Modified Legendre-Gauss Collocation Method for Optimal Control Problems with Nonsmooth Solutions

Gabriela Abadia-Doyle[1] and Anil V. Rao[2]

***Abstract*— A modified form of Legendre-Gauss orthogonal direct collocation is developed for solving optimal control problems whose solutions are nonsmooth due to control discontinuities. This new method adds switch time variables, control variables, and collocation conditions at both endpoints of a mesh interval, whereas these new variables and collocation conditions are not included in standard Legendre-Gauss orthogonal collocation. The modified Legendre-Gauss collocation method alters the search space of the resulting nonlinear programming problem and optimizes the switch point of the control solution. The transformed adjoint system of the modified Legendre-Gauss collocation method is then derived and shown to satisfy the necessary conditions for optimality. Finally, an example is provided where the optimal control is bang-bang and contains multiple switches. This method is shown to be capable of solving complex optimal control problems with nonsmooth solutions.**

I. INTRODUCTION

Optimal control problems whose solutions are nonsmooth due to discontinuous control structures pose significant computational challenges. In particular, a priori knowledge of the precise discontinuity locations is seldom known, but detection and optimization of these locations are critical to obtaining high accuracy solutions. The objective of this paper is to develop a method that determines accurate solutions to optimal control problems with bang-bang control solutions.

Over the past few decades, direct collocation methods have become popular for solving general constrained optimal control problems numerically. More recently, the class of Gaussian quadrature direct orthogonal collocation methods has received a great deal of attention [1]–[5]. In a Gaussian quadrature orthogonal collocation method, the state is often approximated using a basis of Lagrange polynomials with Gaussian quadrature points as the support points for the Lagrange polynomials. The resulting finite-dimensional Gaussian quadrature collocation method then forms a nonlinear programming problem (NLP) that can be solved using well-known nonlinear optimization software. Well-developed Gaussian quadrature methods employ Legendre-Gauss (LG) points [1], Legendre-Gauss-Radau (LGR) points [2], [4], or Legendre-Gauss-Lobatto (LGL) points [5]. Additionally, convergence theory for Gaussian quadrature collocation methods that collocate the dynamics at LG or LGR points has demonstrated that, under assumptions of smoothness and coercivity, these methods converge to a local minimizer of the optimal control problem at an exponential rate as a function of the polynomial degree of the approximation [6].

When the solution of an optimal control problem is nonsmooth, both the standard Gaussian quadrature methods and the associated convergence theory are no longer applicable. For this reason, extensive research has been conducted into developing *hp* mesh refinement algorithms that adjust the number and width of mesh intervals and/or adjust the degree of the polynomial approximation within a mesh interval [7]–[9]. However, these methods tend to place an unnecessarily large number of collocation points and mesh intervals in the neighborhood of control discontinuities. Furthermore, some of the discretization schemes employed do not approximate both the left-hand and right-hand limits of a bang-bang optimal control at a single discrete switch time. Another approach is to introduce a variable mesh such that parameters corresponding to the switching structure are included as variables to be optimized. Assuming structure detection has been performed, the optimal control problem may be parameterized with switch time variables and additional optimal control law constraints as either multi-stage direct shooting methods [10], [11] or multi-phase direct collocation methods [12], [13]. An alternate approach for constructing a variable mesh is through nested direct transcription [14] which first solves an inner NLP on a static mesh and then solves an outer NLP to determine grid size and enforce additional optimality conditions. However when adding a degree of freedom through the inclusion of a switch time variable, the search space of the NLP solver is altered and the Lavrentiev phenomenon may occur. Such a phenomenon is observed where a numerical approximation of a continuous optimization problem leads to an optimal objective value that differs from the true optimal value [15]. A recently developed modified LGR collocation method [16] addresses the Lavrentiev gap by introducing defect constraints at one end of each mesh interval alongside variable switch times.

This paper presents a modified LG direct collocation method with the purpose of optimizing the locations of control discontinuities and obtaining accurate solutions to the state, control, and costate. This method extends the idea of Ref. [16] to LG collocation in order to take numerical advantage of the symmetry of the LG points as well as the increased accuracy of Gauss quadrature using LG points. It also addresses the drawback of multi-interval standard LG

*The authors gratefully acknowledge support for this research from the U.S. National Science Foundation under the Graduate Research Fellowship Program and grant CMMI-2031213, from the U.S. Office of Naval Research under grant N00014-22-1-2397, and from the U.S. Air Force Research Laboratory under contract FA8651-21-F-1041.

[1]Gabriela Abadia-Doyle is with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA gabadia97@ufl.edu

[2]Anil V. Rao is with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA anilvrao@ufl.edu

collocation related to absent discrete control values at adjacent interval interfaces. Unlike the modified LGR collocation method which introduces a new collocation constraint and control variable at just one endpoint of each mesh interval, the modified LG collocation method introduces additional collocation constraints and control variables at *both* the initial and terminal endpoint of each mesh interval. By comparing the Karush-Kuhn-Tucker (KKT) conditions from the NLP with the first-order variational conditions of the continuous optimal control problem, a costate mapping is obtained for this modified LG collocation scheme. Furthermore, the discrete and continuous adjoint systems of the modified LG collocation scheme are equivalent, unlike in the LGL collocation scheme (which inherently collocates both endpoints).

II. Bolza Optimal Control Problem

Without loss of generality, consider the following Bolza form of an optimal control problem. Determine the state, $\mathbf{x}(T) \in \mathbb{R}^{n_x}$ and $\mathbf{v}(T) \in \mathbb{R}^{n_v}$, the control $\mathbf{u}(T) \in \mathbb{R}^{n_u}$, the initial time, $t_0 \in \mathbb{R}$, and the final time, $t_f \in \mathbb{R}$, that minimize the objective functional

$$\mathcal{J} = \mathcal{M}(\mathbf{x}(-1), \mathbf{v}(-1), \mathbf{x}(+1), \mathbf{v}(+1), t_0, t_f) + \alpha \int_{-1}^{+1} \mathcal{L} \mathrm{d}T, \quad (1)$$

subject to the dynamic constraints

$$\mathbf{x}' = \alpha \mathbf{f}_x(\mathbf{x}, \mathbf{v}), \qquad \mathbf{v}' = \alpha \mathbf{f}_v(\mathbf{x}, \mathbf{v}, \mathbf{u}), \quad (2)$$

the boundary conditions

$$\mathbf{b}(\mathbf{x}(-1), \mathbf{v}(-1), \mathbf{x}(+1), \mathbf{v}(+1), t_0, t_f) = \mathbf{0}, \quad (3)$$

and the control inequality constraints

$$\mathbf{c}(\mathbf{u}) \leq \mathbf{0}, \quad (4)$$

where $\mathbf{z}'$ is the derivative of some vector-valued function $\mathbf{z}$ with respect to T, $\alpha := (t_f - t_0)/2$ is a domain scaling factor from $t \in [t_0, t_f]$ onto $T \in [-1, +1]$, and the functions $\mathcal{M}$, $\mathcal{L}$, $\mathbf{f}_x$, $\mathbf{f}_v$, $\mathbf{b}$, and $\mathbf{c}$ are defined by the mappings $\mathcal{M} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $\mathcal{L} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_u} \to \mathbb{R}$, $\mathbf{f}_x : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \to \mathbb{R}^{n_x}$, $\mathbf{f}_v : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_v}$, $\mathbf{b} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_v} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{n_b}$, and $\mathbf{c} : \mathbb{R}^{n_u} \to \mathbb{R}^{n_c}$. The modified Legendre-Gauss collocation method exploits the separation of those differential equations that explicitly depend on the control and those that do not. Furthermore, no generality is lost with such a decomposition since $n_x = 0$ is a special case of the dynamics in Eq. (2).

III. Legendre-Gauss Collocation

The Bolza optimal control problem of Section II may be partitioned into a mesh consisting of K mesh intervals, $\mathcal{I}_k = [T_{k-1}, T_k]$, $(k = 1, \ldots, K)$, where $-1 = T_0 < T_1 < \ldots < T_{K-1} < T_K = +1$. The mesh intervals have the property that $\bigcup_{k=1}^{K} \mathcal{I}_k = [-1, +1]$ and $\mathcal{I}_k \cap \mathcal{I}_{k+1} = \{T_k\}, (k = 1, \ldots, K-1)$. Each individual interval is mapped from the computational domain T to a new independent variable $\tau \in [-1, +1]$ via the affine transformation $\tau = 2(T - T_{k-1})/(T_k - T_{k-1}) - 1$, which implies that $\mathrm{d}T/\mathrm{d}\tau = (T_k - T_{k-1})/2 := \beta_k, (k = 1, \ldots, K)$, where β_k is defined as a secondary domain scaling factor.

The multiple interval Legendre-Gauss (LG) direct orthogonal collocation method for optimal control [1], [2] is based on approximating the state in each interval using Lagrange interpolating polynomials and enforcing the state dynamics at collocation points. For simplicity of notation, it is assumed that the number of collocation points, denoted by N, is the same in each mesh interval. Next, let $(\tau_1, \tau_2, \ldots, \tau_N)$ be the N LG nodes on the interval $(-1, +1)$ while $\tau_0 = -1$ and $\tau_{N+1} = +1$ are located at the endpoints of each interval. Now, let the state in each interval be approximated by a polynomial of degree at most N using a basis of Lagrange polynomials, $\ell_j(\tau)$, such that

$$\mathbf{x}^{(k)} \approx \sum_{j=0}^{N} \mathbf{X}_j^{(k)} \ell_j(\tau), \qquad \mathbf{v}^{(k)} \approx \sum_{j=0}^{N} \mathbf{V}_j^{(k)} \ell_j(\tau), \quad (5)$$

where $\ell_j(\tau)$ are the Lagrange polynomials

$$\ell_j(\tau) = \prod_{\substack{i=0 \\ i \neq j}}^{N} \frac{\tau - \tau_i}{\tau_j - \tau_i}, \quad (j = 0, \ldots, N), \quad (6)$$

whose support points are the initial endpoint, τ_0, and the N LG nodes, $(\tau_1, \ldots, \tau_N)$. The row-vectors $\mathbf{X}_i^{(k)} \in \mathbb{R}^{n_x}$ and $\mathbf{V}_i^{(k)} \in \mathbb{R}^{n_v}$ satisfy $\mathbf{x}^{(k)}(\tau_i) \approx \mathbf{X}_i^{(k)}$ and $\mathbf{v}^{(k)}(\tau_i) \approx \mathbf{V}_i^{(k)}$, $(i = 0, \ldots, N;\ k = 1, \ldots, K)$ due to the isolation property of the Lagrange polynomials

$$\ell_j(\tau_i) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (7)$$

Differentiating $\mathbf{x}^{(k)}$ and $\mathbf{v}^{(k)}$ leads to

$$\dot{\mathbf{x}}^{(k)} \approx \sum_{j=0}^{N} \mathbf{X}_j^{(k)} \dot{\ell}_j(\tau), \qquad \dot{\mathbf{v}}^{(k)} \approx \sum_{j=0}^{N} \mathbf{V}_j^{(k)} \dot{\ell}_j(\tau), \quad (8)$$

where $\dot{\mathbf{z}}$ is the derivative of some vector-valued function $\mathbf{z}$ with respect to τ. Next, let the state dynamics of Eq. (2) be discretized at the LG nodes and denoted by ${\mathbf{f}_x}_i^{(k)} := \mathbf{f}_x(\mathbf{X}_i^{(k)}, \mathbf{V}_i^{(k)})$ and ${\mathbf{f}_v}_i^{(k)} := \mathbf{f}_v(\mathbf{X}_i^{(k)}, \mathbf{V}_i^{(k)}, \mathbf{U}_i^{(k)})$, $(i = 1, \ldots, N;\ k = 1, \ldots, K)$, where the row vector $\mathbf{U}_i^{(k)} \in \mathbb{R}^{n_u}$, corresponds to the discrete control components at τ_i, $(i = 1, \ldots, N)$. The state derivative approximation of Eq. (8) is collocated with the right-hand side of the dynamic constraints at the N LG points of each mesh interval, producing the following defect constraints,

$$\sum_{j=0}^{N} \mathbf{D}_{(i,j)} \mathbf{X}_j^{(k)} = \alpha \beta_k {\mathbf{f}_x}_i^{(k)}, \quad \sum_{j=0}^{N} \mathbf{D}_{(i,j)} \mathbf{V}_j^{(k)} = \alpha \beta_k {\mathbf{f}_v}_i^{(k)}, \quad (9)$$

where $\mathbf{D}_{(i,j)} := \dot{\ell}_j(\tau_i)$, $(i = 1, \ldots, N;\ j = 0, \ldots, N)$, are the elements of the $N \times (N+1)$ *standard LG differentiation matrix*. It can be seen in Eq. (9) that the dynamic constraints are only collocated at the LG points and *not* at the boundary points. Since the Lagrange interpolating polynomials are used to approximate the state at the initial endpoint of an interval and the collocation points, the approximation of the state at the terminal endpoint of each mesh interval is obtained via the Gauss quadrature constraint,

$$\begin{aligned}\mathbf{X}_{N+1}^{(k)} &= \mathbf{X}_0^{(k)} + \alpha\beta_k \sum_{i=1}^{N} w_i \mathbf{f}_{x_i}^{(k)},\\ \mathbf{V}_{N+1}^{(k)} &= \mathbf{V}_0^{(k)} + \alpha\beta_k \sum_{i=1}^{N} w_i \mathbf{f}_{v_i}^{(k)},\end{aligned} \tag{10}$$

where w_i, $(i = 1, \ldots, N)$, are the Gauss quadrature weights.

The aforementioned discretization leads to the following nonlinear programming problem (NLP) that approximates the optimal control problem given in Section II. Minimize the objective function

$$\mathcal{J} = \mathcal{M}(\mathbf{X}_0^{(1)}, \mathbf{V}_0^{(1)}, \mathbf{X}_{N+1}^{(K)}, \mathbf{V}_{N+1}^{(K)}, t_0, t_f) + \alpha \sum_{k=1}^{K} \sum_{i=1}^{N} \beta_k w_i \mathcal{L}_i^{(k)}, \tag{11}$$

subject to

$$\mathbf{D}_{(i,:)}\mathbf{X}_{0:N}^{(k)} - \alpha\beta_k \mathbf{f}_{x_i}^{(k)} = \mathbf{0}, \tag{12}$$

$$\mathbf{D}_{(i,:)}\mathbf{V}_{0:N}^{(k)} - \alpha\beta_k \mathbf{f}_{v_i}^{(k)} = \mathbf{0}, \tag{13}$$

$$\mathbf{X}_{N+1}^{(k)} - \mathbf{X}_0^{(k)} - \alpha\beta_k \sum_{i=1}^{N} w_i \mathbf{f}_{x_i}^{(k)} = \mathbf{0}, \tag{14}$$

$$\mathbf{V}_{N+1}^{(k)} - \mathbf{V}_0^{(k)} - \alpha\beta_k \sum_{i=1}^{N} w_i \mathbf{f}_{v_i}^{(k)} = \mathbf{0}, \tag{15}$$

$$\mathbf{b}(\mathbf{X}_0^{(1)}, \mathbf{V}_0^{(1)}, \mathbf{X}_{N+1}^{(K)}, \mathbf{V}_{N+1}^{(K)}, t_0, t_f) = \mathbf{0}, \tag{16}$$

$$\mathbf{c}(\mathbf{U}_i^{(k)}) \leq \mathbf{0}, \tag{17}$$

where $\mathcal{L}_i^{(k)} := \mathcal{L}(\mathbf{X}_i^{(k)}, \mathbf{V}_i^{(k)}, \mathbf{U}_i^{(k)})$, $(i = 1, \ldots, N;\ k = 1, \ldots, K)$. Continuity in the state is enforced implicitly by using the same variable for the pair $\mathbf{X}_{N+1}^{(k)}$ and $\mathbf{X}_0^{(k+1)}$ and the pair $\mathbf{V}_{N+1}^{(k)}$ and $\mathbf{V}_0^{(k+1)}$ at each interior mesh point.

IV. MODIFIED LEGENDRE-GAUSS COLLOCATION

Additional variables and corresponding constraints are now augmented to the standard LG collocation method in order to improve the approximation of nonsmoothness in the solution to the optimal control problem. In particular, interior mesh points are treated as variables, control variables are introduced at the previously non-collocated endpoints of each mesh interval, and collocation constraints are added at both endpoints of each mesh interval.

A. New Decision Variables

The modified LG collocation method introduces new decision variables corresponding to the location of interior mesh points as well as new decision variables corresponding to the value of the control at the endpoints of each mesh interval. The interior mesh point variables are denoted T_k, $(k = 1, \ldots, K-1)$. The values of the control approximation at the start and end of each mesh interval are denoted, respectively, by $\mathbf{U}_0^{(k)}$ and $\mathbf{U}_{N+1}^{(k)}$, $(k = 1, \ldots, K)$. It is important to note that $\mathbf{U}_{N+1}^{(k)}$ and $\mathbf{U}_0^{(k+1)}$, $(k = 1, \ldots, K-1)$, correspond to the same mesh point T_k. Unlike the state approximation which implicitly maintains continuity at the mesh points by using the same variable for the pair $\mathbf{X}_{N+1}^{(k)}$ and $\mathbf{X}_0^{(k+1)}$ and the pair $\mathbf{V}_{N+1}^{(k)}$ and $\mathbf{V}_0^{(k+1)}$, $(k = 1, \ldots, K)$, the control approximation needs not be continuous, particularly in the case of nonsmoothness in the solution. Therefore, the dual values of the control at a mesh point T_k allow the left-hand and right-hand limits of the control at T_k be approximated, i.e., $\mathbf{u}^{(k)}(T_k^-) \approx \mathbf{U}_{N+1}^{(k)}$ and $\mathbf{u}^{(k+1)}(T_k^+) \approx \mathbf{U}_0^{(k+1)}$.

B. New Constraints

Additional constraints are now added to appropriately modify the search space such that the values of the new decision variables can be accurately approximated. These additional constraints consist of collocation constraints at the endpoints of each mesh interval, exclusively applied to those differential equations that are an explicit function of the control. It is important to note that the standard LG collocation method uses the initial endpoint and the LG nodes to formulate a basis of Lagrange polynomials for the purpose of approximating the state. Evaluating the derivative of this same basis of Lagrange polynomials at the endpoints of each interval results in a *modified LG differentiation matrix* of the form

$$\tilde{\mathbf{D}} = \begin{bmatrix} \left[\dot{\ell}_0(\tau_0), \ldots, \dot{\ell}_N(\tau_0)\right] \\ \mathbf{D} \\ \left[\dot{\ell}_0(\tau_{N+1}), \ldots, \dot{\ell}_N(\tau_{N+1})\right] \end{bmatrix} \in \mathbb{R}^{(N+2)\times(N+1)}, \tag{18}$$

where $\mathbf{D} \in \mathbb{R}^{N\times(N+1)}$ is the standard LG differentiation matrix. The resulting collocation constraints at the initial endpoint and terminal endpoint of each mesh interval are then given by

$$\begin{aligned}\tilde{\mathbf{D}}_{(0,:)}\mathbf{V}_{0:N}^{(k)} - \alpha\beta_k \mathbf{f}_{v_0}^{(k)} &= \mathbf{0},\\ \tilde{\mathbf{D}}_{(N+1,:)}\mathbf{V}_{0:N}^{(k)} - \alpha\beta_k \mathbf{f}_{v_{N+1}}^{(k)} &= \mathbf{0},\end{aligned} \tag{19}$$

where $\mathbf{f}_{v_0}^{(k)} := \mathbf{f}_v(\mathbf{X}_0^{(k)}, \mathbf{V}_0^{(k)}, \mathbf{U}_0^{(k)})$ and $\mathbf{f}_{v_{N+1}}^{(k)} := \mathbf{f}_v(\mathbf{X}_{N+1}^{(k)}, \mathbf{V}_{N+1}^{(k)}, \mathbf{U}_{N+1}^{(k)})$, $(k = 1, \ldots, K)$, and $\tilde{\mathbf{D}}_{(0,:)}$ and $\tilde{\mathbf{D}}_{(N+1,:)}$ correspond to the first row and last row of $\tilde{\mathbf{D}}$, respectively. Note, these new constraints only correspond to components of $\mathbf{v}$ since $\mathbf{f}_x(\mathbf{x}, \mathbf{v})$ is not an explicit function of control.

In addition to the endpoint collocation constraints given by Eq. (19), the control inequality constraints in Eq. (17) are augmented to include the new control variables using

$$\mathbf{c}(\mathbf{U}_i^{(k)}) \leq \mathbf{0}, \quad (i = 0, \ldots, N+1;\ k = 1, \ldots, K). \tag{20}$$

Lastly, the inclusion of variable mesh points requires that the mesh interval scaling factors be constrained by $\sum_{k=1}^{K} \beta_k = 1$, where $\beta_k > 0$, $(k = 1, \ldots, K)$, ensures that the timespan of each mesh interval is strictly increasing. These mesh interval scaling factors can also be thought of as fractions of the mesh, summing to unity. The standard Legendre-Gauss collocation method given by Eqs. (11)-(16) together with the collocation constraints in Eq. (19), the control path constraints in Eq. (20), and the scaling factor constraints is referred to as the *modified Legendre-Gauss collocation method*.

V. FIRST-ORDER NECESSARY CONDITIONS OF THE CONTINUOUS BOLZA PROBLEM

The transformed adjoint system of the modified LG collocation method can be derived by relating the Karush-Kuhn-Tucker (KKT) conditions of the NLP to the first-order optimality conditions of the continuous optimal control problem. These necessary conditions for optimality are derived using a variational approach which employs calculus of variations and Pontryagin's minimum principle [17] on the optimal control problem defined in Section II. To simplify the derivation, it is assumed that the control inequality path constraint of Eq. (4) can be omitted from the problem formulation because it can be enforced implicitly by the NLP variable bounds. The continuous augmented Hamiltonian is defined as $\mathcal{H}(\mathbf{x},\mathbf{v},\boldsymbol{\lambda}_x,\boldsymbol{\lambda}_v,\mathbf{u}) = \mathcal{L}(\mathbf{x},\mathbf{v},\mathbf{u}) + \boldsymbol{\lambda}_x\mathbf{f}_x^{\mathsf{T}}(\mathbf{x},\mathbf{v}) + \boldsymbol{\lambda}_v\mathbf{f}_v^{\mathsf{T}}(\mathbf{x},\mathbf{v},\mathbf{u})$, where $\boldsymbol{\lambda}_x \in \mathbb{R}^{n_x}$ and $\boldsymbol{\lambda}_v \in \mathbb{R}^{n_v}$ are the costates associated with $\mathbf{x}$ and $\mathbf{v}$, respectively. The continuous first-order optimality conditions are given by

$$(\mathbf{x}',\mathbf{v}') = \alpha(\mathbf{f}_x,\mathbf{f}_v), \tag{21}$$

$$(\boldsymbol{\lambda}_x',\boldsymbol{\lambda}_v') = -\alpha(\nabla_x(\mathcal{H}),\nabla_v(\mathcal{H})), \tag{22}$$

$$\mathbf{0} = \alpha\nabla_u(\mathcal{H}), \tag{23}$$

$$(\boldsymbol{\lambda}_x(-1),\boldsymbol{\lambda}_v(-1)) = (\nabla_{x_0}(\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}}-\mathcal{M}),\nabla_{v_0}(\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}}-\mathcal{M})), \tag{24}$$

$$(\boldsymbol{\lambda}_x(+1),\boldsymbol{\lambda}_v(+1)) = (\nabla_{x_f}(\mathcal{M}-\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}}),\nabla_{v_f}(\mathcal{M}-\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}})), \tag{25}$$

$$(\mathcal{H}(t_0),\mathcal{H}(t_f)) = (\nabla_{t_0}(\mathcal{M}-\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}}),\nabla_{t_f}(\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}}-\mathcal{M})), \tag{26}$$

where $\boldsymbol{\psi} \in \mathbb{R}^{n_b}$ is the Lagrange multiplier associated with the boundary condition $\mathbf{b}$. Furthermore, it has been shown in [18] that the augmented Hamiltonian at the initial and final times can be written, respectively, as

$$\mathcal{H}(t_0) = -\alpha\int_{-1}^{1}\frac{\partial\mathcal{H}}{\partial t_0}\mathrm{d}t + \frac{1}{2}\int_{-1}^{1}\mathcal{H}\mathrm{d}t, \tag{27}$$

$$\mathcal{H}(t_f) = \alpha\int_{-1}^{1}\frac{\partial\mathcal{H}}{\partial t_f}\mathrm{d}t + \frac{1}{2}\int_{-1}^{1}\mathcal{H}\mathrm{d}t. \tag{28}$$

VI. KKT CONDITIONS OF THE NLP

The KKT conditions of the NLP associated with the modified LG collocation method are obtained by setting equal to zero the derivatives of the augmented cost function, or Lagrangian, with respect to each variable. The Lagrangian associated with modified LG collocation is given as

$$\begin{aligned}
\mathcal{J}_a = \mathcal{J} &- \sum_{k=1}^{K}\sum_{i=1}^{N}\left\langle {\boldsymbol{\Lambda}_x}_i^{(k)}, \mathbf{D}_{(i,:)}\mathbf{X}_{0:N}^{(k)} - \alpha\beta_k{\mathbf{f}_x}_i^{(k)}\right\rangle \\
&- \sum_{k=1}^{K}\sum_{i=0}^{N+1}\left\langle {\tilde{\boldsymbol{\Lambda}}_v}_i^{(k)}, \tilde{\mathbf{D}}_{(i,:)}\mathbf{V}_{0:N}^{(k)} - \alpha\beta_k{\mathbf{f}_v}_i^{(k)}\right\rangle \\
&- \sum_{k=1}^{K}\left\langle {\boldsymbol{\Lambda}_x}_{N+1}^{(k)}, \mathbf{X}_{N+1}^{(k)} - \mathbf{X}_0^{(k)} - \alpha\beta_k\sum_{i=1}^{N}w_i{\mathbf{f}_x}_i^{(k)}\right\rangle \\
&- \sum_{k=1}^{K}\left\langle {\boldsymbol{\Lambda}_v}_{N+1}^{(k)}, \mathbf{V}_{N+1}^{(k)} - \mathbf{V}_0^{(k)} - \alpha\beta_k\sum_{i=1}^{N}w_i{\mathbf{f}_v}_i^{(k)}\right\rangle \\
&- \boldsymbol{\Psi}\mathbf{b}^{\mathsf{T}}(\mathbf{X}_0^{(1)},\mathbf{V}_0^{(1)},\mathbf{X}_{N+1}^{(K)},\mathbf{V}_{N+1}^{(K)},t_0,t_f) - \Theta\Big(\sum_{k=1}^{K}\beta_k - 1\Big),
\end{aligned} \tag{29}$$

where $\boldsymbol{\Lambda}_x^{(k)} \in \mathbb{R}^{(N+1)\times n_x}$, $\boldsymbol{\Lambda}_v^{(k)} \in \mathbb{R}^{(N+1)\times n_v}$, ${\tilde{\boldsymbol{\Lambda}}_v}_0^{(k)} \in \mathbb{R}^{n_v}$, ${\tilde{\boldsymbol{\Lambda}}_v}_{N+1}^{(k)} \in \mathbb{R}^{n_v}$, $\boldsymbol{\Psi} \in \mathbb{R}^{n_b}$, and $\Theta \in \mathbb{R}$ are the Lagrange multipliers, $\tilde{\boldsymbol{\Lambda}}_v^{(k)} := [{\tilde{\boldsymbol{\Lambda}}_v}_0^{(k)}, {\boldsymbol{\Lambda}_v}_{1:N}^{(k)}, {\tilde{\boldsymbol{\Lambda}}_v}_{N+1}^{(k)}]^{\mathsf{T}} \in \mathbb{R}^{(N+2)\times n_v}$, and $\langle\cdot,\cdot\rangle$ denotes the standard inner product between two vectors. Furthermore, ${\boldsymbol{\Lambda}_x}_i^{(k)}$ and ${\boldsymbol{\Lambda}_v}_i^{(k)}$ denote the ith rows of $\boldsymbol{\Lambda}_x^{(k)}$ and $\boldsymbol{\Lambda}_v^{(k)}$, respectively.

Next, the following theorem is introduced that will allow the terms involving ${\mathbf{f}_v}_0^{(k)}$ and $\tilde{\mathbf{D}}_{(0,1:N)}$ in Eq. (29) to be written as functions of $\mathbf{X}_{1:N}^{(k)}$, $\mathbf{V}_{1:N}^{(k)}$, $\mathbf{U}_{1:N}^{(k)}$, and $\mathbf{D}$. For the remainder of this discussion, let $\mathbf{W} = \mathrm{diag}(w_1,\ldots,w_N)$ be a diagonal matrix of LG quadrature weights

Theorem 1: Let $(\tau_1,\ldots,\tau_N)$ be the Legendre-Gauss points on the interval $(-1,+1)$ and let $\tau_0 = -1$ and $\tau_{N+1} = +1$. Furthermore, let $L_j(\tau)$ be a Lagrange basis polynomial, given by

$$L_j(\tau) = \prod_{\substack{i=0\\ i\neq j}}^{N+1}\frac{\tau-\tau_i}{\tau_j-\tau_i}, \quad (j=0,\ldots,N+1), \tag{30}$$

with support points at $(\tau_0,\tau_1,\ldots,\tau_{N+1})$. Then, if $f(\tau)$ is a polynomial of degree at most $N-1$ on the interval $\tau \in [-1,+1]$, it is the case that

$$\int_{-1}^{+1} f(\tau)\dot{L}_0(\tau)\mathrm{d}\tau = -f(-1). \tag{31}$$

Proof: The left-hand side of Eq. (31) can be integrated by parts as

$$\int_{-1}^{+1} f(\tau)\dot{L}_0(\tau)\mathrm{d}\tau = f(\tau)L_0(\tau)\Big|_{-1}^{+1} - \int_{-1}^{+1}\dot{f}(\tau)L_0(\tau)\mathrm{d}\tau. \tag{32}$$

Because $f(\tau)$ is a polynomial of degree at most $N-1$, it follows that $\dot{f}(\tau)$ is a polynomial of degree at most $N-2$. Furthermore, because $L_0(\tau)$ is a polynomial of at most degree $N+1$, then the integrand on the right-hand side of Eq. (32) is at most degree $2N-1$. Since LG quadrature is exact for polynomials of degree $2N-1$ or less, the integral on the right-hand side of Eq. (32) can be evaluated exactly using LG quadrature as

$$\int_{-1}^{+1}\dot{f}(\tau)L_0(\tau)\mathrm{d}\tau = \sum_{i=1}^{N} w_i\dot{f}(\tau_i)L_0(\tau_i), \tag{33}$$

where w_i is the ith LG quadrature weight. Then, since the Lagrange polynomials given by Eq. (30) satisfy the isolation property $L_j(\tau_i) = \delta_{ij}$ (see Eq. (7)), every term $L_0(\tau_i), (i = 1,\ldots,N+1)$, is zero which implies that

$$\int_{-1}^{+1} f(\tau)\dot{L}_0(\tau)\mathrm{d}\tau = f(\tau)L_0(\tau)\Big|_{-1}^{+1} = -f(-1). \tag{34}$$

□

Corollary 1: The row vector $\tilde{\mathbf{D}}_{(0,1:N)}$ obtained from the modified LG differentiation matrix can be written as $-\boldsymbol{\Gamma}_0\mathbf{W}\mathbf{D}_{(:,1:N)}$, where $\boldsymbol{\Gamma}_0 := [\dot{L}_0(\tau_1),\dot{L}_0(\tau_2),\ldots,\dot{L}_0(\tau_N)]$.

Proof: Replacing $f(\tau)$ from Theorem 1 with $\dot{\ell}_j(\tau)$, $(j=1,\ldots,N)$, results in

$$\int_{-1}^{+1}\dot{\ell}_j(\tau)\dot{L}_0(\tau)\mathrm{d}\tau = -\dot{\ell}_j(-1) = -\tilde{\mathbf{D}}_{(0,j)}. \tag{35}$$

Furthermore, since $\dot{\ell}_j(\tau)\dot{L}_0(\tau)$ is a polynomial of degree at most $2N-1$, the left-hand side of Eq. (35) can be replaced exactly with an LG quadrature as

$$\int_{-1}^{+1} \dot{\ell}_j(\tau)\dot{L}_0(\tau)\mathrm{d}\tau = \sum_{i=1}^{N} w_i \dot{\ell}_j(\tau_i)\dot{L}_0(\tau_i). \quad (36)$$

Relating Eq. (35) and Eq. (36), $\tilde{\mathbf{D}}_{(0,1:N)}$ can be written as

$$\tilde{\mathbf{D}}_{(0,1:N)} = -\mathbf{\Gamma}_0 \mathbf{W}\mathbf{D}_{(:,1:N)}. \quad (37)$$

□

Corollary 2: Suppose that $(\mathbf{X}_i^{(k)}, \mathbf{V}_i^{(k)}, \mathbf{U}_i^{(k)})$, $(i = 0,\dots,N+1)$ satisfy the collocation constraints given in Eq. (9) and Eq. (19). Following the definitions in Corollary 1, the row vector $\mathbf{f}_{v0}^{(k)}$ can be written as $-\mathbf{\Gamma}_0\mathbf{W}\mathbf{f}_{v1:N}^{(k)}$, where $\mathbf{f}_{v1:N}^{(k)} := \left[\mathbf{f}_{v1}^{(k)}, \mathbf{f}_{v2}^{(k)}, \dots, \mathbf{f}_{vN}^{(k)}\right]^\mathsf{T} \in \mathbb{R}^{N\times n_v}$.

Proof: Replacing $f(\tau)$ from Theorem 1 with the vector function $\mathbf{F}(\tau) = \sum_{j=0}^{N} \dot{\ell}_j(\tau)\mathbf{V}_j^{(k)}$ results in

$$\int_{-1}^{+1} \dot{L}_0(\tau)\mathbf{F}(\tau)\mathrm{d}\tau = -\mathbf{F}(-1) = -\alpha\beta_k \mathbf{f}_{v0}^{(k)}. \quad (38)$$

Furthermore, since the integrand in Eq. (38) is a polynomial of degree at most $2N-1$, it can be replaced exactly with an LG quadrature as

$$\int_{-1}^{+1} \dot{L}_0(\tau)\mathbf{F}(\tau)\mathrm{d}\tau = \sum_{i=1}^{N} w_i \dot{L}_0(\tau_i)\mathbf{F}(\tau_i), \quad (39)$$

where $\mathbf{F}(\tau_i)$ is equal to the discrete state dynamics of $\mathbf{v}(\tau)$ as given by the right-hand side of Eq. (9). Combining Eq. (38) and Eq. (39), $\mathbf{f}_{v0}^{(k)}$ can be written as

$$\mathbf{f}_{v0}^{(k)} = -\mathbf{\Gamma}_0\mathbf{W}\mathbf{f}_{v1:N}^{(k)}. \quad (40)$$

□

A process similar to that of Theorem 1 is used to derive the expressions $\tilde{\mathbf{D}}_{(N+1,1:N)} = \mathbf{\Gamma}_{N+1}\mathbf{W}\mathbf{D}_{(:,1:N)}$ and $\mathbf{f}_{vN+1}^{(k)} = \mathbf{\Gamma}_{N+1}\mathbf{W}\mathbf{f}_{v1:N}^{(k)}$, where $\mathbf{\Gamma}_{N+1} := [\dot{L}_{N+1}(\tau_1), \dot{L}_{N+1}(\tau_2), \dots, \dot{L}_{N+1}(\tau_N)]$. Such details are beyond the scope of this paper and can be found in Ref. [19].

The previously derived expressions can be substituted into the Lagrangian of Eq. (29), and then the KKT conditions are found by setting equal to zero the derivatives of the Lagrangian with respect to $\mathbf{X}_{0:N+1}^{(k)}$, $\mathbf{V}_{0:N+1}^{(k)}$, $\mathbf{U}_{1:N}^{(k)}$, $\mathbf{\Lambda}_{x1:N+1}^{(k)}$, $\mathbf{\Lambda}_{v1:N+1}^{(k)}$, $\tilde{\mathbf{\Lambda}}_{v0}^{(k)}$, $\tilde{\mathbf{\Lambda}}_{vN+1}^{(k)}$, $\mathbf{\Psi}$, Θ, β_k, t_0, and t_f. Along with the constraints given in Sections III and IV, the solution to the NLP of the modified LG collocation method must satisfy the following KKT conditions:

$$\mathbf{D}_{(:,i)}^\mathsf{T}\mathbf{\Lambda}_{x1:N}^{(k)} = \alpha\beta_k \nabla_{\mathbf{X}_i}\left(w_i\bar{\mathcal{H}}_i^{(k)}\right), \quad (41)$$

$$\begin{aligned}\mathbf{D}_{(:,i)}^\mathsf{T}\mathbf{\Lambda}_{v1:N}^{(k)} &= \alpha\beta_k \nabla_{\mathbf{V}_i}\left(w_i\bar{\mathcal{H}}_i^{(k)}\right) + \mathbf{\Gamma}_0\mathbf{W}\mathbf{D}_{(:,i)}\tilde{\mathbf{\Lambda}}_{v0}^{(k)} \\ &\quad -\mathbf{\Gamma}_{N+1}\mathbf{W}\mathbf{D}_{(:,i)}\tilde{\mathbf{\Lambda}}_{vN+1}^{(k)},\end{aligned} \quad (42)$$

$$\mathbf{0} = \alpha\beta_k \nabla_{\mathbf{U}_i}\left(w_i\bar{\mathcal{H}}_i^{(k)}\right), \quad (43)$$

$$\nabla_{\mathbf{X}_0}(\mathbf{\Psi}\mathbf{b}^\mathsf{T} - \mathcal{M}) = \mathbf{\Lambda}_{xN+1}^{(k)} - \mathbf{D}_{(:,0)}^\mathsf{T}\mathbf{\Lambda}_{x1:N}^{(k)}, \quad (44)$$

$$\begin{aligned}\nabla_{\mathbf{V}_0}(\mathbf{\Psi}\mathbf{b}^\mathsf{T} - \mathcal{M}) &= \mathbf{\Lambda}_{vN+1}^{(k)} - \mathbf{D}_{(:,0)}^\mathsf{T}\mathbf{\Lambda}_{v1:N}^{(k)} \\ &\quad -\tilde{\mathbf{D}}_{(0,0)}\tilde{\mathbf{\Lambda}}_{v0}^{(k)} - \tilde{\mathbf{D}}_{(N+1,0)}\tilde{\mathbf{\Lambda}}_{vN+1}^{(k)},\end{aligned} \quad (45)$$

$$\nabla_{\mathbf{X}_{N+1}}(\mathcal{M} - \mathbf{\Psi}\mathbf{b}^\mathsf{T}) = \mathbf{\Lambda}_{xN+1}^{(k)}, \quad (46)$$

$$\nabla_{\mathbf{V}_{N+1}}(\mathcal{M} - \mathbf{\Psi}\mathbf{b}^\mathsf{T}) = \mathbf{\Lambda}_{vN+1}^{(k)}, \quad (47)$$

$$\nabla_{t_0}(\mathcal{M} - \mathbf{\Psi}\mathbf{b}^\mathsf{T}) = \frac{1}{2}\sum_{k=1}^{K}\beta_k\sum_{i=1}^{N} w_i\bar{\mathcal{H}}_i^{(k)}, \quad (48)$$

$$\nabla_{t_f}(\mathbf{\Psi}\mathbf{b}^\mathsf{T} - \mathcal{M}) = \frac{1}{2}\sum_{k=1}^{K}\beta_k\sum_{i=1}^{N} w_i\bar{\mathcal{H}}_i^{(k)}, \quad (49)$$

$$\Theta = \alpha\sum_{i=1}^{N} w_i\bar{\mathcal{H}}_i^{(k)}, \quad (50)$$

where $\bar{\mathcal{H}}_i^{(k)} := \mathcal{L}_i^{(k)} + \langle \mathbf{f}_{xi}^{(k)}, \mathbf{\Lambda}_{xi}^{(k)}/w_i + \mathbf{\Lambda}_{xN+1}^{(k)}\rangle + \langle \mathbf{f}_{vi}^{(k)}, \mathbf{\Lambda}_{vi}^{(k)}/w_i + \mathbf{\Lambda}_{vN+1}^{(k)} - \mathbf{\Gamma}_{0,i}\tilde{\mathbf{\Lambda}}_{v0}^{(k)} + \mathbf{\Gamma}_{N+1,i}\tilde{\mathbf{\Lambda}}_{vN+1}^{(k)}\rangle$, $(i = 1,\dots,N;\ k = 1,\dots,K)$ is the discrete-time augmented Hamiltonian.

VII. TRANSFORMED ADJOINT SYSTEM

The transformed adjoint variables in the kth interval corresponding to the modified LG collocation method can be expressed as follows:

$$\boldsymbol{\lambda}_{x0}^{(k)} = \mathbf{\Lambda}_{xN+1}^{(k)} - \mathbf{D}_{(:,0)}^\mathsf{T}\mathbf{\Lambda}_{x1:N}^{(k)}, \quad (51)$$

$$\boldsymbol{\lambda}_{x1:N}^{(k)} = \mathbf{W}^{-1}\mathbf{\Lambda}_{x1:N}^{(k)} + \mathbf{1}\mathbf{\Lambda}_{xN+1}^{(k)}, \quad (52)$$

$$\boldsymbol{\lambda}_{xN+1}^{(k)} = \mathbf{\Lambda}_{xN+1}^{(k)}, \quad (53)$$

$$\begin{aligned}\boldsymbol{\lambda}_{v0}^{(k)} &= \mathbf{\Lambda}_{vN+1}^{(k)} - \mathbf{D}_{(:,0)}^\mathsf{T}\mathbf{\Lambda}_{v1:N}^{(k)} \\ &\quad - \tilde{\mathbf{D}}_{(0,0)}\tilde{\mathbf{\Lambda}}_{v0}^{(k)} - \tilde{\mathbf{D}}_{(N+1,0)}\tilde{\mathbf{\Lambda}}_{vN+1}^{(k)},\end{aligned} \quad (54)$$

$$\begin{aligned}\boldsymbol{\lambda}_{v1:N}^{(k)} &= \mathbf{W}^{-1}\mathbf{\Lambda}_{v1:N}^{(k)} + \mathbf{1}\mathbf{\Lambda}_{vN+1}^{(k)} \\ &\quad - \mathbf{\Gamma}_0^\mathsf{T}\tilde{\mathbf{\Lambda}}_{v0}^{(k)} + \mathbf{\Gamma}_{N+1}^\mathsf{T}\tilde{\mathbf{\Lambda}}_{vN+1}^{(k)},\end{aligned} \quad (55)$$

$$\boldsymbol{\lambda}_{vN+1}^{(k)} = \mathbf{\Lambda}_{vN+1}^{(k)}, \quad (56)$$

$$\boldsymbol{\psi} = \mathbf{\Psi}. \quad (57)$$

Finally, let $\mathbf{D}^\dagger$ be the $N\times(N+1)$ matrix derived in [18] given by

$$\mathbf{D}_{(i,j)}^\dagger = -\frac{w_j}{w_i}\mathbf{D}_{(j,i)}, \quad \mathbf{D}_{(i,N+1)}^\dagger = \sum_{i=1}^{N}\frac{w_j}{w_i}\mathbf{D}_{(j,i)}, \quad (58)$$

$(i,j = 1,\dots,N)$. Substituting the costate estimates of Eqs. (51)-(57) and the differentiation matrix $\mathbf{D}^\dagger$ of Eq. (58) into the KKT conditions given by Eqs. (41)-(50), the transformed adjoint system is given by

$$\left[\mathbf{D}_{(i,1:N+1)}^\dagger\right]^\mathsf{T}\boldsymbol{\lambda}_{x1:N+1}^{(k)} = -\alpha\beta_k\nabla_{\mathbf{X}_i}\left(\mathcal{H}_i^{(k)}\right), \quad (59)$$

$$\left[\mathbf{D}_{(i,1:N+1)}^\dagger\right]^\mathsf{T}\boldsymbol{\lambda}_{v1:N+1}^{(k)} = -\alpha\beta_k\nabla_{\mathbf{V}_i}\left(\mathcal{H}_i^{(k)}\right), \quad (60)$$

$$\mathbf{0} = \alpha\beta_k\nabla_{\mathbf{U}_i}\left(\mathcal{H}_i^{(k)}\right), \quad (61)$$

$$\boldsymbol{\lambda}_{x0}^{(k)} = \nabla_{\mathbf{X}_0}\left(\boldsymbol{\psi}\mathbf{b}^\mathsf{T} - \mathcal{M}\right), \quad (62)$$

$$\boldsymbol{\lambda}_{v0}^{(k)} = \nabla_{\mathbf{V}_0}\left(\boldsymbol{\psi}\mathbf{b}^\mathsf{T} - \mathcal{M}\right), \quad (63)$$

$$\boldsymbol{\lambda}_{xN+1}^{(k)} = \nabla_{\mathbf{X}_{N+1}}\left(\mathcal{M} - \boldsymbol{\psi}\mathbf{b}^\mathsf{T}\right), \quad (64)$$

$$\boldsymbol{\lambda}_{vN+1}^{(k)} = \nabla_{\mathbf{V}_{N+1}}\left(\mathcal{M} - \boldsymbol{\psi}\mathbf{b}^\mathsf{T}\right), \quad (65)$$

$$\nabla_{t_0}\left(\mathcal{M} - \boldsymbol{\psi}\mathbf{b}^\mathsf{T}\right) = \frac{1}{2}\sum_{k=1}^{K}\beta_k\sum_{i=1}^{N} w_i\mathcal{H}_i^{(k)}, \quad (66)$$

$$\nabla_{t_f}\left(\boldsymbol{\psi}\mathbf{b}^{\mathsf{T}} - \mathcal{M}\right) = \frac{1}{2}\sum_{k=1}^{K}\beta_k\sum_{i=1}^{N} w_i\mathcal{H}_i^{(k)}, \quad (67)$$

such that Eqs. (59)-(67) are the multiple-interval discrete representations of the continuous first-order optimality conditions from Eqs. (22)-(28).

VIII. NUMERICAL EXAMPLE

Consider the following optimal control problem. Minimize the final time, t_f, of a robotic arm reorientation maneuver [20] subject to the dynamic constraints

$$y_1' = \alpha y_2, \quad y_2' = \alpha u_1/L, \quad y_3' = \alpha y_4$$
$$y_4' = \alpha u_2/I_\theta, \quad y_5' = \alpha y_6, \quad y_6' = \alpha u_3/I_\phi,$$

the boundary conditions

$$(y_1(-1), y_1(+1)) = (4.5, 4.5), \quad (y_2(-1), y_2(+1)) = (0, 0),$$
$$(y_3(-1), y_3(+1)) = (0, 2\pi/3), \quad (y_4(-1), y_4(+1)) = (0, 0),$$
$$(y_5(-1), y_5(+1)) = (\pi/4, \pi/4), \quad (y_6(-1), y_6(+1)) = (0, 0),$$

and the control inequality constraints $-1 \le u_i \le +1$, $(i = 1, 2, 3)$, where $t_0 = 0$, t_f is free, $I_\phi = ((L - y_1)^3 + y_1^3)/3$, $I_\theta = I_\phi \sin^2(y_5)$, and $L = 5$. The optimal control for this example exhibits a bang-bang structure with five discontinuities located at $T \approx \{-0.5000, -0.3882, 0.0000, 0.3882, 0.5000\}$ with an optimal cost of $t_f^* \approx 9.1409$.

Figure 1 shows the optimal control solution obtained using the modified Legendre-Gauss (mLG) collocation method presented in this work alongside the multi-interval standard Legendre-Gauss (LG) collocation method with linear interpolation between points. The approximate locations of the five discontinuities and their uncertainty bounds are first determined using a jump function approximation [7] on an initial static mesh of ten uniformly spaced intervals with four collocation points in each interval. Setting the five discontinuity approximations as variable mesh points, a new variable mesh is generated consisting of six intervals with four collocation points in each interval. The resulting NLP is then solved using SNOPT [21] with the NLP error tolerance set to 10^{-6} and first-derivatives supplied using the automatic differentiation software ADiGator [22]. The first observation is that discrete controls are obtained at the left and right limits of each switch time only for the mLG method. Furthermore, the accuracy of the discontinuity locations is higher using the mLG method when compared with the LG method with variable mesh points. The higher accuracy of the mLG method demonstrates the improvement that the mLG method has on the NLP search space. The objective obtained using the LG method is $t_f = 9.1407 < t_f^*$. The lower objective obtained using the LG method can be interpreted as Lavrentiev phenomenon [15] because the additional degrees of freedom introduced via the inclusion of variable mesh points leads to an optimal objective value that differs from the true optimal value. The constraints imposed by the mLG method reduce the size of the larger NLP search space that is attributed to the added degrees of freedom. The augmented search space corresponding to the mLG method prevents convergence to a pseudo-minimizer and enables accurate computation of the optimal control solution.

Next, let $\boldsymbol{\lambda}_v = [\lambda_{y_2}, \lambda_{y_4}, \lambda_{y_6}]$ be the costates associated with the state dynamics that explicitly depend on the control. The approximations to these costate components are shown in Fig. 2. It can be seen that each control switch time, T_i, $i \in \{1, \dots, 5\}$, coincides with one of these costate components being zero, which is in agreement with Pontryagin's minimum principle. That is, because the Hamiltonian is linear in the control, the switching structure of the optimal control is determined by the switching function $\partial\mathcal{H}/\partial u_i$, $(i = 1, 2, 3)$, where $\partial\mathcal{H}/\partial u_i = 0$ at a control switch time (assuming no control is singular).

Lastly, the solution accuracy of the mLG method is compared to that of equivalent meshes discretized with modified LGR collocation (mLGR) [16] and standard LG collocation in Table I. Additionally, the problem was solved using the *hp*-Legendre(1) mesh refinement method of [7] with standard

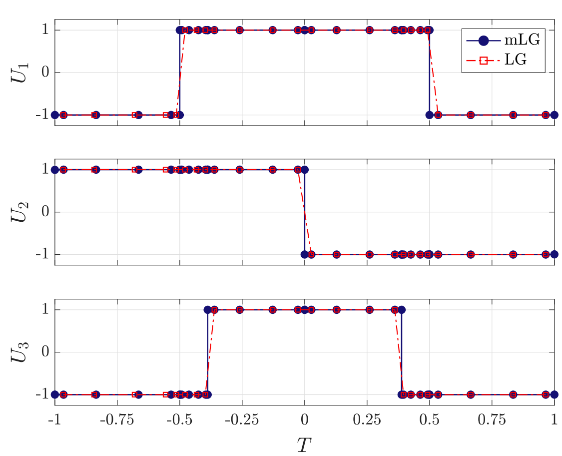

Fig. 1: Comparison of control solutions for robot arm reorientation maneuver.

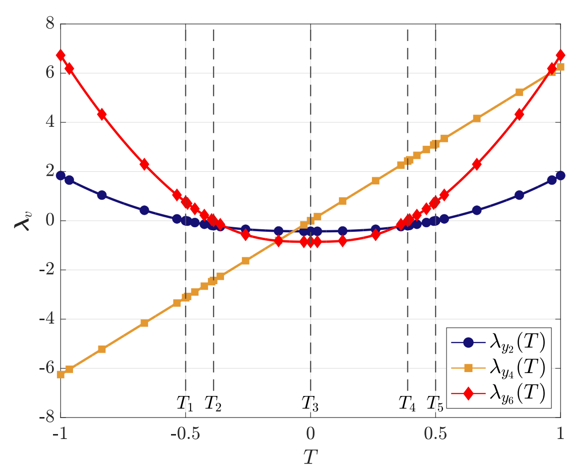

Fig. 2: Costate solutions, $\boldsymbol{\lambda}_v$, for robot arm reorientation maneuver obtained with modified LG collocation method.

TABLE I: Relative errors obtained using modified LG collocation compared to various methods.

	mLG	mLGR	LG	*hp*-Legendre(1)
E_{t_f}	2.8×10^{-7}	3.9×10^{-7}	2.4×10^{-5}	3.0×10^{-6}
E_{T_1}	2.1×10^{-7}	4.2×10^{-14}	4.7×10^{-2}	9.1×10^{-3}
E_{T_2}	4.7×10^{-7}	8.6×10^{-5}	5.4×10^{-5}	2.4×10^{-2}
E_{T_3}	8.9×10^{-8}	3.6×10^{-5}	2.8×10^{-5}	8.9×10^{-3}
E_{T_4}	6.6×10^{-7}	8.4×10^{-7}	5.5×10^{-5}	7.5×10^{-3}
E_{T_5}	2.6×10^{-7}	1.3×10^{-14}	1.1×10^{-4}	9.1×10^{-3}
$\max_{i \in \{1,\ldots,6\}} E_{y_i}$	6.4×10^{-5}	1.3×10^{-4}	9.5×10^{-3}	5.2×10^{-3}
$\max_{i \in \{1,2,3\}} E_{u_i}$	4.7×10^{-3}	1.1×10^{-2}	1.0×10^{-6}	8.4×10^{-1}

LGR collocation. The *hp*-Legendre(1) method uses jump function approximations to bracket detected discontinuities, but the refined mesh remains static. All relative errors were computed using a baseline solution that was obtained with $\mathbb{GPOPS}-\mathbb{II}$ [23] in multi-phase mode with the controls fixed to their optimal values. The mLG method results in the most accurate final cost and state approximations. The control solutions and switch times are also computed with high accuracy. While the maximum control error appears smallest with the LG collocation method, it is important to remark that the control is noticeably absent at the mesh points. Finally, the accuracy of the solution obtained with the *hp*-Legendre(1) method is negatively impacted by the static nature of the mesh, i.e. the mesh density increases near the control discontinuities but the switch times themselves are not optimized by the NLP solver.

IX. CONCLUSIONS

A modified Legendre-Gauss collocation method has been described for solving optimal control problems with nonsmooth solutions. The method augments the standard Legendre-Gauss direct collocation method by introducing additional control variables and variable mesh points as well as enforcing the dynamics at the previously non-collocated interval endpoints. It was shown that the KKT conditions from the NLP obtained via the modified Legendre-Gauss collocation method satisfy the variational conditions of the continuous optimal control problem. The method was demonstrated on a complex optimal control problem with multiple control switches. The results obtained in this paper demonstrate the viability of the modified Legendre-Gauss collocation method for solving optimal control problems with nonsmooth solutions when variable mesh points are located in the neighborhood of corresponding control discontinuities.

REFERENCES

[1] D. A. Benson, G. T. Huntington, T. P. Thorvaldsen, and A. V. Rao, "Direct trajectory optimization and costate estimation via an orthogonal collocation method," *Journal of Guidance, Control, and Dynamics*, vol. 29, pp. 1435–1440, Nov.-Dec. 2006.

[2] D. Garg, M. Patterson, W. W. Hager, A. V. Rao, D. A. Benson, and G. T. Huntington, "A unified framework for the numerical solution of optimal control problems using pseudospectral methods," *Automatica*, vol. 46, pp. 1843–1851, Nov. 2010.

[3] D. Garg, W. W. Hager, and A. V. Rao, "Pseudospectral methods for solving infinite-horizon optimal control problems," *Automatica*, vol. 47, pp. 829–837, Apr. 2011.

[4] D. Garg, M. A. Patterson, C. Francolin, C. L. Darby, G. T. Huntington, W. W. Hager, and A. V. Rao, "Direct trajectory optimization and costate estimation of finite-horizon and infinite-horizon optimal control problems using a Radau pseudospectral method," *Computational Optimization and Applications*, vol. 49, pp. 335–358, Oct. 2009.

[5] F. Fahroo and I. M. Ross, "Costate estimation by a Legendre pseudospectral method," *Journal of Guidance, Control, and Dynamics*, vol. 24, pp. 270–277, Mar. 2001.

[6] W. W. Hager, J. Liu, S. Mohapatra, A. V. Rao, and X.-S. Wang, "Convergence rate for a Gauss collocation method applied to constrained optimal control," *SIAM Journal on Control and Optimization*, vol. 56, pp. 1386–1411, Jan. 2018.

[7] A. T. Miller, W. W. Hager, and A. V. Rao, "Mesh refinement method for solving optimal control problems with nonsmooth solutions using jump function approximations," *Optimal Control Applications and Methods*, vol. 42, pp. 1119–1140, Mar. 2021.

[8] J. Zhao and S. Li, "Adaptive mesh refinement method for solving optimal control problems using interpolation error analysis and improved data compression," *Journal of the Franklin Institute*, vol. 357, no. 3, pp. 1603–1627, 2020.

[9] F. Liu, W. W. Hager, and A. V. Rao, "Adaptive mesh refinement method for optimal control using decay rates of Legendre polynomial coefficients," *IEEE Transactions on Control Systems Technology*, vol. 26, pp. 1475–1483, July 2018.

[10] M. Schlegel and W. Marquardt, "Detection and exploitation of the control switching structure in the solution of dynamic optimization problems," *J. Process Control*, vol. 16, pp. 275–290, Mar. 2006.

[11] M. Aghaee and W. W. Hager, "The switch point algorithm," *SIAM J. Control and Optimization*, vol. 59, pp. 2570–2593, Jan. 2021.

[12] P. Wang, C. Yang, and Z. Yuan, "The combination of adaptive pseudospectral method and structure detection procedure for solving dynamic optimization problems with discontinuous control profiles," *Industrial & Engineering Chemistry Research*, vol. 53, pp. 7066–7078, Apr. 2014.

[13] E. R. Pager and A. V. Rao, "Method for solving bang-bang and singular optimal control problems using adaptive Radau collocation," *Computational Optimization and Applications*, Jan. 2022.

[14] W. Chen, Y. Ren, G. Zhang, and L. T. Biegler, "A simultaneous approach for singular optimal control based on partial moving grid," *AIChE Journal*, vol. 65, p. e16584, Mar. 2019.

[15] J. M. Ball and G. Knowles, "A numerical method for detecting singular minimizers," *Numerische Mathematik*, vol. 51, pp. 181–197, 1987.

[16] J. D. Eide, W. W. Hager, and A. V. Rao, "Modified Legendre–Gauss–Radau collocation method for optimal control problems with nonsmooth solutions," *Journal of Optimization Theory and Applications*, vol. 191, pp. 600–633, Feb. 2021.

[17] D. E. Kirk, *Optimal Control Theory: An Introduction*. Courier Corporation, 2004.

[18] D. Benson, *A Gauss pseudospectral transcription for optimal control*. PhD thesis, Massachusetts Institute of Technology, 2005.

[19] G. Abadia-Doyle and A. V. Rao, "Modified Legendre-Gauss collocation method for solving optimal control problems with nonsmooth solutions," arXiv:2406.06824, June 2024.

[20] E. D. Dolan, J. J. More, and T. S. Munson, "Benchmarking optimization software with COPS 3.0.," tech. rep., Argonne National Lab, May 2004.

[21] P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM Review*, vol. 47, pp. 99–131, Jan. 2005.

[22] M. J. Weinstein and A. V. Rao, "Algorithm 984: ADiGator, a toolbox for the algorithmic differentiation of mathematical functions in MATLAB using source transformation via operator overloading," *ACM Transactions on Mathematical Software*, vol. 44, pp. 1–25, Aug. 2017.

[23] M. A. Patterson and A. V. Rao, "GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming," *ACM Transactions on Mathematical Software*, vol. 41, pp. 1–37, Oct. 2014.