 2, \cdots, k$. Finally, we define $\pi^* = \mathcal{T}(w^*) - \mathcal{L}(w^*)$, and we let a^*, s^*, r^*, and b^* denote the four components of π^*, corresponding to the four components of $\mathcal{T}$ and $\mathcal{L}$. The set Π is the set of sequences $\pi = (a, s, r, b) \in \mathcal{Y}$ for which

(24) $\pi \in \mathbf{B}_\sigma(\pi^*),$ $a - a^*, r - r^*, s - s^* \in \mathrm{Lip}_\kappa,$ $b - b^* \in \mathrm{Lip}_\kappa^1,$

where σ is a small positive constant, chosen later in Lemmas 7.3 and 7.4, and κ is a positive constant (not necessarily small) chosen in Lemma 5.1.

4. Approximation preliminaries

To prove Theorem 2.1, we will match the parameters ξ, κ, and σ in such a way that the assumptions of Theorem 3.1 hold for h sufficiently small. The estimate of Theorem 2.1 is a consequence of (20). In verifying the assumptions of Theorem 3.1, we utilize various approximation properties for piecewise polynomial spaces, and various embeddings, and we engage in an interplay between discrete sequences and continuous functions. In this section, we pull together results that are exploited throughout the analysis.

First, we recall standard properties of piecewise polynomial interpolants (see [4], [6], or [37]). In particular, given a sufficiently smooth function χ, if χ^I denotes the linear function with the property that $\chi^I(t_j) = \chi(t_j)$ for $j = i$ and $j = i + 1$, then the error in linear interpolation satisfies

(25)
$$\|\chi^I - \chi\|_{W^{l,\infty}[t_i,t_{i+1}]} \leq ch^{k-l}\|\chi^{(k)}\|_{L^\infty[t_i,t_{i+1}]} \quad \text{for all } 0 \leq l \leq k \leq 2.$$

When a time interval appears in any norm, the domain is restricted to the given interval. So if χ is essentially bounded and z is a discrete sequence, then

$$\|\chi\|_{L^\infty[r,s]} = \text{essential} \sup_{r \leq t \leq s} |\chi(t)| \quad \text{and} \quad \|z\|_{L^\infty[r,s]} = \max_{r \leq t_i \leq s} |z_i|.$$

Of course, (25) holds in other norms besides L^∞ norms; however, in our analysis, we will only use this property in the L^∞ norm. If χ^I is the quadratic function with the property that $\chi^I(t_j) = \chi(t_j)$ for $j = i - 1$, i, and $i + 1$, then the error in quadratic interpolation satisfies

(26)
$$\|\chi^I - \chi\|_{W^{l,\infty}[t_{i-1},t_{i+1}]} \leq ch^{k-l}\|\chi^{(k)}\|_{L^\infty[t_{i-1},t_{i+1}]} \quad \text{for all } 0 \leq l \leq k \leq 3.$$

Given a sequence $x_0, x_1, \ldots, x_N$, let y denote the associated continuous, piecewise linear interpolant that satisfies $y(t_i) = x_i$ for each i. On any grid interval, the absolute maximum of y is attained at either end of the interval. Hence, the continuous and the discrete L^∞ norms are equal:

(27)
$$\|y\|_{L^\infty} = \|x\|_{L^\infty}.$$

Since $y(\cdot)^2$ is a convex, nonnegative function on any mesh interval, and since the trapezoidal rule overestimates the integral of such a function, we have

(28)
$$\|y\|_{L^2}^2 \leq \sum_{i=0}^{N-1} h(x_i^2 + x_{i+1}^2)/2 \leq \|x\|_{L^2}^2.$$

In other words, the discrete L^2 norm is an upper bound for the continuous L^2 norm of the associated interpolant. Since $\dot{y}(t) = x_i'$ for all $t \in [t_i, t_{i+1}]$, it follows that

(29)
$$\|\dot{y}\|_{L^2} = \|x'\|_{L^2} \quad \text{and} \quad \|\dot{y}\|_{L^\infty} = \|x'\|_{L^\infty}.$$

Consequently, by (27), we have

(30)
$$\|y\|_{W^{1,\infty}} = \|x\|_{W^{1,\infty}},$$

while (28) implies that

(31)
$$\|y\|_{H^1} \leq \|x\|_{H^1}.$$

Letting x^I denote the continuous, piecewise linear interpolant of the optimal state x^*, it follows from (25) and (30) that

(32)
$$\|y - x^*\|_{W^{1,\infty}} \le \|y - x^I\|_{W^{1,\infty}} + \|x^I - x^*\|_{W^{1,\infty}} \le \|x - x^*\|_{W^{1,\infty}} + ch.$$

Analogously, if u^* is the optimal control, u is a discrete sequence, and v is the continuous, piecewise linear interpolant that satisfies $v(t_i) = u_i$ for each i, then it follows from (25) and (27) that

(33)
$$\|v - u^*\|_{L^\infty} \le \|u - u^*\|_{L^\infty} + ch.$$

For a sequence $x_0, x_1, \ldots, x_N$, the quadratic interpolant q on $[t_{i-1}, t_{i+1}]$ with the property that $q(t_j) = x_j$ for $j = i - 1, i$, and $i + 1$, satisfies

(34)
$$\|q\|_{L^\infty[t_{i-1}, t_{i+1}]} \le \frac{5}{4} \|x\|_{L^\infty[t_{i-1}, t_{i+1}]}.$$

The derivative of this quadratic interpolant at the ends of the interval $[t_{i-1}, t_{i+1}]$ can be expressed

$$\dot{q}(t_{i-1}) = x'_{i-1} - \frac{h}{2}x''_i \quad \text{and} \quad \dot{q}(t_{i+1}) = x_i' + \frac{h}{2}x''_i.$$

Since $\dot{q}$ is linear, its maximum value on $[t_{i-1}, t_{i+1}]$ is attained at either $t = t_{i-1}$ or $t = t_{i+1}$, which implies that

$$\|\dot{q}\|_{L^\infty[t_{i-1}, t_{i+1}]} \le \|x'\|_{L^\infty[t_{i-1}, t_i]} + \frac{h}{2}|x''_i|.$$

Combining this with (34) gives

(35)
$$\|q\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]} \le \frac{5}{4} \|x\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]} + \frac{h}{2}|x''_i|.$$

For any continuous function $z \in H^1$, $\|z\|_{L^\infty} \le \sqrt{2}\|z\|_{H^1}$. Combining this with (27) and (31) gives

$$\|x\|_{L^\infty}^2 = \|y\|_{L^\infty}^2 \le 2\|y\|_{H^1}^2 \le 2\|x\|_{H^1}^2.$$

Hence, we have

(36)
$$\|x\|_{L^\infty} \le \sqrt{2}\|x\|_{H^1}.$$

In [19, Lemma 3.1] we proved the following reverse Hölder-type inequality: If $y \in W^{1,\infty}$ and $\|\dot{y}\|_{L^\infty} \le \xi$, then

(37)
$$\|y\|_{L^\infty} \le \max\{\sqrt{3}\|y\|_{L^2}, \sqrt[3]{3\xi}\|y\|_{L^2}^{2/3}\}.$$

If y denotes the continuous, piecewise linear interpolant associated with the sequence $x_0, x_1, \ldots, x_N$ in Lip_ξ, then $\|\dot{y}\|_{L^\infty} \le \xi$, and (37) is applicable. Combining this with (27) and (28) gives the discrete version of (37):

(38)
$$\|x\|_{L^\infty} \le \max\{\sqrt{3}\|x\|_{L^2}, \sqrt[3]{3\xi}\|x\|_{L^2}^{2/3}\}.$$

For an N-element sequence $u_0, u_1, \ldots, u_{N-1}$, we form the associated continuous, piecewise linear interpolant v on $[0, 1 - h]$ and apply (37) to obtain the following discrete analogue:

(39)
$$\|u\|_{L^\infty} \le \max\{\sqrt{3/(1-h)}\|u\|_{L^2}, \sqrt[3]{3\xi}\|u\|_{L^2}^{2/3}\}.$$

The inequalities (38) and (39) allow us to convert L^2 neighborhoods in Lip_ξ into L^∞ neighborhoods. By Smoothness, the optimal control u^* is Lipschitz continuous

with Lipschitz constant bounded by $\|\dot{u}^*\|_{L^\infty}$. By (39) it follows that for any $\gamma > 0$, there exists $r > 0$ such that

$$\text{(40)} \qquad \|u - u^*\|_{L^\infty} \leq \gamma$$

for all sequences $u \in \text{Lip}_\xi$ with $\|u - u^*\|_{L^2} \leq r$. Similarly, due to (36),

$$\text{(41)} \qquad \|x - x^*\|_{L^\infty} \leq \gamma$$

for all sequences x with $\|x - x^*\|_{H^1} \leq \gamma/\sqrt{2}$. Applying (38) to $(x - x^*)'$, we conclude that for any $\gamma > 0$, there exists $r > 0$ such that

$$\|(x - x^*)'\|_{L^\infty} \leq \gamma$$

for all sequences $x \in \text{Lip}_\xi^1$ with $\|x - x^*\|_{H^1} \leq r$. Combining this with (41), we see that for any $\gamma > 0$, there exists $r > 0$ such that

$$\text{(42)} \qquad \|x - x^*\|_{W^{1,\infty}} \leq \gamma$$

for all sequences $x \in \text{Lip}_\xi^1$ with $\|x - x^*\|_{H^1} \leq r$.

To convert from divided differences of discrete sequences to derivatives of smooth functions, we utilize two integral representations. The first is simply the fundamental theorem of calculus: If $\zeta \in W^{1,1}$ has the property that $\zeta(t_j) = z_j$ for $j = i$ and $j = i + 1$, then

$$\text{(43)} \qquad z_i' = \frac{1}{h} \int_{t_i}^{t_{i+1}} \dot{\zeta}(s) \, ds.$$

Our second formula

$$\text{(44)} \qquad z_{i+1} - 2z_i + z_{i-1} = \int_0^h \int_{t_{i-1}+s}^{t_{i+1}-s} \ddot{\zeta}(t) \, dt \, ds$$

relates the second-order divided difference to the second derivative of a function $\zeta \in W^{2,1}$ that satisfies $\zeta(t_j) = z_j$ for $j = i - 1$, i and $i + 1$.

Let x^I denote the quadratic interpolant of x^* on the interval $[t_{i-1}, t_{i+1}]$, and let q be a quadratic chosen so that $q(t_j) = x_j$ for $j = i - 1$, i, and $i + 1$. By (26), we have

$$\|q - x^*\|_{W^{1,\infty}} \leq \|q - x^I\|_{W^{1,\infty}} + \|x^I - x^*\|_{W^{1,\infty}} \leq \|q - x^I\|_{W^{1,\infty}} + ch\|\ddot{x}^*\|_{L^\infty}.$$

Combining this with (35) and with the representation (44), which connects the second-order divided difference to the second derivative of an interpolant, we obtain

(45)
$$\|q - x^*\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]} \leq \frac{5}{4} \|x - x^*\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]} + ch(|x_i''| + \|\ddot{x}^*\|_{L^\infty}).$$

Referring to (42), it follows that for any given $\gamma > 0$, there exists r and $\bar{h}$ such that

$$\text{(46)} \qquad \|q - x^*\|_{W^{1,\infty}} \leq \gamma$$

for all $h \leq \bar{h}$ and for all $x \in \text{Lip}_\xi^1$ with $\|x - x^*\|_{H^1} \leq r$.

5. Analysis of residual and deviation from linearity

In this section, we begin to show that Theorem 3.1 is applicable to the Euler discretization, with the identifications given in Section 3, by verifying assumptions (P1), (P2), and (P4). Our first step is to show that when κ is sufficiently large, (P1) is satisfied for h sufficiently small.

Lemma 5.1. *If Smoothness holds, then we have $\mathcal{T}(w^*) + \delta^* \in \mathcal{F}(w^*)$ and there exists constants c and κ, independent of h, such that*

$$\text{(47)} \qquad \|\delta^*\|_{\mathcal{Y}} \leq ch \quad and \quad \delta^* \in (Lip_\kappa)^3 \times Lip_\kappa^1.$$

Proof. The inclusion $\mathcal{T}(w^*) + \delta^* \in \mathcal{F}(w^*)$ is trivial for the first three components. The only case where the last component does not vanish is the case that $g_j(x^*(t)) < 0$ for all $t \in [t_i, t_{i+1}]$. However, by the complementary slackness condition (6), $\nu_j^*(t)$ is constant on $[t_i, t_{i+1}]$ in this case, so the inclusion $\mathcal{T}(w^*) + \delta^* \in \mathcal{F}(w^*)$ is valid for the fourth component too.

Now consider the estimate for the norm of the residual. Since $\delta^* \in \mathcal{Y} = (L^2)^3 \times H^1$, the norm in (47) is interpreted relative to L^2 for the first three components and H^1 for the last component. Since $f(x_i^*, u_i^*) = \dot{x}_i^*$ and $x^* \in W^{2,\infty}$ by Smoothness, the first component of δ^* satisfies

$$\text{(48)} \qquad |(x_i^*)' - f(x_i^*, u_i^*)| = |(x_i^*)' - \dot{x}_i^*| \leq ch, \ \ i = 0, 1, \ldots, N-1,$$

by (25). Since the L^2 norm is bounded by the L^∞ norm, (48) implies that the L^2 norm of the first component of δ^* satisfies the first inequality in (47). The forward difference of the first component is $(x_i^*)'' - (\dot{x}_i^*)'$. By (43) and (44), both $|(x_i^*)''|$ and $|(\dot{x}_i^*)'|$ are bounded by $\|x^*\|_{W^{2,\infty}}$. Hence, the first component of δ^* satisfies (47) when κ is sufficiently large.

After utilizing (3), the second component of δ^* is expressed

$$(\psi_{i-}^*)' - \dot{\psi}_i^* + \nabla_x(\nu_i^* \nabla g(x) f(x, u_i^*))|_{x=x_i^*} - \nu_{i+1} \nabla g(x_i^*)' - \nu_{i+1} K_{i+1} A_i,$$

where $A_i = A(t_i)$ and $K_i = K(t_i)$. It can be verified that

$$\nu^*(t) \nabla \dot{g}(x^*(t)) + \nu^*(t) \nabla g(x^*(t)) \nabla_x f(x^*(t), u^*(t))$$
$$= \nabla_x[\nu^*(t) \nabla g(x^*(t)) f(x^*(t), u^*(t))],$$

where $\nabla \dot{g}(x_i^*)$ stands for the time derivative of $\nabla g(x^*(t))$ evaluated at $t = t_i$. With this substitution, the second component of δ^* can be expressed

(49)

$$\left[(\psi_{i-}^*)' - \dot{\psi}_i^*\right] + \left[\nu_i^* \nabla \dot{g}(x_i^*) - \nu_{i+1}^* \nabla g(x_i^*)'\right] + \left[(\nu_i^* K_i - \nu_{i+1}^* K_{i+1}) A_i\right].$$

Due to the assumed smoothness, each of the terms in brackets is bounded by ch. Hence, the second component of δ^* satisfies the first inequality in (47). Moreover, when the difference operator $'$ is applied to (49), the resulting expression is bounded in terms of $\|\psi^*\|_{W^{2,\infty}}$, $\|x^*\|_{W^{2,\infty}}$, and $\|\nu^*\|_{W^{1,\infty}}$, so the second component of δ^* satisfies (47) for κ sufficiently large.

The treatment of the third component of δ^* is similar to that of the first and second components. The last component δ^* is only nonzero when there exists $s \in (t_i, t_{i+1})$ such that $g_j(x^*(s)) = 0$. Since $x^* \in W^{2,\infty}$ by Smoothness, the inequality $g_j(x^*(t)) \leq 0$ for all $t \in [0, 1]$ implies that $\frac{d}{dt} g_j(x^*(s)) = 0$. From a Taylor expansion around s, we conclude that

$$|g_j(x^*(t))| \leq c|s - t|^2 \leq ch^2$$

for all $t \in (t_i, t_{i+1})$. Hence, the L^∞ norm of the last component of δ^* is bounded by ch^2. It follows that the H^1 norm is bounded by ch, the first-order divided difference of the last component is bounded by ch, and the second-order divided difference of the last component is bounded by c. This completes the proof of (47). □

Next, we establish condition (P2) of Theorem 3.1.

Lemma 5.2. *If Smoothness holds, then for each ξ and $\varepsilon > 0$, there exist $\bar{h}$ and $r > 0$ such that if $\pi^i = (\mathcal{T} - \mathcal{L})(w^i)$, $i = 1, 2$, then*

$$\|\pi^1 - \pi^2\|_{\mathcal{Y}} \leq \varepsilon \|w^1 - w^2\|_{\mathcal{X}} \tag{50}$$

for all $w^1, w^2 \in \mathbf{B}_r(w^*)$ and $h \leq \bar{h}$.

(Recall that $\mathbf{B}_r(w^*)$ is a ball in the space $\mathcal{X}_\xi$.)

Proof. Suppose that ξ and $\varepsilon > 0$ are given, and let $w^1 = (x^1, \psi^1, u^1, \nu^1)$ and $w^2 = (x^2, \psi^2, u^2, \nu^2)$. By Smoothness, for any given $\eta > 0$, there exists r such that

$$|\nabla_x f(x_i, u_i) - A_i| + |\nabla_u f(x_i, u_i) - B_i| < \eta \tag{51}$$

whenever $|x_i - x_i^*| + |u_i - u_i^*| \leq r$, where $A_i = A(t_i)$ and $B_i = B(t_i)$, $i = 0, 1, \ldots, N - 1$. By Smoothness we have, for all $z_i^1, z_i^2 \in \Omega$,

$$f(z_i^1) - f(z_i^2) = \int_0^1 \nabla f(z_i^1 + t(z_i^2 - z_i^1)) dt (z_i^1 - z_i^2),$$

where z_i^1 and z_i^2 are identified with the pairs (x_i^1, u_i^1) and (x_i^2, u_i^2), respectively. Combining this with (51), we have

$$|f(z_i^1) - f(z_i^2) - A_i(x_i^1 - x_i^2) - B_i(u_i^1 - u_i^2)| \leq \eta |z_i^1 - z_i^2|$$

whenever $|x_i - x_i^*| + |u_i - u_i^*| \leq r$. It follows that the first component of $\pi^1 - \pi^2$ has the estimate

$$\begin{aligned}
\|f(x^1, u^1) - f(x^2, u^2) - A(x^1 - x^2) - B(u^1 - u^2)\|_{L^2} \\
\leq \eta(\|x^1 - x^2\|_{H^1} + \|u^1 - u^2\|_{L^2})
\end{aligned} \tag{52}$$

whenever $\|x - x^*\|_{L^\infty} + \|u - u^*\|_{L^\infty} \leq r$, where $f(x, u)$, Ax, and Bu denote the sequences whose i-th elements are $f(x_i, u_i)$, $A_i x_i$, and $B_i u_i$, respectively. Combining this with (40) and (42), we conclude that for r sufficiently small, (52) holds for all $u \in \text{Lip}_\xi$ and $x \in \text{Lip}_\xi^1$ with $\|x - x^*\|_{H^1} + \|u - u^*\|_{L^2} \leq r$.

Now consider the fourth component of $\pi^1 - \pi^2$, which can be expressed as $g(x^1) - g(x^2) - K(x^1 - x^2)$. The same approach used in the analysis of the first component of $\pi^1 - \pi^2$ implies that for any given η, there exists r such that

$$\|g(x^1) - g(x^2) - K(x^1 - x^2)\|_{L^2} \leq \eta \|x^1 - x^2\|_{H^1}$$

for all $x \in \text{Lip}_\xi^1$ with $\|x - x^*\|_{H^1} \leq r$. Since the fourth component of $\mathcal{Y}$ is equipped with the H^1 norm, we also need to consider $(g(x^1) - g(x^2) - K(x^1 - x^2))'$. Utilizing (43) gives

$$\begin{aligned}
(g(x_i^1) - g(x_i^2) &- K_i(x_i^1 - x_i^2))' \\
&= \int_{t_i}^{t_{i+1}} \frac{d}{dt} \left(g(y^1(t)) - g(y^2(t)) - K(t)(y^1(t) - y^2(t)) \right) dt,
\end{aligned}$$

where y^1 and y^2 are continuous, piecewise linear interpolants associated with x^1 and x^2, respectively. Defining

$$\bar{K}(t) = \int_0^1 \nabla g((1-s)y^1(t) + sy^2(t))ds,$$

a Taylor expansion yields

$$(g(x_i^1) - g(x_i^2) - K_i(x_i^1 - x_i^2))' = \frac{1}{h}\int_{t_i}^{t_{i+1}} \frac{d}{dt}\left((\bar{K}(t) - K(t))(y^1(t) - y^2(t))\right)ds.$$

Utilizing (32) and the bound $\|y^1 - y^2\|_{H^1} \leq \|x^1 - x^2\|_{H^1}$ given in (31), we obtain

$$\|(g(x^1) - g(x^2) - K(x^1 - x^2))'\|_{L^2}$$
$$\leq c\|\bar{K} - K\|_{W^{1,\infty}}\|y^1 - y^2\|_{H^1}$$
$$\leq c(\|y^1 - x^*\|_{W^{1,\infty}} + \|y^2 - x^*\|_{W^{1,\infty}})\|y^1 - y^2\|_{H^1}$$
$$\leq c(h + \|x^1 - x^*\|_{W^{1,\infty}} + \|x^2 - x^*\|_{W^{1,\infty}})\|x^1 - x^2\|_{H^1}.$$

Again by (40) and (42), there exists $\bar{h}$ and $r > 0$ such that

$$\|g(x^1) - g(x^2) - K(x^1 - x^2)\|_{H^1} \leq \eta\|x^1 - x^2\|_{H^1}$$

for all x^1 and $x^2 \in \text{Lip}_\xi^1$ with $\|x^1 - x^*\|_{H^1} + \|x^2 - x^*\|_{H^1} \leq r$ and for all $h \leq \bar{h}$. Since η was arbitrary in this analysis, it follows that for η sufficiently small, (50) holds for the first and last components of $\pi^1 - \pi^2$. The analysis of the second and third components of $\mathcal{T} - \mathcal{L}$ is similar to the analysis of the first and last components. That is, discrete sequences are converted to continuous functions using piecewise polynomial interpolation, Taylor expansions are performed, and the resulting expressions are analyzed using the finite element estimates of Section 4.[1]

$\square$

We now verify condition (P4) of Theorem 3.1. That is, we will show that for r sufficiently small, $(\mathcal{T} - \mathcal{L})\mathbf{B}_r(w^*) \subset \Pi$, where Π is the set of sequences $\pi = (a, s, r, b)$ satisfying (24) for some given σ and κ. By Lemma 5.2 with $w^2 = w^*$ and $\pi^1 = (\mathcal{T} - \mathcal{L})(w)$, we have

$$\|(\mathcal{T} - \mathcal{L})(w) - \pi^*\|_{\mathcal{Y}} \leq \varepsilon\|w - w^*\|_{\mathcal{X}}$$

for all $w \in \mathbf{B}_r(w^*)$, where $\pi^* = (\mathcal{T} - \mathcal{L})(w^*)$. It follows that $(\mathcal{T} - \mathcal{L})\mathbf{B}_r(w^*) \subset \mathbf{B}_\sigma(\pi^*)$ for r sufficiently small. To finish the verification of (P4), we must show that

(53) $(\mathcal{T} - \mathcal{L})\mathbf{B}_r(w^*) - \pi^* \subset (\text{Lip}_\kappa)^3 \times \text{Lip}_\kappa^1.$

Lemma 5.3. *If Smoothness holds, then for each choice of ξ, κ, $\sigma > 0$, there exist $\bar{h}$ and $r > 0$ such that (53) holds, uniformly in $h \leq \bar{h}$.*

Proof. Given $w = (x, \psi, u, \nu) \in \mathbf{B}_r(w^*)$, the first component of $\pi^* - (\mathcal{T} - \mathcal{L})(w)$ is

$$f(x_i, u_i) - f(x_i^*, u_i^*) - A_i(x_i - x_i^*) - B_i(u_i - u_i^*), \quad 0 \leq i \leq N - 1.$$

---

[1] For this lemma as well as Lemma 5.3 and Lemma 7.1, where part of the proof is omitted, we provide an appendix on our web site (http://www.math.ufl.edu/~hager) that fills in additional details.

Proceeding as in Lemma 5.2, introducing continuous, piecewise linear interpolants y and v of the sequences x and u, respectively, and applying (43) gives

$$(f(x_i, u_i) - f(x_i^*, u_i^*) - A_i(x_i - x_i^*) - B(u_i - u_i^*))'$$

$$= \frac{1}{h} \int_{t_i}^{t_{i+1}} \frac{d}{dt}(f(y, v) - f(x^*, u^*) - A(y - x^*) - B(v - u^*))dt$$

$$= \frac{1}{h} \int_{t_i}^{t_{i+1}} ((\nabla_x f(y, v) - A)\dot{y} + (\nabla_u f(y, v) - B)\dot{v} - \dot{A}(y - x^*) - \dot{B}(v - u^*))dt.$$

Given any $\eta > 0$, it follows from Smoothness, (32), (33), (40), and (41) that for $\bar{h}$ and $r > 0$ sufficiently small, we have

$$\|\nabla_x f(y, v) - A\|_{L^\infty} + \|\nabla_u f(y, v) - B\|_{L^\infty} \leq \eta,$$

for all $h \leq \bar{h}$, $u \in \text{Lip}_\xi$ and $x \in \text{Lip}_\xi^1$ with $\|u - u^*\|_{L^2}$ and $\|x - x^*\|_{H^1} \leq r$. Hence, for any given $\eta > 0$, there exists $\bar{h}$ and $r > 0$ such that

$$\sup_i |(f(x_i, u_i) - f(x_i^*, u_i^*) - A_i(x_i - x_i^*) - B(u_i - u_i^*))'| \leq \eta$$

for all $h \leq \bar{h}$ and $u \in \text{Lip}_\xi$ and $x \in \text{Lip}_\xi^1$ with $\|u - u^*\|_{L^2} + \|x - x^*\|_{H^1} \leq r$.

Now consider the last component of $(\mathcal{T} - \mathcal{L})(w) - \pi^*$. In this case, we need to analyze a second-order divided difference. We do this by applying (44) with the following identification: If $q(t)$ denotes the quadratic on $[t_{i-1}, t_{i+1}]$ for which $q(t_j) = x_j$ for $j = i - 1, i,$ and $i + 1$, then we set

$$\zeta(t) = g(q(t)) - g(x^*(t)) - K(t)(q(t) - x^*(t))$$

to obtain

$$(g(x_i) - g(x_i^*) - K_i(x_i - x_i^*))''$$

$$= \frac{1}{h^2} \int_0^h \int_{t_{i-1}+s}^{t_{i+1}-s} \frac{d^2}{dt^2} \left(g(q(t)) - g(x^*(t)) - K(t)(q(t) - x^*(t)) \right) dt\, ds.$$

Expanding further,

$$g(q(t)) - g(x^*(t)) = \bar{K}(t)(q(t) - x^*(t)), \quad \bar{K}(t) = \int_0^1 \nabla g((1 - \tau)q(t) + \tau x^*(t))\, d\tau,$$

and we have

$$(g(x_i) - g(x_i^*) - K_i(x_i - x_i^*))''$$

$$(54) \qquad = \frac{1}{h^2} \int_0^h \int_{t_{i-1}+s}^{t_{i+1}-s} \frac{d^2}{dt^2} \left((\bar{K}(t) - K(t))(q(t) - x^*(t)) \right)\, dt\, ds.$$

Notice that

$$|\ddot{q}(t) - \ddot{x}^*(t)| = |x_i'' - \ddot{x}^*(t)| \leq \xi + \|x^*\|_{W^{2,\infty}}$$

for all $t \in [t_{i-1}, t_{i+1}]$ and $x \in \text{Lip}_\xi^1$. In a similar manner, $|\ddot{\bar{K}}(t) - \ddot{K}(t)| \leq c$, where c depends on ξ, $\|x^*\|_{W^{2,\infty}}$, and the first three derivatives of g on Ω. Since

$$\|\bar{K} - K\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]} \leq c\|q - x^*\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]},$$

we conclude, using (45) and (54), that

$$|(g(x_i) - g(x_i^*) - K_i(x_i - x_i^*))''| \leq c\|q - x^*\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]}$$
$$\leq c\|x - x^*\|_{W^{1,\infty}[t_{i-1}, t_{i+1}]} + ch.$$

Consequently, for any $\eta > 0$, there exists $\bar{h}$ and r such that

$$|(g(x_i) - g(x_i^*) - K_i(x_i - x_i^*))''| \leq \eta$$

for all $x \in \mathrm{Lip}_\xi^1$ with $\|x - x^*\|_{H^1} \leq r$ and for all $h \leq \bar{h}$. Since η was arbitrary in this analysis, it follows that for η sufficiently small, (53) holds for the first and last components. The analysis of the second and third components of $\mathcal{T} - \mathcal{L}$ is similar to the analysis of the first and the last components (see the web site cited earlier). $\quad\square$

6. A LINEAR-QUADRATIC PROBLEM

At this point, we have shown that for suitably chosen constants, (P1), (P2), and (P4) hold for the control problem. Next, we will verify that the map $(\mathcal{F} - \mathcal{L})^{-1}$ is single-valued and Lipschitz continuous when restricted to Π. Our approach is roughly the following. We first relate the elements of $(\mathcal{F} - \mathcal{L})^{-1}\pi$ to the solution of a linear-quadratic control problem in which the parameter π appears in the constraints and in the cost function. We show that the linear-quadratic problem has a unique solution depending Lipschitz continuously on the parameter. From this its follows that $(\mathcal{F} - \mathcal{L})^{-1}$ is single-valued and Lipschitz continuous. In the final phase of the analysis, we prove that derivatives of the solution to the linear-quadratic problem can be bounded in terms of derivatives of the parameters. From this discrete regularity result, we deduce that $(\mathcal{F} - \mathcal{L})^{-1}\pi$ lies in $\mathcal{X}_\xi$, for an appropriate choice of ξ, when $\pi \in \Pi$. In carrying out this plan, we need to deal with the several technical issues. First, we need to show that the coercivity assumption posed for the continuous control problem implies that the (discrete) linear-quadratic problem satisfies an analogous coercivity condition. Second, we need to observe that the feasible set for the linear-quadratic problem is stable under perturbation.

To begin the analysis, we write down the linearized problem. Since $\mathcal{T}(w)_i$ involves both x_i and x_{i+1} for each i, $\mathcal{L}(w)_i$ involves both x_i and x_{i+1} for each i. For any $\pi = (a, s, r, b)$, when $w = (x, \psi, u, \nu)$ is an element of $(\mathcal{F} - \mathcal{L})^{-1}\pi$, we have $x_{i+1} = x_i + h(A_i x_i + B_i u_i - a_i)$. After using this relation to substitute for x_{i+1}, we find that $w = (x, \psi, u, \nu)$ is an element of $(\mathcal{F} - \mathcal{L})^{-1}\pi$ if and only if

(55) $$x_i' = A_i x_i + B_i u_i - a_i, \quad x_0 = x^0,$$

(56) $$\psi_{i-}' = -\psi_i A_i - x_i^\mathsf{T} Q_i - (M_i u_i)^\mathsf{T} + \nu_{i+1}\hat{K}_i - \bar{s}_i, \quad \psi_{N-1} = 0,$$

(57) $$u_i^\mathsf{T} R_i + x_i^\mathsf{T} M_i + \psi_i B_i - \nu_{i+1} K_{i+1} B_i + \bar{r}_i = 0,$$

(58) $$K_i x_i + b_i \in \mathcal{N}_{\mathbf{R}_+^k}(\nu_i'), \quad \nu_i' \geq 0,$$

where $\hat{K}_i = K_i' + K_{i+1}A_i$ and

$$
\begin{aligned}
Q_i &= \hat{Q}_i - \frac{1}{h}(I + hA_i^\mathsf{T})(G_{i+1}\nu_{i+1}^*)(I + hA_i) + \frac{1}{h}G_i\nu_{i+1}^*, \\
M_i &= \hat{M}_i - (I + hA_i^\mathsf{T})(G_{i+1}\nu_{i+1}^*)B_i, \\
R_i &= \hat{R}_i - hB_i^\mathsf{T}(G_{i+1}\nu_{i+1}^*)B_i, \\
\bar{s}_i &= s_i + a_i^\mathsf{T}(G_{i+1}\nu_{i+1}^*)(I + hA_i), \quad \bar{r}_i = r_i + ha_i^\mathsf{T}(G_{i+1}\nu_{i+1}^*)B_i.
\end{aligned}
$$

The matrices $\hat{Q}_i$, $\hat{M}_i$, and $\hat{R}_i$ are defined by

(59)

$$
\begin{bmatrix} \hat{Q}_i & \hat{M}_i \\ \hat{M}_i^\mathsf{T} & \hat{R}_i \end{bmatrix} = \nabla_{(x,u)}^2 \left(\mathcal{H}(x, u, \psi_i^*) - \nu_{i+1}^* K_{i+1} f(x, u) \right) \Big|_{(x,u)=(x_i^*, u_i^*)},
$$

and G_i is the Hessian $\nabla^2 g(x)$ evaluated at $x = x^*(t_i)$. Above, products of the form $G_i \nu_{i+1}^*$ are defined in the following way. If G_{ij} is $\nabla^2 g_j(x)$ evaluated at $x = x^*(t_i)$ and $\nu \in \mathbf{R}^k$, then

$$G_i \nu = \sum_{j=1}^{k} G_{ij} \nu_j.$$

The system (55)–(58) constitutes the first-order necessary optimality conditions for the following discrete-time linear-quadratic problem:

(60) \qquad \text{minimize} \quad $\mathcal{B}(x, u) + \langle \bar{s}, x \rangle + \langle \bar{r}, u \rangle$

\qquad\qquad \text{subject to} \ $L(x, u) + a = 0, \ x_0 = x^0, \ Kx + b \le 0,$

where

$$L(x, u)_i = x_i' - A_i x_i - B_i u_i,$$

$\langle \cdot, \cdot \rangle$ is the discrete L^2 inner product:

$$\langle x, y \rangle = h \sum_i x_i^\mathsf{T} y_i,$$

and the discrete quadratic cost function is given by

(61) \qquad\qquad $\mathcal{B}(x, u) = \dfrac{h}{2} \sum_{i=0}^{N-1} x_i^\mathsf{T} Q_i x_i + u_i^\mathsf{T} R_i u_i + 2 x_i^\mathsf{T} M_i u_i.$

In order to establish existence and uniqueness results for the solution of (60), we analyze the coercivity of the cost function in this section, and the stability of the feasible set and solution in the next section.

Proposition 6.1. *If Smoothness and Coercivity hold, then for any given $\bar{\alpha} < \alpha$, there exists $\bar{h} > 0$ such that for all $h \le \bar{h}$, we have*

(62) \qquad\qquad $\mathcal{B}(x, u) \ge \bar{\alpha} \|u\|_{L^2}^2 \quad \text{for all } (x, u) \in \mathcal{M},$

where

$$\mathcal{M} = \{(x, u) : L(x, u) = 0, \ x_0 = 0\}.$$

Proof. In [18, Lemma 11], we establish this result in the case that $R_i = R^*(t_i)$, $Q_i = Q^*(t_i)$, and $M_i = M^*(t_i)$. Consequently, if there exists $\bar{h} > 0$ such that

(63) \qquad\qquad $|Q_i - Q^*(t_i)| + |M_i - M^*(t_i)| + |R_i - R^*(t_i)| \le ch,$

for all i and $h \le \bar{h}$, then the proposition follows immediately (after taking into account the fact that $\|x\|_{H^1} \le c\|u\|_{L^2}$ for all $(x, u) \in \mathcal{M}$). The relation (63), for h sufficiently small, will be established for the Q matrix, while the analysis of R and M is similar.

When we compute the difference $Q_i - Q^*(t_i)$, some terms cancel leaving us with the following expression:

(64)

$$Q_i - Q^*(t_i) = \nabla_{xx} \left(\nu_i^* \nabla g(x) f(x, u_i^*) - \nu_{i+1}^* K_{i+1} f(x, u_i^*) \right)_{x = x_i^*}$$
$$- G_i' \nu_{i+1}^* - A_i^\mathsf{T} (G_{i+1} \nu_{i+1}^*) - (G_{i+1} \nu_{i+1}^*) A_i + O(h).$$

The very first term $\nabla_{xx}[\nu_i^*(\nabla g)f]$ in (64) comes from Q^*, while the subsequent terms come from Q_i. Obviously, the $O(h)$ term in (64) can be made arbitrarily small

by taking h sufficiently small. Utilizing index notation, where repeated indices are summed over, the (p, q)-element of $\nabla_{xx}[\nu(\nabla g)f]$ is

(65)

$$
\frac{\partial^2}{\partial x_p \partial x_q} \left(\frac{\partial g_l}{\partial x_j} f_j \nu_l \right)
$$
$$
= \nu_l \frac{\partial g_l}{\partial x_j} \frac{\partial^2 f_j}{\partial x_p \partial x_q} + \nu_l f_j \frac{\partial^3 g_l}{\partial x_p \partial x_q \partial x_j} + \nu_l \frac{\partial^2 g_l}{\partial x_j \partial x_q} \frac{\partial f_j}{\partial x_p} + \nu_l \frac{\partial^2 g_l}{\partial x_j \partial x_p} \frac{\partial f_j}{\partial x_q}.
$$

By Smoothness, when each of these terms is evaluated at $(x, u, \nu) = (x_i^*, u_i^*, \nu_i^*)$, it cancels to within $O(h)$ the corresponding subsequent term in (64). For example, with the first term, we have

$$
\left(\nu_l \frac{\partial g_l}{\partial x_j} \frac{\partial^2 f_j}{\partial x_p \partial x_q} \right)_{(x,u,\nu)=(x_i^*,u_i^*,\nu_i^*)} = (\nu_i^*)_l \frac{\partial g_l(x_i^*)}{\partial x_j} \frac{\partial^2 f_j(x_i^*, u_i^*)}{\partial x_p \partial x_q}
$$
$$
= (\nu_{i+1}^*)_l \frac{\partial g_l(x_{i+1}^*)}{\partial x_j} \frac{\partial^2 f_j(x_i^*, u_i^*)}{\partial x_p \partial x_q} + O(h)
$$
$$
= (\nu_{i+1}^* K_{i+1})_j \frac{\partial^2 f_j(x_i^*, u_i^*)}{\partial x_p \partial x_q} + O(h)
$$
$$
= \frac{\partial^2}{\partial x_p \partial x_q} \left. \left(\nu_{i+1}^* K_{i+1} f(x, u_i^*) \right) \right|_{x=x_i^*} + O(h),
$$

which cancels to within $O(h)$ the second term on the right of (64). This completes the proof. $\qquad \square$

Henceforth, we assume that h is chosen small enough that the discrete coercivity condition (62) holds for some $\bar{\alpha} > 0$. Combining Proposition 6.1 with [18, Lemma 4], we have Lipschitz continuity of the state and control with respect to the parameters r and s:

Corollary 6.2. *If Smoothness and Coercivity hold, then for h sufficiently small, the linear-quadratic problem (60) has a unique solution for each choice of a and b where it is feasible, and if (x^j, u^j) is the solution associated with the parameters a, b, r^j, and s^j, $j = 1, 2$, we have*

(66) $\|x^1 - x^2\|_{H^1} + \|u^1 - u^2\|_{L^2} \le c(\|r^1 - r^2\|_{L^2} + \|s^1 - s^2\|_{L^2})$,

where c is independent of a and b. Conversely, any solution of the first-order conditions (55)–(58) is a solution of the linear-quadratic problem (60).

Proof. By [18, Lemma 4],

$$
\bar{\alpha} \|u^1 - u^2\|_{L^2} \le \|\bar{r}^1 - \bar{r}^2\|_{L^2} + \|\bar{s}^1 - \bar{s}^2\|_{L^2} = \|r^1 - r^2\|_{L^2} + \|s^1 - s^2\|_{L^2}.
$$

Since $(x^1 - x^2)_0 = 0$, it follows that $\|x^1 - x^2\|_{H^1} \le c\|u^1 - u^2\|_{L^2}$, which establishes (66). Since the first-order conditions are sufficient for optimality when the cost function is convex, any solution of the the first-order conditions (55)–(58) is a solution of the linear-quadratic problem (60). This completes the proof. $\qquad \square$

7. Lipschitz continuity in $H^1 \times L^2$

We will analyze the effect of perturbations in a and b by making a change of variables that moves a and b from the constraints to the cost function. This translation is based on the following result, which is a discrete-time version of [19, Lemma 3.6].

Lemma 7.1. *Let $\mathcal{I}$ map $[0, 1]$ to the subsets of $\{1, 2, \ldots, k\}$ and suppose that $\mathcal{I}^{-1}(i)$ is closed for each integer $i \in [1, k]$. If Smoothness and Independence at $\mathcal{I}$ hold, then for h sufficiently small and for every choice of a and b, there exists x and u such that $L(x, u) + a = 0$, $x_0 = x^0$, and*

$$(67) \qquad\qquad (K_i x_i + b_i)_j = 0 \quad \text{for each } j \in \mathcal{I}(t_i),$$

$i = 1, 2, \ldots, N$. This (x, u) pair is an affine function of (a, b), and

$$(68) \qquad \|x^1 - x^2\|_{H^1} + \|u^1 - u^2\|_{L^2} \le c(\|a^1 - a^2\|_{L^2} + \|b^1 - b^2\|_{H^1})$$

where (x^j, u^j) is the pair associated with (a^j, b^j) for $j = 1, 2$.

Proof. In [19, Lemma 3.6] we prove the continuous version of this result where the parameters a and b lie in the continuous spaces L^2 and H^1 and the solutions x and u lie in the continuous spaces H^1 and L^2. The same proof works for the discrete result Lemma 7.1, but with obvious changes. For example, the variable $t \in [0, 1]$ in [19, Lemma 3.6] is replaced by the discrete variable t_i, while derivatives that appear in [19, Lemma 3.6] are replaced by divided differences. For completeness, the proof details are given in the appendix at the web site cited earlier. ∎

Lemma 7.2. *If Smoothness and Independence at $\mathcal{A}$ hold, then there exists $\eta > 0$ with the following property: For any $\varepsilon > 0$ there exists $\bar{h} > 0$ and $(\bar{x}, \bar{u}) \in \mathbf{B}_\varepsilon(x^*, u^*)$ such that*

$$(69) \qquad L(\bar{x}, \bar{u}) + a^* = 0, \quad \bar{x}_0 = x^0, \quad K\bar{x} + b^* \le -\eta\mathbf{1}$$

for all $h \le \bar{h}$, where $\mathbf{1}$ denotes the vector of appropriate dimension whose entries are all one.

Proof. In [18, Lemma 3] we show that there exist $\tau > 0$, $v \in L^\infty$, and $y \in W^{1, \infty}$ (here the spaces are the continuous spaces, not the discrete versions) such that

$$\dot{y} - Ay - Bv = 0, \quad y(0) = 0, \quad \text{and} \quad Ky + g(x^*) \le -\tau\mathbf{1}.$$

Since the infinitely differentiable functions are dense in L^2, there is no loss of generality in assuming that v is continuously differentiable. Let $\bar{y}$ and $\bar{v}$ denote the sequences defined by $\bar{y}_i = y(t_i)$ and $\bar{v}_i = v(t_i)$. Since v is continuously differentiable, a Taylor expansion yields

$$(70) \qquad\qquad L(\bar{y}, \bar{v})_i = O(h).$$

Finally, we set $\bar{u} = u^* + \gamma\bar{v}$, and we let $\bar{x}$ be the solution to

$$(71) \qquad\qquad L(\bar{x}, \bar{u}) + a^* = 0, \quad \bar{x}_0 = x^0.$$

We will show that for γ and h small enough, $(\bar{x}, \bar{u})$ lies in $\mathbf{B}_\varepsilon(x^*, u^*)$.

The leading component of the residual, denoted δ^{*1}, was chosen so that

$$L(x^*, u^*) + a^* + \delta^{*1} = 0, \quad x_0^* = x^0.$$

Subtracting this and γ times (70) from (71), and utilizing the relation $\|\delta^{*1}\|_{L^2} = O(h)$ from Lemma 5.1, we have

$$\|\bar{x} - x^* - \gamma\bar{y}\|_{L^\infty} \le \|\bar{x} - x^* - \gamma\bar{y}\|_{H^1} = O(h)(1 + \gamma).$$

Since $\|\bar{x} - x^*\|_{H^1} \leq O(h)(1 + \gamma) + \gamma\|\bar{y}\|_{H^1}$ and $\|\bar{u} - u^*\|_{L^2} \leq \gamma\|\bar{v}\|_{L^2}$, it follows that $(\bar{x}, \bar{u})$ lies in $\mathbf{B}_\varepsilon(x^*, u^*)$ when γ and h are small enough. Since $b^* = g(x^*) - Kx^*$, we have

$$
\begin{aligned}
K\bar{x} + b^* &= K(x^* + \gamma\bar{y}) + b^* + O(h)(1 + \gamma) \\
&= \gamma K\bar{y} + g(x^*) + O(h)(1 + \gamma) \\
&= \gamma K\bar{y} + \gamma g(x^*) + (1 - \gamma)g(x^*) + O(h)(1 + \gamma) \\
&\leq -\gamma\tau\mathbf{1} + O(h)(1 + \gamma),
\end{aligned}
$$
(72)

assuming $\gamma \leq 1$. Now decrease h further if necessary so that the $O(h)(1 + \gamma)$ term in (72) is smaller than $\gamma\tau/2$. Taking $\eta = \gamma\tau/2$, the proof is complete. $\qquad\square$

Utilizing Lemma 7.1, we partially establish (P3) by showing that for a neighborhood of π^*, the function $(\mathcal{F} - \mathcal{L})^{-1}$ is single-valued and Lipschitz continuous relative to the norms of $\mathcal{X}$ and $\mathcal{Y}$. In Corollary 6.2, we have already established this result for the state and the control and perturbations in r and s. Now we consider perturbations in a and b, and we analyze the stability of multipliers as well. Our first result focuses on solution stability.

Lemma 7.3. *If Smoothness, Independence at $\mathcal{A}$, and Coercivity hold, then there exist constants $\bar{h}$ and $\sigma > 0$ with the property that whenever*

$$
\|\pi - \pi^*\|_{\mathcal{Y}} \leq \sigma \quad and \quad h \leq \bar{h},
$$
(73)

the linear-quadratic problem (60) has a unique solution. If for $j = 1$, 2, (x^j, u^j) is the solution corresponding to the parameter $\pi = \pi^j$ that satisfies (73), then we have

$$
\|x^1 - x^2\|_{H^1} + \|u^1 - u^2\|_{L^2} \leq c\|\pi^1 - \pi^2\|_{\mathcal{Y}}.
$$
(74)

Proof. Let $\mathcal{A}_\varepsilon$ be the index set for the ε-active constraints

$$
\mathcal{A}_\varepsilon(t) = \{i : g_i(x^*(t)) \geq -\varepsilon\}.
$$

In [19, p. 711] we show that there exists $\varepsilon > 0$ such that Independence at $\mathcal{A}_\varepsilon$ holds. Let us consider the following linear-quadratic problem:

$$
\text{minimize} \quad \mathcal{B}(x, u) + \langle \bar{s}, x \rangle + \langle \bar{r}, u \rangle
$$
(75)
$$
\text{subject to} \quad L(x, u) + a = 0, \quad x_0 = x^0, \quad (Kx + b)_{\mathcal{A}_\varepsilon} \leq 0,
$$

where $(y_{\mathcal{A}_\varepsilon})_i = (y_i)_{\mathcal{A}_\varepsilon(t_i)}$. This quadratic program is gotten by imposing only those constraints, at each time t_i, that are associated with the index set $\mathcal{A}_\varepsilon(t_i)$.

We now make some important observations. First, for $\pi = \delta^* + \pi^*$, where δ^* is the residual, the optimal solution to (60) is simply $(x, u) = (x^*, u^*)$. This can be confirmed by checking that for this choice of the parameter π, $(x, u) = (x^*, u^*)$ satisfies the first-order necessary conditions for (60), which are sufficient for optimality when the discrete coercivity condition (62) holds. Second, there exists a $\gamma > 0$ with the property that if (x, u) is feasible in (75) and

$$
\|x - x^*\|_{L^\infty} < \gamma, \quad \|b - b^*\|_{L^\infty} < \gamma,
$$
(76)

then (x, u) is feasible in (60). In particular, if

$$
\gamma(\|K\|_{L^\infty} + 1) < \varepsilon,
$$

then the constraints corresponding to j in the complement of $\mathcal{A}_\varepsilon(t_i)$ are satisfied automatically since

$$
\begin{aligned}
(K_i x_i + b_i)_j &= (K_i x_i^* + b_i^*)_j + (K_i(x_i - x_i^*))_j + (b_i - b_i^*)_j \\
&\leq (K_i x_i^* + b_i^*)_j + \gamma(|K_i| + 1) \\
&= g_j(x_i^*) + \gamma(|K_i| + 1) < -\varepsilon + \gamma(|K_i| + 1) < 0,
\end{aligned}
$$

for all $j \in \mathcal{A}_\varepsilon(t_i)^c$. Therefore, if $(\hat{x}, \hat{u})$ is a local minimizer for (75) and $x = \hat{x}$ and b satisfy (76), then $(\hat{x}, \hat{u})$ is a local minimizer for (60). Finally, for $\pi = \delta^* + \pi^*$, (x^*, u^*) is the solution of (75) (as well of (60)). That is, if (x, u) is feasible in (75) and (76) holds, then (x, u) is feasible in (60). Since (x^*, u^*) is the unique minimizer of (60), it follows that (x^*, u^*) is a local minimizer in (75). Since a local minimizer is a global minimizer for a convex program, we conclude that (x^*, u^*) is the unique minimizer of (75) when $\pi = \delta^* + \pi^*$.

Since Independence at $\mathcal{A}_\varepsilon$ holds, it follows from Lemma 7.1 that for any given a and b, there exist associated $x(a, b)$ and $u(a, b)$, which are affine functions of a and b, such that $x = x(a, b)$ and $u = u(a, b)$ satisfy $L(x, u) + a = 0$, $x_0 = x^0$, and

$$
(K_i x_i + b_i)_j = 0 \quad \text{for each } j \in \mathcal{A}_\varepsilon(t_i).
$$

Substituting $x = y + x(a, b)$ and $u = v + u(a, b)$, we transform (75) to an equivalent problem of the form

$$
\begin{aligned}
(77) \qquad \text{minimize} \quad &\mathcal{B}(y, v) + \langle \tilde{s}, y \rangle + \langle \tilde{r}, v \rangle \\
\text{subject to} \quad &L(y, v) = 0, \ x_0 = 0, \ (Ky)_{\mathcal{A}_\varepsilon} \leq 0,
\end{aligned}
$$

where

$$
(78) \qquad
\begin{aligned}
\tilde{s}_i &= \bar{s}_i + (Q_i x_i(a, b) + M_i u_i(a, b))^\mathsf{T} \\
&= s_i + a_i^\mathsf{T}(G_{i+1} \nu_{i+1}^*)(I + hA_i) + (Q_i x_i(a, b) + M_i u_i(a, b))^\mathsf{T},
\end{aligned}
$$

$$
(79) \qquad
\begin{aligned}
\tilde{r}_i &= \bar{r}_i + (R_i u_i(a, b) + M_i x_i(a, b))^\mathsf{T} \\
&= r_i + ha_i^\mathsf{T}(G_{i+1} \nu_{i+1}^*)B_i + (R_i u_i(a, b) + M_i x_i(a, b))^\mathsf{T}.
\end{aligned}
$$

Observe that $y = 0$ and $v = 0$ are feasible in this quadratic program. Hence, by Coercivity there exists a unique minimizer for each choice of $\tilde{s}$ and $\tilde{r}$. As in Corollary 6.2, it follows from Coercivity and [18, Lemma 4] that the solution change $(\delta y, \delta v)$ associated with the parameter change $(\delta \tilde{r}, \delta \tilde{s})$ satisfies the following estimate:

$$
\|\delta y\|_{H^1} + \|\delta v\|_{L^2} \leq c(\|\delta \tilde{r}\|_{L^2} + \|\delta \tilde{s}\|_{L^2}).
$$

Taking into account (78) and (79) and the bound (68) of Lemma 7.1 gives

$$
\|\delta y\|_{H^1} + \|\delta v\|_{L^2} \leq c\|\delta \pi\|_\mathcal{Y}.
$$

Since the solution pairs $(\hat{y}, \hat{v})$ for (77) and $(\hat{x}, \hat{u})$ for (75) satisfy $\hat{x} = \hat{y} + x(a, b)$ and $\hat{u} = \hat{v} + u(a, b)$, we conclude that the solution change $(\delta x, \delta u)$ corresponding to the parameter change $\delta \pi$ in (75) satisfies an estimate of the same form:

$$
(80) \qquad \|\delta x\|_{H^1} + \|\delta u\|_{L^2} \leq c\|\delta \pi\|_\mathcal{Y}.
$$

We now show that this result on solution stability for (75) yields solution stability for (60) as well. Let us consider the parameters $\pi = \pi^* + \delta^*$ in (75), for which the

solution is (x^*, u^*), and $\pi = \pi^1$, for which the associated solution is denoted (x^1, u^1). By (80) we have

$$\|x^1 - x^*\|_{H^1} + \|u^1 - u^*\|_{L^2} \leq c(\|\pi^1 - \pi^*\|_{\mathcal{Y}} + \|\delta^*\|_{\mathcal{Y}}).$$

We observed earlier that any solution to (75) for which (76) holds is a solution of (60). By Lemma 5.1, $\|\delta^*\|_{\mathcal{Y}}$ tends to zero as h tends to zero. Hence, for h sufficiently small and π^1 sufficiently close to π^*, the associated solution (x^1, u^1) to (75) satisfies (76); as a result, (x^1, u^1) is also the solution to (60). This completes the proof. $\qquad\square$

Next, we consider the stability of the multipliers. We will show that the change $(\delta\psi, \delta\nu)$ in the multipliers corresponding to a change $\delta\pi$ in the parameters satisfies an estimate of the form

(81) $$\|\delta\psi\|_{H^1} + \|\delta\nu\|_{L^2} \leq c(\|\delta x\|_{L^2} + \|\delta u\|_{L^2} + \|\delta\pi\|_{\mathcal{Y}}),$$

where $(\delta x, \delta u)$ is the solution change. Applying Lemma 7.3, we obtain the following result:

Lemma 7.4. *If Smoothness, Independence at $\mathcal{A}$, and Coercivity hold, then there exist constants $\bar{h}$ and $\sigma > 0$ with the property that whenever π and h satisfy (73), the linear-quadratic problem (60) has a unique solution and unique associated Lagrange multipliers. If for $j = 1, 2$, (ψ^j, ν^j) are the multipliers corresponding to the parameter $\pi = \pi^j$ that satisfies (73), then we have*

(82) $$\|\psi^1 - \psi^2\|_{H^1} + \|\nu^1 - \nu^2\|_{L^2} \leq c\|\pi^1 - \pi^2\|_{\mathcal{Y}}.$$

Proof. In [19, p. 711] we show not only that there exists $\varepsilon > 0$ such that Independence at $\mathcal{A}_\varepsilon$ holds, but also that for some $\bar{\beta} > 0$, there exists subsets $J_1, J_2, \ldots, J_l$ of $\{1, 2, \ldots, k\}$, corresponding points $0 = \tau_1 < \tau_2 < \cdots < \tau_{l+1} = 1$, and a constant $0 < \eta < \min_q(\tau_{q+1} - \tau_q)$ such that whenever $t \in [\tau_q - \eta, \tau_{q+1} + \eta] \cap [0, 1]$ for some $1 \leq q \leq l$, we have $\mathcal{A}_\varepsilon(t) \subset J_q$ and

$$\left|\sum_{j \in J_q} v_j (K(t)B(t))_j\right| \geq \bar{\beta}|v_{J_q}|$$

for every choice of v. Since $K_i = K_{i+1} + O(h)$, let us choose $\bar{h} < \eta$ small enough that

(83) $$\left|\sum_{j \in J_q} v_j (K_{i+1}B_i)_j\right| \geq .5\bar{\beta}|v_{J_q}|$$

for each $t_i \in [\tau_q - \eta, \tau_{q+1} + \eta] \cap [0, 1]$ and $h \leq \bar{h}$, and for every choice of v. Choose σ and $\bar{h}$ smaller if necessary so that when π and h satisfy (73), the associated solution (x, u) of (60) has the property that

(84) $$(K_i x_i + b_i)_{J_q^c} < 0 \quad \text{for all} \quad t_i \in [\tau_q - \eta, \tau_{q+1} + \eta] \cap [0, 1].$$

By the complementary slackness condition, we know that the multipliers associated with inactive constraints must vanish. It follows that

$$(\mu_i)_{J_q^c} = 0 \quad \text{for all} \quad t_i \in [\tau_q - \eta, \tau_{q+1} + \eta] \cap [0, 1],$$

while $(\nu_i)_{J_q^c}$ is constant for these i. From (57), we have

(85) $$\delta u_i^\mathsf{T} R_i + \delta x_i^\mathsf{T} M_i + \delta\psi_i B_i - \delta\nu_{i+1} K_{i+1} B_i + \delta\bar{r}_i = 0.$$

In [22, Lemma 2] we showed that the smallest eigenvalue of $R^*(t)$ is bounded from below by the positive constant α appearing in Coercivity. Since $R_i - R^*(t_i) = O(h)$ (see the proof of Proposition 6.1), the smallest eigenvalue of R_i is bounded from below by say $\bar{\alpha} < \alpha$ for h sufficiently small, and $|R_i^{-1}| \leq 1/\bar{\alpha}$. Focusing on i for which $t_i \in [\tau_q - \eta, \tau_{q+1} + \eta]$, (85) implies that

$$(86)$$
$$(\delta\nu_{i+1})_{J_q}(K_{i+1}B_i)_{J_q}R_i^{-1}(K_{i+1}B_i)_{J_q}^{\mathsf{T}}$$
$$= (\delta u_i^{\mathsf{T}}R_i + \delta x_i^{\mathsf{T}}M_i + \delta\psi_i B_i - (\delta\nu_{i+1})_{J_q^c}(B_i K_{i+1})_{J_q^c} + \delta\bar{r}_i)R_i^{-1}(K_{i+1}B_i)_{J_q}^{\mathsf{T}}.$$

The coefficient matrix for $(\delta\nu_{i+1})_{J_q}$ is invertible by (83). From the adjoint equation (56), we obtain the relation

$$(87)$$
$$\delta\psi_{i-}' = -\delta\psi_i A_i - \delta x_i^{\mathsf{T}}Q_i - (M_i\delta u_i)^{\mathsf{T}} + \delta\nu_{i+1}\hat{K}_i - \delta\bar{s}_i = 0, \quad \delta\psi_{N-1} = 0.$$

When this is combined with (86) and when the bound of Lemma 7.3 for $(\delta x, \delta u)$ is utilized, we conclude that

$$(88)$$
$$\|\delta\psi\|_{H^1([\tau_q-\eta,\tau_{q+1}])} + \|\delta\nu\|_{L^2([\tau_q-\eta,\tau_{q+1}])}$$
$$\leq c(|\delta\psi(t_p)| + \|\delta\pi\|_{\mathcal{Y}} + \|(\delta\nu)_{J_q^c}\|_{L^2([\tau_q-\eta,\tau_{q+1}])}),$$

where p is the largest integer i such that $t_i \in [\tau_q, \tau_{q+1}]$. Since $\delta\psi_{N-1} = 0$ and $\nu_N = 0$, it follows that for $q = l$, we have $p = N - 1$, $\delta\psi(t_p) = 0$, and $(\delta\nu_i)_{J_q^c} = 0$ when $\tau_l - \eta \leq t_i \leq 1$. In this case, (88) gives us the estimate

$$\|\delta\psi\|_{H^1([\tau_l-\eta,1])} + \|\delta\nu\|_{L^2([\tau_l-\eta,1])} \leq c\|\delta\pi\|_{\mathcal{Y}}.$$

Proceeding by induction (on q), suppose that

$$(89)\qquad \|\delta\psi\|_{H^1([\tau_{q+1}-\eta,1])} + \|\delta\nu\|_{L^2([\tau_{q+1}-\eta,1])} \leq c\|\delta\pi\|_{\mathcal{Y}}.$$

We just established this in the case $q = l - 1$. Since the L^∞ norm is bounded by the H^1 norm, we have

$$|\delta\psi(t_p)| \leq \|\delta\psi\|_{H^1([\tau_{q+1}-\eta,1])} \leq c\|\delta\pi\|_{\mathcal{Y}},$$

where p is the largest integer i such that $t_i \in [\tau_q, \tau_{q+1}]$. Also, by (89) we have the estimate

$$(90)\qquad \|(\delta\nu)_{J_q^c}\|_{L^2([\tau_{q+1}-\eta,\tau_{q+1}])} \leq c\|\delta\pi\|_{\mathcal{Y}}.$$

Since $(\delta\nu_i)_{J_q^c}$ is constant for $t_i \in [\tau_q - \eta, \tau_{q+1}]$, the L^2 norm of $(\delta\nu_i)_{J_q^c}$ over the interval $[\tau_q - \eta, \tau_{q+1}]$ is a multiple of the L^2 norm over the interval $[\tau_{q+1} - \eta, \tau_{q+1}]$. Hence, (90) implies that

$$\|(\delta\nu)_{J_q^c}\|_{L^2([\tau_q-\eta,\tau_{q+1}])} \leq c\|\delta\pi\|_{\mathcal{Y}}.$$

And when this, together with the bound $|\delta\psi(t_p)| \leq c\|\delta\pi\|_{\mathcal{Y}}$ is inserted in (88), the induction step is complete. $\qquad\square$

Corollary 7.5. *If Smoothness, Independence at $\mathcal{A}$, and Coercivity hold, then there exist constants $\bar{h}$ and $\sigma > 0$ with the property that whenever π and h satisfy (73), a unique solution to the first-order system (55)–(58) exists and the change in the solution corresponding to a change in the parameters satisfies (74) and (82).*

Proof. In the convex setting, the first-order system (55)–(58) is necessary and sufficient for optimality in the linear-quadratic problem (60). Hence, Lemmas 7.3 and 7.4 yield the claimed result. $\square$

8. Lipschitz continuity in discrete time

To complete the verification of condition (P3), we now prove a regularity result for the solution to the discrete linear-quadratic problem (60), establishing bounds for discrete derivatives of the solution in terms of discrete derivatives of the parameters.

Lemma 8.1. *If Smoothness, Independence at $\mathcal{A}$, and Coercivity hold, then for any given $\kappa > 0$, there exists ξ with the property that for all σ and h sufficiently small, and for all $\pi \in \Pi$, defined in (24), the linear-quadratic problem (60), and the associated first-order optimality system (55)–(58), have a unique solution (x, u) and associated multipliers (ψ, ν) with x and $\psi \in \mathrm{Lip}_\xi^1$ and u and $\nu \in \mathrm{Lip}_\xi$.*

Proof. Throughout this proof, (x, u) denotes the solution to (60) corresponding to $\pi \in \Pi$, while ψ and ν are the associated multipliers. As at the start of the proof of Lemma 7.4, we assume that σ and $\bar{h}$ are small enough that (83) holds for $h \leq \bar{h}$ and (84) holds for the solution to (60) when π and h satisfy (73). Finally, we assume that $\bar{h}$ is small enough that $\|\delta^*\|_\mathcal{Y} \leq \sigma$ when $h \leq \bar{h}$. As noted in the proof of Lemma 7.3, (x^*, u^*) is the solution and ψ^* and ν^* are the associated multipliers in (60) when $\pi = \pi^* + \delta^*$. By the Lipschitz estimates (74) and (82) with $\pi^2 = \delta^* + \pi^*$, we have

$$\|x - x^*\|_{H^1} + \|u - u^*\|_{L^2} \leq c(\|\pi - \pi^*\|_\mathcal{Y} + \|\delta^*\|_\mathcal{Y})$$

and

$$\|\psi - \psi^*\|_{H^1} + \|\nu - \nu^*\|_{L^2} \leq c(\|\pi - \pi^*\|_\mathcal{Y} + \|\delta^*\|_\mathcal{Y})$$

for all $\pi \in \Pi$. Hence, x and ψ are uniformly bounded in H^1 and u and ν are uniformly bounded in L^2 for all $\pi \in \Pi$. Since the L^∞ norm is bounded by the H^1 norm, we conclude that x and ψ are uniformly bounded in L^∞ as well.

Next, we establish a uniform L^∞ bound for ν. Taking a fixed $\varepsilon > 0$, there exist $\eta > 0$ by Lemma 7.2 and $(\bar{x}, \bar{u})$ in $\mathbf{B}_\varepsilon(x^*, u^*)$ such that

$$L(\bar{x}, \bar{u}) + a^* = 0, \quad \bar{x}_0 = x^0, \quad K\bar{x} + b^* \leq -\eta \mathbf{1}.$$

For each a, let $\bar{x}(a)$ be chosen to satisfy

$$L(\bar{x}(a), \bar{u}) + a = 0 \quad \text{and} \quad \bar{x}(a)_0 = x^0.$$

Taking η smaller if necessary, it follows that for a in an L^2 neighborhood of a^* and b in an H^1 neighborhood of b^*, we have

$$(91) \qquad L(\bar{x}(a), \bar{u}) + a = 0, \quad \bar{x}(a)_0 = x^0, \quad K\bar{x}(a) + b \leq -\eta \mathbf{1}.$$

Choose σ smaller if necessary so that (91) holds whenever $\|a - a^*\|_{L^2} + \|b - b^*\|_{H^1} \leq \sigma$.

Assuming h is chosen small enough to comply with Proposition 6.1, it follows from this discrete coercivity result that there exists a unique solution to the following problem:

$$(92) \qquad \begin{aligned} \text{minimize} \quad & \mathcal{B}(y, v) + \langle \bar{s}, y \rangle + \langle \bar{r}, v \rangle + \langle Ky + b, \mu \rangle \\ \text{subject to} \quad & L(y, v) + a = 0, \quad y_0 = x^0. \end{aligned}$$

If $\mu = \nu'$, where ν is the multiplier corresponding to the solution of (60), the solution to (92) coincides with the solution to (60) (since the first-order necessary conditions for (92) are satisfied by the solution to (60), and the first-order conditions are sufficient for optimality when the cost function is convex). For all $\pi \in \Pi$, we have observed already that the solution to (60) is uniformly bounded in L^2. Hence, the minimum cost in (60) is bounded from below by a constant χ independent of $\pi \in \Pi$. Since the optimal cost in (92) is equal to the optimal cost in (60), and since the cost in (92) corresponding to $(y, v) = (\bar{x}(a), \bar{u})$ cannot be smaller than the optimal cost, we have

$$\chi \le \mathcal{B}(\bar{x}(a), \bar{u}) + \langle \bar{s}, \bar{x}(a) \rangle + \langle \bar{r}, \bar{u} \rangle + \langle K\bar{x}(a) + b, \mu \rangle$$

for all choice of $\pi \in \Pi$. Rearranging this inequality and utilizing the relation $K\bar{x}(a) + b \le -\eta \mathbf{1}$ from (91), we have

$$(93) \qquad -\langle \nu_0, \mathbf{1} \rangle = h \sum_{i=0}^{N-1} \langle \mu_i, \mathbf{1} \rangle \le (\mathcal{B}(\bar{x}(a), \bar{u}) + \langle \bar{s}, \bar{x}(a) \rangle + \langle \bar{r}, \bar{u} \rangle - \chi)/\eta.$$

Since $(\bar{x}, \bar{u}) \in \mathbf{B}_\varepsilon(x^*, u^*)$ and $\|a - a^*\|_{L^2} \le \sigma$ when $\pi \in \Pi$, we conclude that the right side of (93) is bounded uniformly in $\pi \in \Pi$. Since ν_i is a nondecreasing function of i, we have $\nu_0 \le \nu_i \le \nu_N = 0$ for all $0 \le i \le N$. Thus ν is bounded in L^∞, uniformly in $\pi \in \Pi$.

We now show that this L^∞ bound for ν implies corresponding uniform L^∞ bounds for u, x', and ψ'. First, we noted in the proof of Lemma 7.4 that for $\bar{\alpha} < \alpha$ and for h sufficiently small, $|R_i^{-1}| \le 1/\bar{\alpha}$. Also, by (39) $\bar{r}$ is bounded in L^∞, uniformly in $\pi \in \Pi$. Hence, by (57) u is bounded, uniformly in $\pi \in \Pi$. By (55) x' is bounded, uniformly in $\pi \in \Pi$. And by (56) ψ' is bounded, uniformly in $\pi \in \Pi$.

First-order divided differences of u and ν and second-order divided differences of x and ψ are now estimated. In order to estimate $\nu_i' = \mu_i$, we focus on the active constraints at time level i since the components of μ_i associated with the inactive constraints vanish by complementary slackness. For a fixed i, let S_j and c_j, for $j = i-1$, i, and $i+1$, denote the submatrix of K_j and the subvector of b_j, respectively, associated with the active state constraints at time level i. Hence, we have

$$S_{i-1}x_{i-1} + c_{i-1} \le 0, \quad S_i x_i + c_i = 0, \quad \text{and} \quad S_{i+1}x_{i+1} + c_{i+1} \le 0.$$

Subtracting the equality $S_i x_i + c_i = 0$ from the inequality $S_{i+1}x_{i+1} + c_{i+1} \le 0$ gives

$$S_i' x_{i+1} + S_i x_i' \le -c_i'.$$

Substituting $x_i' = A_i x_i + B_i u_i - a_i$ from the state equation (55), we obtain

$$(94) \qquad S_i' x_{i+1} + S_i(A_i x_i + B_i u_i - a_i) \le -c_i'.$$

Solving for u_i in (57), we have

$$(95) \qquad u_i^\mathsf{T} = -(x_i^\mathsf{T} M_i + \psi_i B_i - \nu_{i+1} K_{i+1} B_i + \bar{r}_i) R_i^{-1},$$

and combining this with (94) gives

$$(96) \qquad \begin{aligned} S_i B_i R_i^{-1}(\nu_{i+1} K_{i+1} B_i)^\mathsf{T} &\le S_i B_i R_i^{-1}(x_i^\mathsf{T} M_i + \psi_i B_i + \bar{r}_i)^\mathsf{T} \\ &\quad - S_i(A_i x_i - a_i) - S_i' x_{i+1} - c_i'. \end{aligned}$$

In a similar manner, subtracting the equality $S_i x_i + c_i = 0$ from $S_{i-1}x_{i-1} + c_{i-1} \le 0$ yields

$$-S_{i-}' x_{i-1} - S_i x_{i-}' \le c_{i-}'.$$

Substituting $x'_{i-} = A_{i-1}x_{i-1} + B_{i-1}u_{i-1} - a_{i-1}$, we obtain

$$-S'_{i-}x_{i-1} - S_i(A_{i-1}x_{i-1} + B_{i-1}u_{i-1} - a_{i-1}) \leq c'_{i-}$$

and combining this with (95) gives

(97)
$$-S_iB_{i-1}R_{i-1}^{-1}(\nu_iK_iB_{i-1})^\mathsf{T} \leq -S_iB_{i-1}R_{i-1}^{-1}(x_{i-1}^\mathsf{T}M_{i-1} + \psi_{i-1}B_{i-1} + \bar{r}_{i-1})^\mathsf{T}$$
$$+ S_i(A_{i-1}x_{i-1} - a_{i-1}) + S'_{i-}x_{i-1} + c'_{i-}.$$

Adding (96) to (97), substituting $\nu_i = \nu_{i+1} - h\mu_i$, and rearranging the result, we obtain

(98)
$$S_iB_{i-1}R_{i-1}^{-1}(\mu_iK_iB_{i-1})^\mathsf{T} \leq S_i(B_iR_i^{-1}B_i^\mathsf{T}K_{i+1}^\mathsf{T})'_{-}\nu_{i+1}^\mathsf{T} - c''_i - S''_ix_i$$
$$- S'_ix'_i + S'_{i-}x'_{i-} + S_i\left(B_iR_i^{-1}(x_i^\mathsf{T}M_i + \psi_iB_i + \bar{r}_i)^\mathsf{T} - (A_ix_i - a_i)\right)'_{-}.$$

Let μ_i^+ denote the subvector of μ_i associated with the active state constraints at time level i. Since the other components of μ_i vanish by complementary slackness, it follows that $\mu_iK_i = \mu_i^+S_i$. Observe that

(99)
$$\mu_i^+S_iB_{i-1}R_{i-1}^{-1}(\mu_iK_iB_{i-1})^\mathsf{T} = \mu_i^+S_iB_{i-1}R_{i-1}^{-1}(\mu_i^+S_iB_{i-1})^\mathsf{T}$$
$$\geq \gamma|\mu_i^+S_iB_{i-1}|^2,$$

where γ is the smallest eigenvalue of R_{i-1}^{-1}, which can be bounded in the following way:

$$\gamma \geq \frac{1}{\max_i |R_i|}.$$

By (83) we have

$$|\mu_i^+S_{i+1}B_i| \geq \bar{\beta}|\mu_i^+| = \bar{\beta}|\mu_i|.$$

By Smoothness, $S_iB_{i-1} = S_{i+1}B_i + O(h)$. Hence, for h sufficiently small,

(100)
$$|\mu_i^+S_iB_{i-1}| \geq .5\bar{\beta}|\mu_i|.$$

Multiplying (98) by $\mu_i^{+\mathsf{T}}$ on the left and utilizing (99) and (100), we obtain

(101)
$$.25\gamma\bar{\beta}^2|\mu_i| \leq c\left(\|\nu\|_{L^\infty} + \|x\|_{W^{1,\infty}} + \|\psi\|_{W^{1,\infty}} + \|b''\|_{L^\infty} + \|\bar{r}'\|_{L^\infty} + \|a'\|_{L^\infty}\right).$$

Earlier in the proof, we established uniform bounds for $\|\nu\|_{L^\infty}$, $\|x\|_{W^{1,\infty}}$, and $\|\psi\|_{W^{1,\infty}}$. Bounds for b'', $\bar{r}'$, and a' can be expressed in terms of the parameter κ that appears in the definition (24) of Π. For example,

$$\|a'\|_{L^\infty} \leq \|(a - a^*)'\|_{L^\infty} + \|a^{*'}\|_{L^\infty} \leq \kappa + \|a^{*'}\|_{L^\infty},$$

where $\|a^{*'}\|_{L^\infty}$ is bounded, uniformly in h, by Smoothness. Since $\nu'_i = \mu_i$, (101) implies that $\|\nu'\|_{L^\infty}$ is bounded, uniformly in $\pi \in \Pi$. Finally, (57) implies that $\|u'\|_{L^\infty}$ is bounded, (55) implies that $\|x''\|_{L^\infty}$ is bounded, and (56) implies that $\|\psi''\|_{L^\infty}$ is bounded. Again, these bounds are uniform in $\pi \in \Pi$. This completes the proof. $\qquad\square$

9. Proof of Theorem 2.1

We now observe that Theorem 2.1 is a consequence of Theorem 3.1. First, note that many of the preceding lemmas contain the qualifier "for $\bar{h}$ sufficiently small." In the discussion that follows, we assume that $\bar{h}$ is chosen sufficiently small, and we do not always mention this point explicitly. Now, κ is chosen in accordance with (47) in Lemma 5.1, and σ is chosen in accordance with Lemmas 7.3 and 7.4. Choose σ smaller if necessary to comply with Lemma 8.1. For the parameter ξ, we take the value given in Lemma 8.1. (P1) follows from Lemma 5.1. With regard to (P3), Lemmas 7.3 and 7.4 imply that the map $(\mathcal{F} - \mathcal{L})^{-1}$, restricted to $\mathbf{B}_\sigma(\pi^*)$, is single-valued and Lipschitz continuous relative to the norms of $\mathcal{Y}$ and $\mathcal{X}$. When we restrict $(\mathcal{F} - \mathcal{L})^{-1}$ to Π, Lemma 8.1 implies that the regularity conditions (22) and (23) of $\mathcal{X}$ are satisfied. Let λ be the Lipschitz constant for $(\mathcal{F} - \mathcal{L})^{-1}$ when restricted to Π. Choose ε small enough that $\varepsilon\lambda < 1$. Choose r in accordance with Lemmas 5.2 and 5.3. (P2) is established in Lemma 5.2. (P4) is established in Lemma 5.3 and the comments preceding the lemma. Finally, choose $\bar{h}$ smaller if necessary so that $r \geq \lambda\|\delta^*\|_\mathcal{Y}/(1 - \varepsilon\lambda)$ whenever $h \leq \bar{h}$, where the residual δ^* satisfies (47). The estimate (20) of Theorem 3.1 coupled with the bound (47) yields the existence of a solution x^h, u^h, ψ^h, and ν^h to (13)–(16) that satisfies the estimate (18) of Theorem 2.1. The L^∞ estimate (19) is a consequence of the reverse Hölder inequalities (38) and (39). After transforming back to the original multipliers p and μ, we obtain the existence of a solution x^h, u^h, p^h, and μ^h to (8)–(11). To complete the proof, we need to show that the pair (x^h, u^h) is a local minimizer for the discrete problem (7).

Let $\mathcal{B}^h$ denote the usual quadratic form associated with the Hessian of the Lagrangian for (7):

$$(102) \qquad \mathcal{B}^h(x, u) = \frac{h}{2} \sum_{i=0}^{N-1} \left(z_i^\mathsf{T} \nabla_z^2 \mathcal{H}(z_i^h, p_i^h) z_i + x_i^\mathsf{T} [G_i^h \mu_i^h] x_i \right),$$

where z_i denotes the pair (x_i, u_i) and G^h is the Hessian of g is evaluated at x^h. Let Q_i^h, M_i^h, and R_i^h be the same as Q_i, M_i and R_i except that the functions are all evaluated at x_i^h, u_i^h, ψ_i^h, and ν_i^h instead of at x_i^*, u_i^*, ψ_i^*, and ν_i^*. And let $\mathcal{M}^h$ be defined by

$$\mathcal{M}^h = \{(x, u) : x_i' = A_i^h x_i + B_i^h u_i, \quad 0 \leq i \leq N - 1, \quad x_0 = 0\},$$

where $A_i^h = \nabla_x f(x_i^h, u_i^h)$ and $B_i^h = \nabla_u f(x_i^h, u_i^h)$. We now observe that when restricted to $\mathcal{M}^h$, $\mathcal{B}^h$ is equal to the following quadratic form:

$$\mathcal{B}^h(x, u) = \frac{h}{2} \sum_{i=0}^{N-1} x_i^\mathsf{T} Q_i^h x_i + u_i^\mathsf{T} R_i^h u_i + 2x_i^\mathsf{T} M_i^h u_i.$$

To see this, first substitute $\mu_i^h = (\nu_{i+1}^h - \nu_i^h)/h$ to obtain the relation

$$(103) \qquad h \sum_{i=0}^{N-1} x_i^\mathsf{T} [G_i^h \mu_i^h] x_i = \sum_{i=0}^{N-1} x_i^\mathsf{T} [G_i^h \nu_{i+1}^h] x_i - x_{i+1}^\mathsf{T} [G_{i+1}^h \nu_{i+1}^h] x_{i+1},$$

assuming $x_0 = 0$. If $(x, u) \in \mathcal{M}^h$, then

(104)
$$\sum_{i=0}^{N-1} x_{i+1}^{\mathsf{T}} [G_{i+1}^h \nu_{i+1}^h] x_{i+1}$$
$$= \sum_{i=0}^{N-1} \left((I + hA_i^h)x_i + hB_i^h u_i \right)^{\mathsf{T}} [G_{i+1}^h \nu_{i+1}^h] \left((I + hA_i^h)x_i + hB_i^h u_i \right).$$

Observe that the terms on the right side of (103) and (104) correspond to the far right terms in the definition of Q_i^h, M_i^h, and R_i^h. Similarly, after substituting $p_i^h = \psi_i^h - \nu_{i+1}^h \nabla g(x_{i+1}^h)$, the first term on the right side of (102) corresponds to the $\hat{Q}_i$, $\hat{M}_i$, and $\hat{R}_i$ matrices in the definition of Q_i^h, M_i^h, and R_i^h.

Utilizing the estimate (19) for the L^∞ distance from x^h, u^h, ψ^h, and ν^h to the continuous analogues x^*, u^*, ψ^*, and ν^*, respectively, it follows, by the same analysis used in the proof of Proposition 6.1, that

$$|Q_i^h - Q_i^*| + |R_i^h - R_i^*| + |M_i^h - M_i^*| = O(h^{2/3}),$$

where $Q_i^* = Q^*(t_i)$, $R_i^* = R^*(t_i)$, and $M_i^* = M^*(t_i)$. Let $\hat{\mathcal{B}}$ be the quadratic form associated with Q_i^*, M_i^*, and R_i^*:

$$\hat{\mathcal{B}}(x, u) = \frac{h}{2} \sum_{i=0}^{N-1} x_i^{\mathsf{T}} Q_i^* x_i + u_i^{\mathsf{T}} R_i^* u_i + 2x_i^{\mathsf{T}} M_i^* u_i.$$

Hence, for all $(x, u) \in \mathcal{M}^h$, we have

(105) $$|\hat{\mathcal{B}}(x, u) - \mathcal{B}^h(x, u)| \le ch^{2/3}(\|x\|_{L^2}^2 + \|u\|_{L^2}^2) \le ch^{2/3}\|u\|_{L^2}^2,$$

where the last inequality comes from the relation $\|x\|_{L^2} \le c\|u\|_{L^2}$ for all $(x, u) \in \mathcal{M}^h$. By [18, Lemma 11], there exists $\bar{\alpha} > 0$ such that

(106) $$\hat{\mathcal{B}}(x, u) \ge \bar{\alpha}\|u\|_{L^2}^2 \quad \text{for all} \ (x, u) \in \mathcal{M}.$$

Given $(x, u) \in \mathcal{M}^h$, we have

(107)
$$\mathcal{B}^h(x, u) = (\mathcal{B}^h(x, u) - \hat{\mathcal{B}}(x, u)) + (\hat{\mathcal{B}}(x, u) - \hat{\mathcal{B}}(y, u)) + \hat{\mathcal{B}}(y, u),$$

where y is the solution to $y_i' = A_i y_i + B_i u_i$, $0 \le i \le N - 1$, $y_0 = 0$. It can be checked that $\|x - y\|_{H^1} \le ch^{2/3}\|u\|_{L^2}$ and $\|y\|_{H^1} \le c\|u\|_{L^2}$, from which it follows that

$$|\hat{\mathcal{B}}(x, u) - \hat{\mathcal{B}}(y, u)| \le ch^{2/3}\|u\|_{L^2}^2$$

for h sufficiently small. Combining this with (105), (106), and (107) gives

$$\mathcal{B}^h(x, u) \ge (\bar{\alpha}/2)\|u\|_{L^2}^2$$

for all $(x, u) \in \mathcal{M}^h$, when h is sufficiently small. Hence, by the standard second-order sufficient optimality condition (see [1]), (x^h, u^h) is a local minimizer of (7). This completes the proof of Theorem 2.1.   □

10. Numerical experiment

For a small example, let us consider the following problem which is extracted from [30, Problem II]:

$$\text{minimize} \quad \frac{1}{2} \int_0^1 (x(t)^2 + u(t)^2) dt$$

$$\text{subject to} \quad \dot{x}(t) = u(t), \quad u(t) \leq 1,$$

$$x(t) \leq \frac{2\sqrt{e}}{1-e}, \quad x(0) = \frac{5e+3}{4(1-e)}.$$

This problem has a known solution (see [30]), while the L^∞ error for various choices of the mesh is given in Table 1. Based on these numerical results, it appears that the L^2 error estimate of Theorem 2.1 is tight, while the L^∞ estimate is not tight (at least in this example).

TABLE 1. L^∞ control error for various choices of the mesh

N	Error
20	.05379
40	.02697
80	.01351
160	.00676
320	.00338

References

[1] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1995.

[2] W. E. Bosarge, Jr. and O. G. Johnson, *Error bounds of high order accuracy for the state regulator problem via piecewise polynomial approximations,* SIAM J. Control, **9** (1971), pp. 15–28. MR **44:**6374

[3] W. E. Bosarge, Jr., O. G. Johnson, R. S. McKnight, and W. P. Timlake, *The Ritz-Galerkin procedure for nonlinear control problems*, SIAM J. Numer. Anal., **10** (1973), pp. 94–111. MR **47:**9827

[4] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 1994. MR **95f:**65001

[5] B. M. Budak, E. M. Berkovich and E. N. Solov'eva, *Difference approximations in optimal control problems*, SIAM J. Control, **7** (1969), pp. 18–31. MR **39:**4721

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978. MR **58:**25001

[7] J. Cullum, *Discrete approximations to continuous optimal control problems,* SIAM J. Control, **7** (1969), pp. 32–49. MR **42:**2341

[8] J. Cullum, *An explicit procedure for discretizing continuous, optimal control problems,* J. Optimization Theory Appl., **8** (1971), pp. 15–34. MR **46:**1029

[9] J. Cullum, *Finite-dimensional approximations of state-constrained continuous optimal control problems*, SIAM J. Control, **10** (1972), pp. 649–670. MR **56:**11284

[10] J. W. Daniel, *On the approximate minimization of functionals,* Math. Comp., **23** (1969), pp. 573–581. MR **40:**1007

[11] J. W. Daniel, *On the convergence of a numerical method in optimal control,* J. Optimization Theory Appl., **4** (1969), pp. 330–342. MR **40:**5120

[12] J. W. Daniel, *The Ritz-Galerkin method for abstract optimal control problems*, SIAM J. Control, **11** (1973), pp. 53–63. MR **48:**1003

[13] J. W. Daniel, *The Approximate Minimization of Functionals*, Wiley-Interscience, New York 1983.

[14] A. L. DONTCHEV, *Error estimates for a discrete approximation to constrained control problems*, SIAM J. Numer. Anal., **18** (1981), pp. 500–514. MR **83a:**49049

[15] A. L. DONTCHEV, *Perturbations, approximations and sensitivity analysis of optimal control systems*, Lecture Notes in Control and Inf. Sc., **52**, Springer, New York, 1983. MR **86m:**49003

[16] A. L. DONTCHEV, *Discrete approximations in optimal control*, in Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control (Minneapolis, MN, 1993), IMA Vol. Math. Appl., 78, Springer, New York, 1996, pp. 59–81. MR **97h:**49043

[17] A. L. DONTCHEV, *An a priori estimate for discrete approximations in nonlinear optimal control*, SIAM J. Control Optim., **34** (1996), pp. 1315–1328. MR **97d:**49034

[18] A. L. DONTCHEV AND W. W. HAGER, *Lipschitzian stability in nonlinear control and optimization*, SIAM J. Control Optim., **31** (1993), pp. 569–603. MR **94d:**49041

[19] A. L. DONTCHEV AND W. W. HAGER, *Lipschitzian stability for state constrained nonlinear optimal control*, SIAM J. Control Optim., **36** (1998), pp. 696–718. MR **99b:**49029

[20] A. L. DONTCHEV AND W. W. HAGER, *A new approach to Lipschitz continuity in state constrained optimal control*, Systems and Control Letters, **35** (1998), pp. 137–143.

[21] A. L. DONTCHEV, W. W. HAGER, AND V. M. VELIOV, *Second-order Runge-Kutta approximations in constrained optimal control*, Department of Mathematics, University of Florida, Gainesville, FL 32611, Dec 29, 1998 (http://www. math.ufl.edu/~hager/papers/rk2.ps).

[22] A. L. DONTCHEV, W. W. HAGER, A. B. POORE, B. YANG, *Optimality, stability and convergence in nonlinear control*, Appl. Math. Optim., **31** (1995), pp. 297–326. MR **95k:**49050

[23] J. C. DUNN, *On L^2 sufficient conditions and the gradient projection method for optimal control problems*, SIAM J. Control Optim., **34** (1996), pp. 1270–1290. MR **97d:**49033

[24] E. FARHI, *Runge-Kutta schemes applied to linear-quadratic optimal control problems*, in Mathematics and mathematical education (Sunny Beach 1984), Bulg. Acad. Sc., Sofia, 1984, pp. 464–472. MR **85e:**00015

[25] W. W. HAGER, *The Ritz-Trefftz method for state and control constrained optimal control problems*, SIAM J. Numer. Anal., **12** (1975), pp. 854–867. MR **54:**3550

[26] W. W. HAGER, *Rate of convergence for discrete approximations to unconstrained control problems*, SIAM J. Numer. Anal., **13** (1976), pp. 449–471. MR **58:**18058

[27] W. W. HAGER, *Convex control and dual approximations,* Control and Cybernetics, **8** (1979), pp. 1–22, 73–86. MR **83b:**49024a, MR **83b:**49024b

[28] W. W. HAGER, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim., **17** (1979), pp. 321–337. MR **80d:**49022

[29] W. W. HAGER, *Runge-Kutta methods in optimal control and the transformed adjoint system*, Department of Mathematics, University of Florida, Gainesville, FL 32611, January 4, 1999 (http://www.math.ufl.edu/~hager/papers/rk.ps).

[30] W. W. HAGER AND G. D. IANCULESCU, *Dual approximations in optimal control,* SIAM J. Control Optim., **22** (1984), pp. 423–465. MR **86c:**49019

[31] R. F. HARTL, S. P. SETHI, R. G. VICKSON, *A survey of the maximum principles for optimal control problems with state constraints*, SIAM Review, **37** (1995), pp. 181–218. MR **96j:**49019

[32] K. MALANOWSKI, C. BÜSKENS, AND H. MAURER, *Convergence of approximations to nonlinear optimal control problems*, in Mathematical Programming with Data Perturbations, Ed. A. V. Fiacco, Lecture Notes in Pure and Appl. Math, vol. 195, Marcel Dekker, New York, 1997, pp. 253–284. MR **98f:**49033

[33] B. MORDUKHOVICH, *On difference approximations of optimal control systems,* J. Appl. Math. Mech., **42** (1978), pp. 452–461. MR **84i:**49064

[34] E. POLAK, *A historical survey of computations methods in optimal control*, SIAM Review, **15** (1973), pp. 553–548. MR **53:**1956

[35] E. POLAK, *Optimization: Algorithms and Consistent Approximation*, Springer, New York, 1997. MR **98g:**49001

[36] A. L. SCHWARTZ AND E. POLAK, *Consistent approximations for optimal control problems based on Runge-Kutta integration,* SIAM J. Control Optim., **34** (1996), pp. 1235–1269. MR **97h:**49045

[37] G. STRANG AND G. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, NJ, 1973. Republished by Wellesley-Cambridge Press, Wellesley, MA, 1997. MR **56:**1747

[38] V. VELIOV, *On the time-discretization of control systems,* SIAM J. Control Optim., **35** (1997), pp. 1470–1486. MR **98f:**49034

[39] S. E. WRIGHT, *Consistency of primal-dual approximations for convex optimal control problems*, SIAM J. Control Optim., **33** (1995), pp. 1489–1509. MR **96h:**49057

[40] V. ZEIDAN, *Sufficient conditions for variational problems with variable endpoints: coupled points*, Appl. Math. Optim., **27** (1993), pp. 191–209. MR **94b:**49034

MATHEMATICAL REVIEWS, ANN ARBOR, MICHIGAN 48107
E-mail address: ald@ams.org

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA 32611
E-mail address: hager@math.ufl.edu
URL: http://www.math.ufl.edu/~hager