

# A Nonlinear Modal-Based Framework for Low Computational Cost Optimal Control of 3D Very Flexible Structures

Marc Artola<sup>a</sup>, Andrew Wynn<sup>b</sup> and Rafael Palacios<sup>c</sup>

**Abstract**—A nonlinear modal-based reduced-order model, equipped with an efficient adjoint-sensitivity analysis, is presented as a low computational cost framework for optimal control of very flexible structures, with particular focus on efficiently computing finite rotations. Multiple shooting is shown to improve convergence of a highly nonlinear problem when compared to the single shooting case, with optimisation further accelerated via parallelisation, which suggests the presented approach may be employed for real-time control of very flexible structures.

## I. INTRODUCTION

Highly flexible structures are challenging to control since they are infinite dimensional systems with nontrivial nonlinear geometric couplings, meaning that linear structural models may not capture their fundamental kinematics. For slender structures, one-dimensional geometrically-exact beam theories (GEBT) were introduced to describe 3D dynamics [1] providing a good trade-off between model-fidelity and computational efficiency. Relevant examples include flexible aircraft wings or wind turbine blades, which are increasingly designed to be slender to improve aerodynamic efficiency, or certain robotic manipulators.

While linear control methodologies such as LQG or  $\mathcal{H}_\infty$  have been, and are still being applied [2], to control flexible structures, nonlinear model predictive control (NMPC) represents a versatile optimal control strategy which can explicitly handle nonlinear dynamics and control saturation [3, ch. 9]. A challenge of implementing NMPC, however, is the computational cost associated with solving its underlying optimisation problems. To address this problem, and to enable efficient online computation of control inputs for systems with fast dynamics, the *real-time iteration* scheme has been proposed [4], based on the multiple shooting algorithm of [5]. To date, application of NMPC strategies to the control of very flexible structures has yet to be fully exploited, mainly due to the high dimensionality of the systems involved. Recent applications to very flexible aircraft, addressing gust-load alleviation [6] and trajectory tracking [7], still rely on successive linearisations of the underlying system.

In this work, we present an optimisation framework to solve optimal control problems for very flexible structures. The approach employs a modal-based reduced order model

of the intrinsic beam equations [8], coupled with an efficient adjoint-based sensitivity analysis. Inspired by real-time NMPC, a multiple shooting scheme is implemented to improve computational efficiency. This facilitates parallelisation of simulations and prevents accumulation of nonlinear effects over long time horizons, in contrast to single shooting [9].

The proposed methodology facilitates direct and easy generalisation to 3D dynamics<sup>1</sup> compared to existing optimal control approaches involving GEBT and adjoint-based sensitivity analysis, which have, to date, only considered low dimensional nonlinear systems [11]. Further, the embedded adjoint system significantly improves the computational efficiency of evaluating sensitivities in high-dimensional optimisation problems, compared to finite-difference-based frameworks [12]. This highlights the potential of this framework for both optimal and real-time control of very flexible structures using fully nonlinear simulations.

## II. SYSTEM MODEL

In this section, an intrinsic formulation of the nonlinear structural dynamics followed by a modal-based finite-dimensional approximation is described, which can be generally applied to any 3D problem. However, we will particularise it for a 2D optimal control problem first described by Wang *et al.* [11], and later studied in [12]. The problem consists of finding the optimal control input  $u(t)$  so that the leftward horizontal velocity in the fixed frame of reference ( $e_1^*$ ,  $e_2^*$ ,  $e_3^*$  in Fig. 1) of a planar flexible pendulum's free end is maximised at a certain time  $t = t_f$ .

The energy of the actuation, involving a torque applied at the root of the hinged pendulum, is penalised by

$$P(u(t)) = \frac{1}{2} \int_0^{t_f} (\pi_1 u(t)^2 + \pi_2 \dot{u}(t)^2) dt, \quad (1)$$

where  $\pi_1$  and  $\pi_2$  are tuning constants.

### A. Structural model

A geometrically exact, fully-intrinsic beam model [8] is used. This model uses two state vector variables to describe the structure, namely, the vector of linear and angular inertial velocities  $\mathbf{x}_1(s, t) := [\mathbf{v}^\top, \boldsymbol{\omega}^\top]^\top : [0, L] \times \mathbb{R}_+ \rightarrow \mathbb{R}^6$  and the vector of force and moment sectional resultants  $\mathbf{x}_2(s, t) := [\mathbf{f}^\top, \mathbf{m}^\top]^\top : [0, L] \times \mathbb{R}_+ \rightarrow \mathbb{R}^6$ , with  $s$  and  $t$  denoting space and time variables, respectively, and  $L$  being the arclength of the undeformed structure. Variables  $\mathbf{x}_1$  and  $\mathbf{x}_2$  in this

<sup>\*</sup>The research has received funding from the EU's H2020 R&I programme under the Marie Skłodowska-Curie grant agreement No 765579.

All authors are with the Department of Aeronautics at the Imperial College London, Exhibition Road, London SW7 2AZ, UK

<sup>a</sup>marc.artola16@imperial.ac.uk

<sup>b</sup>a.wynn@imperial.ac.uk

<sup>c</sup>r.palacios@imperial.ac.uk

<sup>1</sup>The intrinsic description uses derivatives (strains and velocities) as primary variables resulting in a Hamiltonian description that extends the structure of rigid-body dynamics into infinite-dimensional systems [10].

formulation are referred to as the intrinsic variables, in contrast to displacements and rotations, which appear as derived variables. The nonlinear equations read

$$M\dot{\mathbf{x}}_1 - \mathbf{x}'_2 - E\mathbf{x}_2 + \mathcal{L}_1(\mathbf{x}_1)M\mathbf{x}_1 + \mathcal{L}_2(\mathbf{x}_2)C\mathbf{x}_2 = \mathbf{f}, \quad (2)$$

$$C\dot{\mathbf{x}}_2 - \mathbf{x}'_1 + E^\top \mathbf{x}_1 - \mathcal{L}_1^\top(\mathbf{x}_1)C\mathbf{x}_2 = 0. \quad (3)$$

Here,  $M(s) : \mathbb{R} \rightarrow \mathbb{R}^{6 \times 6}$ ,  $C(s) : \mathbb{R} \rightarrow \mathbb{R}^{6 \times 6}$ ,  $E(s) : \mathbb{R} \rightarrow \mathbb{R}^{6 \times 6}$  are the positive definite mass, positive semi-definite compliance and initial curvature and pre-twist matrices, while  $\mathcal{L}_1 : \mathbb{R}^6 \rightarrow \mathbb{R}^{6 \times 6}$  and  $\mathcal{L}_2 : \mathbb{R}^6 \rightarrow \mathbb{R}^{6 \times 6}$  are linear operators, detailed in [10]. Prime and dot superscripts denote partial derivatives with respect to  $s$  and  $t$ , respectively. Finally,  $\mathbf{f}$  represents external forces and moments per unit length acting along the structure.

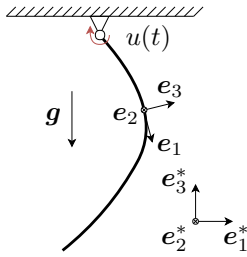


Fig. 1: Frames of reference involved in the pendulum problem.

Here, the main external forcing besides the control actuation is the weight of the structure. Gravity is a fixed vector in the fixed, inertial frame of reference while the state variables  $\mathbf{x}_1$  and  $\mathbf{x}_2$  in equations (2) and (3) are expressed in the local frame of reference ( $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  in Fig. 1), which moves and rotates together with the beam's elastic axis. Therefore, a transformation matrix  $T(s, t) : [0, L] \times \mathbb{R}_+ \rightarrow \mathbb{R}^{3 \times 3}$  from the local to the inertial frame is required. The gravity contribution in the momentum equation (2) is

$$\mathbf{f}_g(s, t) = \mu(s) \begin{bmatrix} I_3 \\ \tilde{\mathbf{r}}_{cm}(s) \end{bmatrix} T(s, t)^\top \mathbf{g}, \quad (4)$$

where  $\mu$  is the structure's mass per unit length,  $I_3$  is the identity matrix in  $\mathbb{R}^3$ ,  $\mathbf{r}_{cm}$  is the offset distance of the sectional centre of mass with respect to the elastic axis. The tilde superscript denotes the linear skew symmetric operator  $\tilde{\mathbf{a}}\mathbf{b} = \mathbf{a} \times \mathbf{b}$ , for two vector fields  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ ,

Rotation quaternions are used to parametrise rotations  $\boldsymbol{\xi}(s, t) = [\xi_0, \boldsymbol{\xi}_v^\top]^\top : [0, L] \times \mathbb{R}_+ \rightarrow \mathbb{R}^4$ , with  $\xi_0$  and  $\boldsymbol{\xi}_v$  being the scalar and vector parts, respectively. They are subject to spatial and temporal evolution equations [13, ch. 26]:

$$\boldsymbol{\xi}' = \mathcal{U}(\boldsymbol{\kappa} + \boldsymbol{\kappa}_0)\boldsymbol{\xi}, \quad (5)$$

$$\dot{\boldsymbol{\xi}} = \mathcal{U}(\boldsymbol{\omega})\boldsymbol{\xi}, \quad (6)$$

where  $\boldsymbol{\kappa}$  is the vector of curvatures and twist, that is, the three last components of the product  $C\mathbf{x}_2$ , and  $\boldsymbol{\kappa}_0$  is the vector of initial curvatures and pre-twist.  $\mathcal{U} : \mathbb{R}^3 \rightarrow \mathbb{R}^{4 \times 4}$  is a linear skew-symmetric operator, defined for a vector field  $\mathbf{a} \in \mathbb{R}^3$  as

$$\mathcal{U}(\mathbf{a}) = \frac{1}{2} \begin{bmatrix} 0 & -\mathbf{a}^\top \\ \mathbf{a} & -\tilde{\mathbf{a}} \end{bmatrix}. \quad (7)$$

The transformation matrix associated to a given set of quaternions  $\boldsymbol{\xi}$  is [13, ch. 6]

$$T(\boldsymbol{\xi}) = (1 - 2\|\boldsymbol{\xi}_v\|^2)I_3 + 2\boldsymbol{\xi}_v\boldsymbol{\xi}_v^\top + 2\xi_0\tilde{\boldsymbol{\xi}}_v. \quad (8)$$

The displacement field  $\mathbf{r}$  can be retrieved *a posteriori* in a similar way in which the rotations are obtained, using the following spatial or time evolution equations

$$\mathbf{r}' = T(\mathbf{e}_1 + \boldsymbol{\gamma}), \quad (9)$$

$$\dot{\mathbf{r}} = T\mathbf{v}, \quad (10)$$

where  $\mathbf{e}_1 = [1, 0, 0]^\top$  and  $\boldsymbol{\gamma}$  is the strain vector in local components, that is, the three first components of  $C\mathbf{x}_2$ .

Given a boundary control input  $u(t)$ , defined as a point moment at the support as shown in Fig. 1, the problem is completely defined by adding to the set of PDEs (2), (3) and (6), the boundary conditions

$$\begin{aligned} \mathbf{x}_1(0, t) &= [0, 0, 0, 0, \cdot, \cdot, 0]^\top, \\ \mathbf{x}_2(0, t) &= [\cdot, \cdot, \cdot, \cdot, -u(t), \cdot]^\top, \\ \mathbf{x}_2(L, t) &= [0, 0, 0, 0, 0, 0]^\top, \end{aligned} \quad (11)$$

and the initial conditions

$$\mathbf{x}_1(s, 0) = \bar{\mathbf{x}}_1(s), \quad \mathbf{x}_2(s, 0) = \bar{\mathbf{x}}_2(s), \quad \boldsymbol{\xi}(s, 0) = \bar{\boldsymbol{\xi}}_1(s). \quad (12)$$

Here,  $\bar{\mathbf{x}}_1(s)$ ,  $\bar{\mathbf{x}}_2(s)$  and  $\bar{\boldsymbol{\xi}}(s)$  describe the (stable) equilibrium position of (2), (3) and (6). The term  $-D\mathbf{x}_1''$ , with  $D = D^\top > 0$ , will be added to the LHS of (2) to model structural damping and the dots in (11) indicate that no value is prescribed.

## B. Modal Reduction

To construct a reduced-order model, the natural modes of the structure are computed [14]. These are eigenfunctions  $\phi_{1j}(s), \phi_{2j}(s) : [0, L] \rightarrow \mathbb{R}^6$  associated with the linearisation of (2) and (3) around the unloaded ( $\mathbf{f} = \mathbf{0}$ ) and undeformed condition ( $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{0}$ ). After  $N_m$  eigenfunctions are computed, approximate state variables are defined by

$$\mathbf{x}_1(s, t) = \phi_{1j}(s)q_{1j}(t), \quad \mathbf{x}_2(s, t) = \phi_{2j}(s)q_{2j}(t) \quad (13)$$

with coefficients  $q_{1j}(t), q_{2j}(t) : \mathbb{R}_+ \rightarrow \mathbb{R}$  (unless otherwise stated, Einstein's summation convention for index  $j=1, \dots, N_m$  is used). The reduced order model is obtained by a Galerkin projection, in which evolution equations for  $q_{1j}$  and  $q_{2j}$  are obtained by substituting (13) into (2) and (3), pre-multiplying by each eigenfunction and integrating over the spatial domain.

Concerning the rotations, a spatial grid with  $N_\xi$  intervals is constructed, giving rise to  $N_\xi + 1$  nodes in which the quaternions are tracked. Focusing on the  $k^{th}$  interval, defined for  $s_k \leq s \leq s_{k+1}$ , the Catmull-Rom spline [13, ch. 25] is a third order polynomial that uses the quaternions  $\boldsymbol{\xi}_{k-1}(t), \boldsymbol{\xi}_k(t), \boldsymbol{\xi}_{k+1}(t), \boldsymbol{\xi}_{k+2}(t)$  to construct a curve matching the value of the central ones,  $\boldsymbol{\xi}_k = \boldsymbol{\xi}(s_k)$  and  $\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi}(s_{k+1})$ . In addition, this interpolation matches the spatial derivatives at the central nodes with the derivatives of the splines constructed on the adjacent intervals, which ensures smooth

matching of all tracked quaternions. In this way, the spatial distribution of quaternions reads

$$\boldsymbol{\xi}(s, t) = \Lambda(s)\boldsymbol{\Xi}(t), \quad (14)$$

where  $\boldsymbol{\Xi}(t) = [\boldsymbol{\xi}_1^\top, \dots, \boldsymbol{\xi}_{N_\xi+1}^\top]^\top : \mathbb{R}_+ \rightarrow \mathbb{R}^{4(N_\xi+1)}$  is the set of tracked quaternions and  $\Lambda(s) : [0, L] \rightarrow \mathbb{R}^{4 \times 4(N_\xi+1)}$  is an interpolation matrix with piece-wise defined cubic polynomial entries.

When (14) is inserted in (8), cross-product quadratic terms between the tracked quaternions give the gravity contribution (4) on modal coordinates,  $\boldsymbol{\eta}_g(\boldsymbol{\Xi})$ , the following structure:

$$\boldsymbol{\eta}_g(\boldsymbol{\Xi}) = \begin{bmatrix} \eta_{0,1} \\ \vdots \\ \eta_{0,N_m} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Xi}^\top G_1 \boldsymbol{\Xi} \\ \vdots \\ \boldsymbol{\Xi}^\top G_{N_m} \boldsymbol{\Xi} \end{bmatrix}. \quad (15)$$

Constant terms  $\eta_{0,1}, \dots, \eta_{0,N_m}$  in (15) arise from the identity matrix in (8), while matrices  $G_1, \dots, G_{N_m}$  present a  $4 \times 4$  block heptadiagonal structure (4 quaternions, with 4 components each, take part in each interval's interpolation).

Regarding the control input, it is incorporated into the model as a point force term added to the RHS of (2), with entries of the modal contribution vector

$$[\boldsymbol{\eta}_u]_i = \int_0^L \boldsymbol{\phi}_{1i}^\top e_5 \delta(s) ds = \boldsymbol{\phi}_{1i}(0)^\top e_5, \quad (16)$$

where  $e_5$  is the fifth element of the canonical basis of  $\mathbb{R}^6$  and  $\delta(s)$  represents the delta function at  $s = 0$ .

The resulting system of ordinary differential equations is

$$\dot{\boldsymbol{q}} = W\boldsymbol{q} + N(\boldsymbol{q})\boldsymbol{q} + \begin{bmatrix} \boldsymbol{\eta}_g(\boldsymbol{\Xi}) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta}_u \\ \mathbf{0} \end{bmatrix} u(t), \quad (17)$$

$$\dot{\boldsymbol{\Xi}} = S(\boldsymbol{q}_1)\boldsymbol{\Xi} := \begin{bmatrix} S_1(\boldsymbol{q}_1) & & \\ & \ddots & \\ & & S_{N_\xi+1}(\boldsymbol{q}_1) \end{bmatrix} \boldsymbol{\Xi}, \quad (18)$$

where the expansion coefficients  $q_{1i}$  and  $q_{2i}$  have been grouped in a column vector  $\boldsymbol{q}(t) : \mathbb{R}_+ \rightarrow \mathbb{R}^{2N_m}$ . The term  $N(\boldsymbol{q})$  is linear in  $\boldsymbol{q}$ , that is  $N(\boldsymbol{q}) = q_{1j}\Gamma_{1j} + q_{2j}\Gamma_{2j}$  for some constant matrices  $\Gamma_{1j}, \Gamma_{2j} \in \mathbb{R}^{2N_m \times 2N_m}$ , making (17) quadratic in  $\boldsymbol{q}$ . Further details of its structure are found in [10]. The block-diagonal entries in (18) are given by

$$S_k(\boldsymbol{q}_1) := U_{kj}q_{1j} \quad (19)$$

with coefficient matrices  $U_{kj} \in \mathbb{R}^{4 \times 4}$

$$U_{kj} := \mathcal{U}(\Pi_\omega \boldsymbol{\phi}_{1j}(s_k)). \quad (20)$$

Here,  $\Pi_\omega := [0, I_3]$  selects the components of  $\boldsymbol{\phi}_{1j}$  referring to angular velocities. Initial conditions for (17) are found by projecting (12) onto the orthogonal [15] modal basis.

The main difference of our approach with that of [16] lays with the fact that only a few points of the structure, where the quaternions' time-evolution is evaluated, need to be selected in order to time-march the resulting system of equations, reducing computational cost without significantly compromising the accuracy of the rotation distribution, as seen in section IV.

A consequence of using the modal solution, which is global in space, to locally evaluate the time evolution equations for rotations (6) and displacements (10) is that errors are introduced which accumulate over the time-integration and result in the individually evaluated points in the structure drifting apart [17]. Periodic regularisations in the rotation and displacement fields are performed to keep these errors within some tolerable margin, updating them by integrating equations (5) and (9) in space [17].

### III. OPTIMAL PROBLEM SOLUTION

#### A. Multiple shooting

In a multiple shooting scheme the time horizon  $t_f$  in which the simulation and optimisation are performed is split into several independent time intervals, with independent initial conditions  $\boldsymbol{q}_0^m = \boldsymbol{q}^m(t_{m-1})$ ,  $\boldsymbol{\Xi}_0^m = \boldsymbol{\Xi}^m(t_{m-1})$  for each interval. These are added to the set of optimisation parameters with additional state continuity constraints [5]. Superscript  $m$  denotes that the variables belong to the  $m^{\text{th}}$  interval, with  $t_{m-1} \leq t \leq t_m$ . The multiple shooting version of the optimal problem in [16] is written in terms of the variables of the reduced-order model as

$$\min_{u(t), \boldsymbol{q}_0^m, \boldsymbol{\Xi}_0^m} \kappa_1 \boldsymbol{e}_1^\top T(\Pi_L \boldsymbol{\Xi}(t_f)) \mathcal{L}_{v_L} \boldsymbol{q}_1(t_f) + \kappa_2 P(u(t)) \quad \text{subject to:} \quad (21a)$$

$$-\dot{\boldsymbol{q}}^m + W\boldsymbol{q}^m + N(\boldsymbol{q}^m)\boldsymbol{q}^m + [\boldsymbol{\eta}_g^\top(\boldsymbol{\Xi}^m) + \boldsymbol{\eta}_u^\top u(t), \mathbf{0}^\top]^\top - \dot{\boldsymbol{\Xi}}^m + S(\boldsymbol{q}_1^m)\boldsymbol{\Xi}^m = \mathbf{0}, \quad (21b)$$

$$\boldsymbol{q}^m(t_{m-1}) = \boldsymbol{q}_0^m, \quad \boldsymbol{\Xi}^m(t_{m-1}) = \boldsymbol{\Xi}_0^m, \quad (21c)$$

$$\boldsymbol{q}_0^m = \boldsymbol{q}_f^{m-1}, \quad \boldsymbol{\Xi}_0^m = \boldsymbol{\Xi}_f^{m-1}, \quad m > 2, \quad (21d)$$

$$u(0) = u(T) = 0, \quad -M_{max} \leq u(t) \leq M_{max}, \quad (21e)$$

where  $\kappa_1$  and  $\kappa_2$  are weights which trade off performance and actuation penalty. The projection operator  $\Pi_L = [0, 0, \dots, I_4]$  selects the quaternion located at  $s = L$ . The linear operator  $\mathcal{L}_{v_L} := \Pi_v \boldsymbol{\phi}_{1j}(L)q_{1j} : \mathbb{R}^{N_m} \rightarrow \mathbb{R}^3$  maps the coefficients  $\boldsymbol{q}_1$  to the velocity at the end node, with  $\Pi_v = [I_3, 0]$  used to select the entries referring to linear velocities. Variables  $\boldsymbol{q}_f^m = \boldsymbol{q}^m(t_m)$  and  $\boldsymbol{\Xi}_f^m = \boldsymbol{\Xi}^m(t_m)$  are the final states of each time interval.

#### B. Adjoint-based sensitivity analysis

We can reformulate the objective function as

$$F := e(\boldsymbol{x}(t_f), \boldsymbol{p}) + \int_0^{t_f} f(\boldsymbol{x}(t), \boldsymbol{p}) dt, \quad (22)$$

where  $f$  and  $e$  are functionals  $f, e : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ , for some vector variables  $\boldsymbol{x} \in \mathbb{R}^n$  and  $\boldsymbol{p} \in \mathbb{R}^m$ . The vector variable  $\boldsymbol{x}$  is time dependent and governed by an ODE, written in implicit form  $\boldsymbol{h}(\dot{\boldsymbol{x}}, \boldsymbol{x}, \boldsymbol{p}) = \mathbf{0}$ , with initial conditions given by  $\boldsymbol{g}(\boldsymbol{x}(0), \boldsymbol{p}) = \mathbf{0}$ , while  $\boldsymbol{p}$  denotes the vector of optimisation parameters. We choose to parametrise the control signal as

$$u(t) = \sum_{l=1}^{N_c} \boldsymbol{\phi}_{ul}(t) u_l, \quad (23)$$

where  $\phi_{ul}$  can be any arbitrary local or global expansion function and  $u_l$  are the expansion coefficients. A sine series is used, in this case, for it automatically satisfies the initial and final constraints on the control signal in (21). Then, our vector of optimisation parameters  $\mathbf{p}$  is formed by the vertical concatenation of  $\mathbf{q}_0^m$ ,  $\Xi_0^m$  and  $u_l$ . The adjoint sensitivity analysis [18] applied to our multiple shooting construction allows us computing the derivative of (22) with respect  $\mathbf{p}$  by

$$\frac{dF}{d\mathbf{p}} = \int_0^{t_f} \left( \frac{\partial f}{\partial \mathbf{p}} + \boldsymbol{\lambda}^\top \frac{\partial \mathbf{h}}{\partial \mathbf{p}} \right) dt + [\boldsymbol{\lambda}^\top(0), 0], \quad (24)$$

with multipliers  $\boldsymbol{\lambda}$  subject to the ODE and final condition

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial \mathbf{h}}{\partial \mathbf{x}}^\top \boldsymbol{\lambda}, \quad \boldsymbol{\lambda}(t_f) = \frac{\partial e}{\partial \mathbf{x}(t_f)}^\top, \quad (25)$$

where the terminal penalty  $e$  in (24) corresponds to the final velocity term in the cost function (21),  $f$  is the actuation penalty term  $P$  and  $\mathbf{x} = [\mathbf{q}^\top, \Xi^\top]^\top$ . Functionals  $\mathbf{h}$  and  $\mathbf{g}$  in (24) are the ODE (21b) and initial conditions (21c) of our problem. Note how the gradient (24), once the multipliers  $\boldsymbol{\lambda}$  are obtained, is easy to compute since partial derivatives with respect the parameters  $\mathbf{p}$  are typically inexpensive to calculate. The adjoint equation (25) must be solved by backward integration, from time  $t = t_f$  to  $t = 0$ .

Relevant terms to solve the adjoint system (25) are easily obtained due to the compact structure of (17) and (18):

$$\frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{bmatrix} W + N(\mathbf{q}) + N^*(\mathbf{q}) & \Xi^\top \begin{bmatrix} (G_1 + G_1^\top) \\ \vdots \\ (G_{N_m} + G_{N_m}^\top) \\ 0 \end{bmatrix} \\ S^*(\Xi) & S(\mathbf{q}_1^m) \end{bmatrix}, \quad (26)$$

where  $N^*(\mathbf{q})$  and  $S^*(\Xi)$  are also linear in  $\mathbf{q}$  and  $\Xi$ , revealing that (26) is linear in both  $\mathbf{q}$  and  $\Xi$ , and

$$\frac{\partial e}{\partial \mathbf{x}(t_f)} = \kappa_1 \left[ \mathbf{e}_1^\top T(\boldsymbol{\xi}_L) \mathcal{L}_{v_L}, \mathbf{0}^\top, \mathbf{e}_1^\top \frac{\partial}{\partial \boldsymbol{\xi}_L} (T(\boldsymbol{\xi}_L) \mathcal{L}_{v_L} \mathbf{q}_1) \right], \quad (27)$$

evaluated at  $t = t_f$  and where  $\mathbf{0}^\top$  has the appropriate dimensions to account for all  $q_{2j}$  coefficients and all tracked quaternions, except for  $\boldsymbol{\xi}_L = \Pi_L \Xi$ .

Finally, the last interesting term appearing in (24) is

$$\frac{\partial \mathbf{h}}{\partial \mathbf{p}} = \left[ 0, 0, \boldsymbol{\eta}_u^* \frac{\partial u}{\partial u_l} \right] = [0, 0, \boldsymbol{\eta}_u^* \phi_{ul}]. \quad (28)$$

Here,  $\boldsymbol{\eta}_u^*$  includes the contribution of the control to  $\mathbf{q}_2$  and to  $\Xi$ , which is zero. The term  $\partial f / \partial \mathbf{p}$  is straightforward to compute since  $f$  depends explicitly and uniquely on  $u_l$ .

### C. Optimisation scheme

The problem (21) is iteratively solved via Sequential Quadratic Programming [19, ch. 18], by which each iteration step  $\Delta \mathbf{p} = [\Delta \mathbf{q}_0^m, \Delta \Xi_0^m, \Delta u_l]^\top$  is solution to the QP

$$\min_{\Delta \mathbf{q}_0^m, \Delta \Xi_0^m, \Delta u_l} \Delta \mathbf{p}^\top \frac{dF}{d\mathbf{p}} + \frac{1}{2} \Delta \mathbf{p}^\top H_i \Delta \mathbf{p} \quad (29a)$$

subject to:

$$(\mathbf{q}_0^m)_i + \Delta \mathbf{q}_0^m = (\mathbf{q}_f^{m-1})_i + \frac{\partial \mathbf{q}_f^{m-1}}{\partial \Delta \mathbf{p}} \Delta \mathbf{p}, \quad m > 2, \quad (29b)$$

$$(\Xi_0^m)_i + \Delta \Xi_0^m = (\Xi_f^{m-1})_i + \frac{\partial \Xi_f^{m-1}}{\partial \Delta \mathbf{p}} \Delta \mathbf{p}, \quad m > 2,$$

$$-M_{max} \leq (u_l)_i + \Delta u_l \leq M_{max}, \quad (29c)$$

constructed by replacing the original cost function (21a) by a second-order approximation around the  $i^{th}$  optimisation iteration, using the gradient of the objective function  $dF/d\mathbf{p}$  and the Hessian of the Lagrangian<sup>2</sup>  $H$ . The constraints are substituted by a first-order approximation [19, ch. 18].

With the adjoint method one is able to construct the optimal problem in the form (29) from (21), embedding the ODE-related constraints in the gradient computation. A similar adjoint system to (25) is needed to obtain the derivatives in (29b). The main differences are that  $f = 0$  in (24), that the adjoint state is a matrix, and that the cost function is now the final state vector (then (27) becomes the identity). For the computation of  $H$ , BFGS approximation [19, ch. 18] is used.

## IV. NUMERICAL RESULTS

The structural parameters have been selected to match [11]:  $M = \text{diag}(\mu, \mu, \mu, 0, 0, 0)$ , with  $\mu = 1 \text{ kgm}^{-1}$ ,  $C = \text{diag}(EA, \infty, GA_3, \infty, EI_2, \infty)^{-1}$ , with  $EA = 10^4 \text{ N}$ ,  $GA_3 = 10^6 \text{ N}$  and  $EI_2 = 1 \text{ Nm}^2$  and  $D = \text{diag}(0, 0, 0, 0, d_b, 0)$ , with  $d_b = 0.01 \text{ Nms}^2$ . The pendulum is straight in its equilibrium, with length  $L = 1 \text{ m}$  and no centre of mass offset.

In section IV-A a convergence analysis is performed using an open-loop simulation with one of the optimal control input solutions in [12], which should trigger highly nonlinear effects and similar dynamics to the ones involved in the optimal control case, while section IV-B solves the optimal problem (21) with,  $t_f = 2 \text{ s}$ ,  $M_{max} = 3.5 \text{ Nm}$ , and tuning constants  $\kappa_1 = 1$ ,  $\kappa_2 = 1$ ,  $\pi_1 = 1$ ,  $\pi_2 = 0.01$ , with appropriate dimensional units.

The presented results have been computed using a model constructed with  $N_m = 21$  modes, selected in ascending frequency order, including 10 axial and 10 bending elastic modes and the only rigid-body mode of the pendulum. A fourth-order Runge-Kutta explicit scheme has been used. Even though the modal-based ROM has been shown to be energy-preserving [10], the RK4 scheme might introduce some numerical dissipation. However, the time step used, equal to 0.1 times the period of the highest frequency mode ( $\Delta t = 2.11 \cdot 10^{-4} \text{ s}$ ), which has been found to suffice convergence criteria, produces an energy loss with respect the input energy<sup>3</sup> of 0.06% in the undamped case ( $d_b = 0$ ).

A uniform grid with  $N_\xi = 10$  has been constructed to keep track of the quaternions and very satisfactory results

<sup>2</sup>The Lagrangian is  $\mathcal{L} = F + \boldsymbol{\lambda}^\top \mathbf{g}$ , where  $F$  is the cost function (21a) and  $\boldsymbol{\lambda}$  the multipliers associated to constraints (21e)–(21d) grouped in  $\mathbf{g} = \mathbf{0}$ .

<sup>3</sup>Input energy is given by  $\epsilon_{in}(t) = \int_0^t u(\tau) \omega_2(0, \tau) d\tau$  while system's energy is  $\epsilon(t) = \frac{1}{2} \mathbf{q}^\top \mathbf{q} + \int_0^L \mu g e_3^\top \mathbf{r} ds$

are observed. Even when the highest space-frequency modes are excited towards the end of the simulation ( $t = 1.83$  s), the interpolation distribution has a relative RMS error of 0.038% with respect to the distribution obtained by the spatial integration of (5), with good unit-norm preservation properties (0.07% maximum error in the norm). Rotation regularisations are performed every 50 time steps. These are performed by integrating (5) using a RK4 scheme with a spatial resolution of  $0.02L$ . The tracked quaternions on each grid point, time-integrated up to that instant (18), are then substituted by those resulting from the spatial integration.

### A. Convergence study

The main convergence exercise involves the choice of the modal basis size  $N_m$ , which should include the relevant modes in the dynamics of the problem making the resulting ROM reasonable. An analysis of the reduction error is shown in Fig. 2, where each basis has been constructed using an equal number of axial and bending modes.

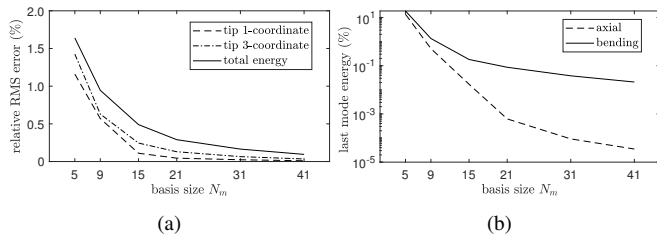


Fig. 2: RMS error (a) and last mode energy contribution (b) vs  $N_m$ .

In Fig. 2a the RMS errors<sup>4</sup> of the tip coordinates and the total energy of the system are displayed for modal expansions with varying base size. The error is computed with respect to a reference solution with  $N_m = 81$ , for which the results are considered fully converged. All simulations have been performed with the same time step and regularisation rate. Fig. 2b displays another convergence metric: the maximum energy of the highest-frequency axial and bending modes (i.e. those defining the truncation threshold). The energies are normalised using the total (maximum) elastic energy (excluding the rigid-body mode). This metric provides us with an idea of how well the expansion resolves the physics, with low values of energy indicating negligible contribution of the last modes in the problem.

It has been considered that a basis with  $N_m = 21$  provides a reasonable computational cost while accurately reproducing the dynamics of the problem. With this basis, the last bending and axial modes represent a 0.086% and a 0.00062% of the maximum elastic energy, in contrast to the clearly unresolved case with  $N_m = 5$ , where these ratios are 18.59% and 13.57%, respectively.

Wang's *et al.* [11] nonlinear optimal problem statement assumes a structure with infinite axial and transverse stiffness. An important feature of the intrinsic formulation is that the change of primary variables linearises the equations of motion, that is, nonlinear effects come from the coupling between axial and bending modes. Setting infinite axial

stiffness  $EA$  automatically forces all axial modes to vanish meaning that nonlinear phenomena (i.e. coupling between bending modes) are lost [15]. Therefore, convergence as  $EA \rightarrow \infty$  with finite axial force has also been studied, with a value of  $EA = 10^4$  already providing converged results.

### B. Multiple shooting results

The optimal problem (21) has been solved using a discrete sine series with  $N_c = 16$  and frequencies  $f_l = 0.5l/t_f$  as a control parametrisation (23), and for a number of multiple shooting time intervals ranging from 4 to 95. A consequence of the rotation regularisations is that the spatial integrations (5) have to be accounted for in the adjoint system, using a similar approach taken when computing the derivatives required in (29b). The multiple shooting time-intervals are chosen to start when a regularisation is scheduled, so that only  $\xi_1 = \xi(s=0)$  is part of the set of optimal parameters  $p$ .

Results are obtained with a Matlab code running on an Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz processor, using Matlab's default parallel pool with 12 workers and with an optimisation stop tolerance of  $10^{-5}$  on  $d\mathcal{L}/dp$ .

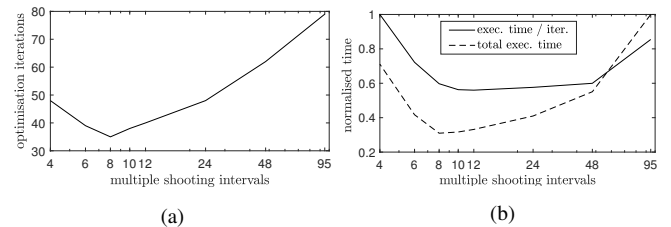


Fig. 3: Optimisation iterations (a), and execution time (norm. by 5,076 s) and execution time/iteration (norm. by 75.19 s) (b) vs number of intervals.

As seen in Fig. 3a, the multiple shooting scheme improves convergence of the nonlinear problem by reducing the number of optimisation (SQP) iterations until 8 intervals, where there is a minimum. Further increasing this number, however, increases the dimension of the optimal problem and convergence is slowed. The total execution time (dashed line in Fig. 3b), also displays a minimum at 8 intervals, which results from the combined effect of the mentioned multiple shooting influence on the convergence (Fig. 3a) and the parallelisation acceleration (solid line in Fig. 3b). This last plot clearly displays how the execution time per optimisation iteration decreases until 12 intervals (note that 12 workers are used), it remains roughly constant for the range 12 – 48 after which it increases again, since the QP solution (29) takes over the forward and backward (adjoint) time-integrations, in terms of computational cost.

A number of intervals below 4 makes the proposed scheme unstable and the single shooting case (i.e. the case in which the problem's time horizon  $t_f$  is taken as a whole and no time intervals are constructed) fails unless initial sub-relaxation is applied. Applying a factor between 0.1 and 0.5 within the first 30 iterations to the step size  $\Delta p$  it converges in 47 iterations towards the same solution with an execution time 8.6 times that of the minimum obtained with the 8-interval case (note this is still a reasonable execution time when compared to the single shooting approach of [12], which required

<sup>4</sup>Given a reference  $x_r(t)$ , the error is  $\sqrt{\int_0^{t_f} (x - x_r)^2 dt / \int_0^{t_f} x_r^2 dt}$

the use of the UK's National Supercomputing resource, ARCHER). Hence, the multiple shooting implementation not only accelerates the simulation but also becomes necessary to ensure convergence for highly nonlinear problems (at least with the proposed optimisation algorithm).

The solution's control input  $u(t)$  of the optimal control problem (21) is shown in Fig. 4a, while the snapshots of the pendulum's motion are presented in Figs. 4b—4d. The solution has a minimum in  $F = -6.64$ , with an end-of-simulation leftward horizontal velocity of  $12.06 \text{ ms}^{-1}$ . These values differ slightly from Maraniello's [12], where a minimum of  $F = -9.63$  and a velocity of  $13.11 \text{ ms}^{-1}$  are obtained. Discrepancies are attributed to the fact that [12] does not account for structural damping but uses a Newmark- $\beta$  integration scheme which adds numerical dissipation.

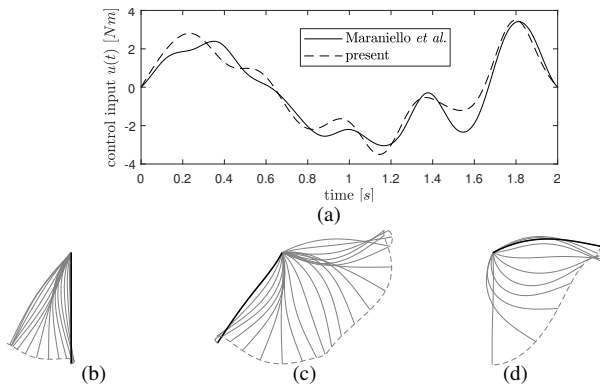


Fig. 4: Optimal control input (a). Snapshots (grey) of the pendulum trajectory at a rate of 25 fps for  $t \in [0, 0.72]$  (b),  $t \in [0.72, 1.60]$  (c) and for  $t \in [1.60, 2]$  (d). The black line represents the initial configuration of the shown interval and the grey, dashed, one the tip's path.

## V. CONCLUSIONS

A multiple shooting approach using a nonlinear modal-based finite-dimensional approximation for the optimal control of flexible structures has been introduced.

A first advantage of the presented framework is the reduction in the complexity of the adjoint system construction, due to the natural integration of rigid-body dynamics into the Hamiltonian framework generated by the intrinsic description. Moreover, nonlinear terms are quadratic and their derivatives are inexpensive to compute.

Second, the multiple shooting scheme has been shown to successfully help in the convergence of highly nonlinear problems as well as in speeding up simulations through parallelisation. An important remark is that multiple shooting increases the dimension of the optimal problem, making the use of an efficient sensitivity-extraction method (such as the adjoint-based) essential, whose cost is roughly independent of the number of parameters as opposed to techniques such as Automatic Differentiation, present in state-of-the-art optimal control packages. Also, the regularisations on the rotation field help to reduce the total number of optimisation parameters.

Finally, even though execution times appear rather high for online control, the modal approach indicates that only

a few modes are required to capture the main sources of nonlinearity, providing a good prediction of the system dynamics at a very low computational-cost. This feature can be exploited in a *real-time iteration* alike NMPC scheme, where model mismatch is allowed and only an estimate of the optimal control solution is required at each sampling time. Improving computational efficiency of the on-going implementation and/or application to systems with somewhat slower dynamics (e.g. aircraft) suggest that nonlinear control of very flexible structures can be attainable.

## REFERENCES

- [1] J. C. Simo, "A Finite Strain Beam Formulation. The Three-Dimensional Dynamic Problem," *Computer Methods in Applied Mechanics and Engineering*, vol. 49, no. 1, pp. 55–70, 1985.
- [2] Y. Wu, B. Hamroun, Y. Le Gorrec, and B. Maschke, "Reduced order controller design for Timoshenko beam: A port Hamiltonian approach," *IFAC PapersOnLine*, vol. 50, no. 1, pp. 7121–7126, 2017.
- [3] E. F. Camacho and C. Bordons Alba, *Model Predictive Control*, 2nd ed. Springer-Verlag, 2007.
- [4] M. Diehl, H. Bock, and J. Schlöder, "A real-time iteration scheme for nonlinear optimization in optimal feedback control," *SIAM Journal on Control and Optimization*, vol. 43, no. 5, pp. 1714–1736, 2005.
- [5] H. Bock and P. Krämer-Eis, "A Multiple Shooting Method for Numerical Computation of Open and Closed Loop Controls in Nonlinear Systems," in *IFAC Proceedings Volumes*, vol. 17, no. 2, Budapest, July 1984, pp. 411–415.
- [6] S. Haghighat, H. H. T. Liu, and J. R. R. A. Martins, "Model-Predictive Gust Load Alleviation Controller for a Highly Flexible Aircraft," *Journal of Guidance, Control, and Dynamics*, vol. 35, no. 6, pp. 1751–1766, 2012.
- [7] Y. Wang, A. Wynn, and R. Palacios, "Nonlinear Aeroelastic Control of Very Flexible Aircraft Using Model Updating," *Journal of Aircraft*, vol. 55, no. 4, pp. 1551–1563, 2018.
- [8] D. H. Hodges, "Geometrically Exact , Intrinsic Theory for Dynamics of Curved and Twisted Anisotropic Beams," *AIAA Journal*, vol. 41, no. 6, pp. 1131–1137, 2003.
- [9] M. Diehl, H. G. Bock, H. Diedam, and P.-B. Wieber, "Fast Direct Multiple Shooting Algorithms for Optimal Robot Control," in *Fast Motions in Biomechanics and Robotics*, Heidelberg, Germany, 2005.
- [10] A. Wynn, Y. Wang, R. Palacios, and P. J. Goulart, "An energy-preserving description of nonlinear beam vibrations in modal coordinates," *Journal of Sound and Vibration*, vol. 332, no. 21, pp. 5543 – 5558, 2013.
- [11] Q. Wang and W. Yu, "Sensitivity Analysis of Geometrically Exact Beam Theory ( GEBT ) Using the Adjoint Method with Hydra," in *52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Denver, 2011.
- [12] S. Maraniello and R. Palacios, "Optimal vibration control and co-design of very flexible actuated structures," *Journal of Sound and Vibration*, vol. 377, no. 2016, pp. 1–21, 2016.
- [13] A. Hanson, *Visualizing quaternions*. Elsevier, 2005.
- [14] R. Palacios and B. Epureanu, "An Intrinsic Description of the Nonlinear Aeroelasticity of Very Flexible Wings," in *52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. Denver, Colorado: AIAA, Apr. 2011.
- [15] R. Palacios, "Nonlinear normal modes in an intrinsic theory of anisotropic beams," *Journal of Sound and Vibration*, vol. 330, no. 8, pp. 1772–1792, 2011.
- [16] Y. Wang, A. Wynn, and R. Palacios, "Nonlinear Modal Aeroservoelastic Analysis Framework for Flexible Aircraft," *AIAA Journal*, vol. 54, no. 10, pp. 3075–3090, 2016.
- [17] —, "Nonlinear Model Reduction for Aeroelastic Control of Flexible Aircraft Described by Large Finite-Element Models," in *55th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. National Harbor, Maryland: AIAA, Jan. 2014.
- [18] Y. Cao, S. Li, L. Petzold, and R. Serban, "Adjoint Sensitivity Analysis for Differential-Algebraic Equations: the Adjoint DAE System and its Numerical Solution," *Journal of Scientific Computing*, vol. 24, no. 3, pp. 1076–1089, 2003.
- [19] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. Springer, 2006.