Modal-Based Nonlinear Model Predictive Control for 3D Very Flexible Structures

Marc Artola, Andrew Wynn, Rafael Palacios

***Abstract*—In this paper a novel NMPC scheme is derived, which is tailored to the underlying structure of the intrinsic description of geometrically exact nonlinear beams (in which velocities and strains are primary variables). This is an important class of PDE models whose behaviour is fundamental to the performance of flexible structural systems (e.g., wind turbines, High-Altitude Long-Endurance aircraft). Furthermore, this class contains the much-studied Euler-Bernoulli and Timoshenko beam models, but has significant additional complexity (to capture 3D effects and arbitrarily large displacements) and requires explicit computation of rotations in the PDE dynamics to account for orientation-dependent forces such as gravity. A challenge presented by this formulation is that uncontrollable modes necessarily appear in any finite dimensional approximation to the PDE dynamics. We show, however, that an NMPC scheme can be constructed in which the error introduced by the uncontrollable modes can be explicitly controlled. Furthermore, in challenging numerical examples exhibiting considerable deformation and nonlinear effects, it is demonstrated that the asymptotic error can be made insignificant (from a practical perspective) using our NMPC scheme and excellent performance is obtained even when applied to a highly resolved numerical simulation of the PDEs. We also present a generalisation of Kelvin-Voigt damping to the intrinsic description of geometrically-exact beams. Finally, special emphasis is placed on constructing a framework suitable for real-time NMPC control, where the particular structure of the underlying PDEs is exploited to obtain both efficient finite-dimensional models and numerical schemes.**

***Index Terms*—Nonlinear MPC, optimal control, real-time control, adjoint-based sensitivity analysis, flexible structures.**

I. Introduction

HIGHLY flexible structures are challenging to control, since linear models fail to capture their fundamental kinematics. In this paper, we consider geometrically-exact beam theories for three-dimensional slender structures [1,2], in which velocities and strains are primary variables. The governing PDE, described in §II, includes nonlinearities to describe geometric couplings arising from arbitrarily large deformations. Importantly, the linearised form of this description contains both the classical Euler-Bernoulli and Timoshenko flexible beam theories, as discussed in §II-D.

Geometrically nonlinear beam models capture the following features absent from linear descriptions, which typically assume small displacements: (i) The geometric stiffening effect, where structural stiffness is increased due to a reorientation of the stress field, of relevance to slender beams and thin plates which possess higher in-plane than out-of-plane stiffness. (ii) Changes to global inertia properties, due to large deflections: a fundamental consequence is that vibration modes and frequencies around deformed configurations can exhibit significant variations from those around unloaded and undeformed equilibria. (iii) The follower force effect, required to describe forces whose orientation is fixed with respect the local orientation of the structure such as aerodynamic forcing. Such geometrically nonlinear effects arise due to couplings between different beam motions, i.e, rigid-body, axial, bending and twisting deformation, and are therefore of particular importance for fully 3D dynamics (see § V-B). In contrast, much work to-date regarding modelling and control of flexible structures has focused on planar cases [3–5]. 3D scenarios have received less interest, with early works based on uncoupled equations for different transverse motions [6,7]. Geometrically-nonlinear models have only recently attracted more attention [8].

Accurate modelling and control of geometrically nonlinear beams is of significance in aerospace applications, in which increasingly slender, and hence flexible, aircraft wings and wind turbine blades are employed for aerodynamic efficiency. Capturing fluid-structure interactions correctly or predicting changes in rigid-body dynamics due to a highly deformed configuration, is crucial for successful vehicle control. Indeed, the lack of accurate nonlinear models was identified as contributing to the failure of NASA's solar-powered Helios prototype aircraft [9,10].

Model predictive control, which can explicitly employ model nonlinearities [11, ch. 9], presents a compelling approach for the control of highly flexible structures, since all the previously mentioned effects are captured. Another advantage of MPC is the direct handling of state hard constraints, which are typical in many flexible structure application problems (e.g. maximum allowable stresses to alleviate fatigue, maximum allowable displacements to prevent collision of flexible components, etc). However, NMPC has not yet been fully exploited in this setting due to its prohibitive computational cost. Indeed, the classical approach of modelling highly flexible geometrically-nonlinear structures with displacement and rotations as primary variables typically requires models with $\mathcal{O}(10^2)$ states for simple geometries [12] to accurately capture the underlying dynamics. To date, this has restricted MPC approaches to employ sequential linearisation of the discretised state-space representation of the system [13,14].

In this paper, we show instead that accurate reduced-order models of highly-flexible structures, using the intrinsic formulation, can be constructed with significantly fewer ($\mathcal{O}(10)$) states. This opens the door to efficient NMPC: we exploit the

M. Artola, Dr. Wynn and Prof. Palacios are with the Department of Aeronautics, Imperial College London, Exhibition Road, London SW7 2AZ, UK. e-mail: a.wynn@imperial.ac.uk.

model structure inherited from the intrinsic formulation and an adjoint-based sensitivity analysis to implement the *real-time iteration* scheme of [15] at low computational cost. It is shown that this approach outperforms NMPC methods which use Automatic Differentiation for sensitivity computations, such as ACADOS [16] or CasADi [17].

A drawback, however, of using the intrinsic formulation is that rotations appear as derived variables. A sufficiently-accurate description of rotations is required, for example, if forces in a fixed global frame (e.g., gravity) are exerted on the structure. To address this, we introduce a modal-based approximation of rotations which maintains the underlying model structure exploited by our NMPC implementation and does not require a significant increase in model order. A further implication of rotations being secondary variables is that linearly unstabilisable modes are introduced in the state-space representation, making it challenging to apply conventional MPC frameworks [18], where stabilisability of the underlying linearised system is required. However, we propose a new NMPC scheme tailored to this particular *partially stabilisable* structure of modal approximations to the intrinsic beam equations. Specifically, we present LMI conditions for the design of the NMPCs stage cost and terminal constraint. If satisfied, provable bounds are then implied on the asymptotic behaviour of both stabilisable and unstabilisable modes under the closed-loop NMPC dynamics. Furthermore, convergence is observed to be robust when the NMPC scheme is applied to a highly resolved simulation of the full PDEs, even in challenging numerical examples. To our knowledge, no such scheme which takes advantage of the internal structure of geometrically exact nonlinear beam models exists in the literature.

The paper is structured as follows: we first introduce the nonlinear PDEs used to describe dynamics of very flexible structures, explain its relation to classical linear beam models, and show in § II-E how Kelvin-Voigt damping can be naturally included in this formulation. Subsequently, in §III, we introduce the modal-based finite dimensional approximation of the intrinsic beam equations [19], and extend it to efficiently approximate finite rotations, studying the unit-norm preservation properties of different approximation schemes. The implementation of the modal-based optimal control problem framework in a nonlinear model predictive control strategy, potentially implementable in real time, is depicted in §IV. The focus is placed on deriving an alternative MPC setup which exploits the particular structure of the underlying system and constitutes the main theoretical contribution of the paper. The computational improvements facilitated by the analytical sensitivity analysis of the nonlinear systems arising from our modal formulation are studied. Supported by numerical examples in §V, the resulting framework paves the way for the application of a real-time-implementable NMPC scheme to control very flexible structures.

II. Structural model

A. Intrinsic beam equations

The geometrically exact, fully-intrinsic nonlinear beam theory of [2] is used. The dynamics of the structure are described

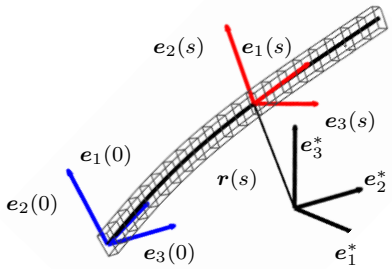


Fig. 1: Fixed frame of reference (black) and local frame of references at $s = 0$ (blue) and at an arbitrary location s (red) in an unconstrained beam configuration. Elastic axis displayed in thick, solid, black line.

by two intrinsic vector states: the vector of linear and angular inertial velocities $\boldsymbol{x}_1(s,t) := [\boldsymbol{v}^\top, \boldsymbol{\omega}^\top]^\top : [0,L] \times \mathbb{R}_+ \to \mathbb{R}^6$ and the vector of force and moment sectional resultants $\boldsymbol{x}_2(s,t) := [\boldsymbol{f}^\top, \boldsymbol{m}^\top]^\top : [0,L] \times \mathbb{R}_+ \to \mathbb{R}^6$, where s is the curvilinear spatial coordinate that defines the elastic axis and t denotes time. The arclength of the undeformed and unloaded structure[1] is L. The equations with no structural damping read

$$M\frac{\partial \boldsymbol{x}_1}{\partial t} - \frac{\partial \boldsymbol{x}_2}{\partial s} - E\boldsymbol{x}_2 + \mathcal{L}_1(\boldsymbol{x}_1)M\boldsymbol{x}_1 + \mathcal{L}_2(\boldsymbol{x}_2)C\boldsymbol{x}_2 = \boldsymbol{f}_e, \quad (1)$$

$$C\frac{\partial \boldsymbol{x}_2}{\partial t} - \frac{\partial \boldsymbol{x}_1}{\partial s} + E^\top \boldsymbol{x}_1 - \mathcal{L}_1^\top(\boldsymbol{x}_1)C\boldsymbol{x}_2 = 0, \quad (2)$$

where all variables and matrices involved are expressed in a local, body-attached, frame of reference ($\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ in Fig. 1), with origin at the beam elastic axis and first component of the basis, $\boldsymbol{e}_1 = [1,0,0]^\top$, tangent to it. Here, $M(s) : \mathbb{R} \to \mathbb{S}^6_{++}$ and $C(s) : \mathbb{R} \to \mathbb{S}^6_{+}$ are the mass and compliance matrices[2], while $E(s) : \mathbb{R} \to \mathbb{R}^{6\times 6}$ is the initial curvature and pre-twist matrix:

$$E = \begin{bmatrix} \widetilde{\boldsymbol{\kappa}}_0 & 0 \\ \widetilde{\boldsymbol{e}}_1 & \widetilde{\boldsymbol{\kappa}}_0 \end{bmatrix}, \quad (3)$$

where $\boldsymbol{\kappa}_0$ is the vector of initial curvatures and pre-twist (for beams with curvature in an unloaded configuration, e.g., arched structures). The linear operators $\mathcal{L}_1 : \mathbb{R}^6 \to \mathbb{R}^{6\times 6}$ and $\mathcal{L}_2 : \mathbb{R}^6 \to \mathbb{R}^{6\times 6}$ are defined as

$$\mathcal{L}_1(\boldsymbol{x}_1) = \begin{bmatrix} \widetilde{\boldsymbol{\omega}} & 0 \\ \widetilde{\boldsymbol{v}} & \widetilde{\boldsymbol{\omega}} \end{bmatrix}, \; \mathcal{L}_2(\boldsymbol{x}_2) = \begin{bmatrix} 0 & \widetilde{\boldsymbol{f}} \\ \widetilde{\boldsymbol{f}} & \widetilde{\boldsymbol{m}} \end{bmatrix}. \quad (4)$$

The matrix operator denoted by the tilde superscript $\widetilde{\boldsymbol{a}}$ is the skew symmetric operator defined as $\widetilde{\boldsymbol{a}}\boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{b}$, for $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$.

The vector $\boldsymbol{f}_e : [0,L] \times \mathbb{R}_+ \to \mathbb{R}^6$ accounts for external forcing, comprising the forces and moments per unit length acting on the structure (e.g., aerodynamic or gravity forces). Natural boundary conditions (i.e, no internal damping and no boundary forcing) for (1) and (2) satisfy

$$x_{1i}(s_\partial, t)x_{2i}(s_\partial, t) = 0, \quad i = 1, \ldots, 6, \quad (5)$$

meaning that either the velocity or force component i can be prescribed to be zero at a boundary $s_\partial \in \{0, L\}$, while

[1] The intrinsic formulation is not limited to a single beam configuration but can be directly applied to multibody configuration and beam networks.

[2] $\mathbb{S}^6_{++}$ and $\mathbb{S}^6_{+}$ denote the set of positive definite and semi-definite symmetric matrices of $\mathbb{R}^{6\times 6}$, respectively.

forced motions or prescribed forces at a boundary are simply enforced as $x_{1i}(s_\partial, t) = \boldsymbol{x}_{1\partial}(t)$ or $x_{2i}(s_\partial, t) = \boldsymbol{x}_{2\partial}(t)$. The total energy (kinetic and potential) of the system is given by

$$\epsilon(t) = \frac{1}{2}\left\langle \boldsymbol{x}_1, M\boldsymbol{x}_1 \right\rangle + \frac{1}{2}\left\langle \boldsymbol{x}_2, C\boldsymbol{x}_2 \right\rangle, \tag{6}$$

where $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \int_0^L \boldsymbol{x}^\top \boldsymbol{y} ds$ denotes the $L^2([0, L], \mathbb{R}^6)$-inner product. If (5) holds, then it is easy to show that the energy satisfies $\dot{\epsilon} = \langle \boldsymbol{x}_1, \boldsymbol{f}_e \rangle$.

B. Intrinsic beam equations in quasi-linear hyperbolic form

An alternative way to write the intrinsic equations (1)–(2), assuming $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are sufficiently smooth, is as second-order quasi-linear hyperbolic partial differential equations, for which results on the convergence of finite-dimensional approximations using the Galerkin Method are readily available in the literature.

We start by defining a new state-space variable

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} M\boldsymbol{x}_1 \\ C\boldsymbol{x}_2 \end{bmatrix} : [0, L] \times \mathbb{R}_+ \to \mathbb{R}^{12}, \tag{7}$$

which allows us to write (1)–(2) in port-Hamiltonian form, similarly to the description of [8],

$$\frac{\partial \boldsymbol{y}}{\partial t}(t, s) = \left(P_1 \frac{\partial}{\partial s} + P_0(\boldsymbol{y}, s) \right) \mathcal{H}(s)\boldsymbol{y}(t, s), \tag{8}$$

with

$$\mathcal{H} = \begin{bmatrix} M^{-1} & 0 \\ 0 & C^{-1} \end{bmatrix}, \qquad P_1 = \begin{bmatrix} 0 & I_6 \\ I_6 & 0 \end{bmatrix} \tag{9}$$

and skew-symmetric nonlinear term

$$P_0(\boldsymbol{y}) = \begin{bmatrix} \mathcal{L}_2(\boldsymbol{y}_1) & \mathcal{L}_1(\boldsymbol{y}_2) + E \\ -\mathcal{L}_1^\top(\boldsymbol{y}_2) - E^\top & 0 \end{bmatrix}. \tag{10}$$

We note that (8) can also be written as a second-order abstract differential equation

$$\frac{\partial^2 \boldsymbol{y}}{\partial t^2} + A\boldsymbol{y} = B(\boldsymbol{y}, \dot{\boldsymbol{y}}). \tag{11}$$

By letting $P_{0,E} = P_0(\boldsymbol{0}, s)$ and $P_{0,y} = P_0(\boldsymbol{y}, s) - P_{0,E}$, the linear operator in (11) is

$$\begin{aligned} A\boldsymbol{y} := &- P_1 \frac{\partial}{\partial s}\left(\mathcal{H} P_1 \frac{\partial}{\partial s}\left(\mathcal{H}\boldsymbol{y} \right) \right) - P_{0,E}\mathcal{H} P_1 \frac{\partial}{\partial s}\left(\mathcal{H}\boldsymbol{y} \right) \\ &- P_1 \frac{\partial}{\partial s}\left(\mathcal{H} P_{0,E} \mathcal{H}\boldsymbol{y} \right) - P_{0,E}\mathcal{H} P_{0,E}\mathcal{H}\boldsymbol{y} \end{aligned} \tag{12}$$

with

$$\begin{aligned} \mathcal{D}(A) := \Bigg\{ &\boldsymbol{y} \in H^2([0, L])^{12} : \\ &[\mathcal{H}\boldsymbol{y}]_i \left[P_1 \mathcal{H} \left(P_1 \frac{\partial}{\partial s}\left(\mathcal{H}\boldsymbol{y} \right) + P_{0,E}\mathcal{H}\boldsymbol{y} \right) \right]_i (s_\partial) = 0 \quad i = 1, \ldots, 12 \Bigg\} \end{aligned} \tag{13}$$

endowed with inner product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \int_0^L \boldsymbol{x}^\top \mathcal{H}\boldsymbol{y}, ds. \tag{14}$$

Boundary conditions in (13) are obtained by differentiating (5) with respect to t and replacing the intrinsic variables $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ by the new state $\boldsymbol{y}$. The nonlinear term reads

$$\begin{aligned} B(\boldsymbol{y}, \dot{\boldsymbol{y}}) := &P_{0,E}\mathcal{H} P_{0,y}\mathcal{H}\boldsymbol{y} + P_{0,y}\mathcal{H} P_{0,E}\mathcal{H}\boldsymbol{y} + P_{0,y}\mathcal{H} P_{0,y}\mathcal{H}\boldsymbol{y} \\ &P_1 \frac{\partial}{\partial s}\left(\mathcal{H} P_{0,y} \right)\mathcal{H}\boldsymbol{y} + \frac{\partial P_{0,y}}{\partial t}\mathcal{H}\boldsymbol{y} + P_{0,y}\mathcal{H} P_1 \frac{\partial}{\partial s}\left(\mathcal{H}\boldsymbol{y} \right). \end{aligned} \tag{15}$$

The results shown below follow from standard arguments directly from the definitions above.

Proposition 1. *Under definitions* (12) *and* (13)*, then*

(i) For any pair $\boldsymbol{y}, \boldsymbol{z} \in \mathcal{D}(A)$, then $\langle A\boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{y}, A^*\boldsymbol{z} \rangle$ and $\mathcal{D}(A) = \mathcal{D}(A^*)$, i.e, operator A is a self-adjoint operator.

(ii) For any $\boldsymbol{y} \in \mathcal{D}(A)$, then $\langle A\boldsymbol{y}, \boldsymbol{y} \rangle \geq 0$, that is, the linear operator A is positive definite.

It is explained in §III that these conditions imply desirable convergence properties of finite dimensional approximations to the underlying PDE (1)–(2).

C. Displacements and rotations

Displacements and rotations appear as secondary variables in the intrinsic formulation and, unless the external forcing $\boldsymbol{f}_e$ depends on them (e.g., gravity), are not needed to solve (1) and (2).

A transformation matrix $T(s, t) : [0, L] \times \mathbb{R}_+ \to \mathbb{R}^{3\times 3}$ from the local to the inertial frame ($\boldsymbol{e}_1^*$, $\boldsymbol{e}_2^*$, $\boldsymbol{e}_3^*$ in Fig. 1) is defined and quaternions are employed to parametrise rotations $\boldsymbol{\xi}(s, t) = [\xi_0, \boldsymbol{\xi}_v^\top]^\top : [0, L] \times \mathbb{R}_+ \to \mathbb{R}^4$, with $\xi_0 \in \mathbb{R}$ and $\boldsymbol{\xi}_v \in \mathbb{R}^3$. They satisfy [20, ch. 26]:

$$\frac{\partial \boldsymbol{\xi}}{\partial s} = \mathcal{U}(\boldsymbol{\kappa} + \boldsymbol{\kappa}_0)\boldsymbol{\xi}, \qquad \frac{\partial \boldsymbol{\xi}}{\partial t} = \mathcal{U}(\boldsymbol{\omega})\boldsymbol{\xi}, \tag{16}$$

where $\boldsymbol{\kappa}$ is the vector of curvatures and twist (i.e, the three last components of the product $C\boldsymbol{x}_2$) and $\mathcal{U}$ is the skew-symmetric operator

$$\mathcal{U}(\boldsymbol{a}) = \frac{1}{2}\begin{bmatrix} 0 & -\boldsymbol{a}^\top \\ \boldsymbol{a} & -\widetilde{\boldsymbol{a}} \end{bmatrix}, \qquad \boldsymbol{a} \in \mathbb{R}^3. \tag{17}$$

Given quaternions $\boldsymbol{\xi}$, the corresponding coordinate transformation matrix is [20, ch. 6]

$$T(\boldsymbol{\xi}) = (1 - 2\left\|\boldsymbol{\xi}_v\right\|^2) I_3 + 2\boldsymbol{\xi}_v\boldsymbol{\xi}_v^\top + 2\xi_0\widetilde{\boldsymbol{\xi}}_v, \tag{18}$$

where I_3 is the identity matrix in $\mathbb{R}^3$. The displacement field $\boldsymbol{r}$ can be computed similarly by solving either of the equations

$$\frac{\partial \boldsymbol{r}}{\partial s} = T(\boldsymbol{e}_1 + \boldsymbol{\gamma}), \qquad \frac{\partial \boldsymbol{r}}{\partial t} = T\boldsymbol{v}, \tag{19}$$

where $\boldsymbol{\gamma}$ is the strain vector (i.e, the three first components of $C\boldsymbol{x}_2$).

D. Intrinsic beam equations and standard formulations

A distinctive characteristic of the intrinsic formulation (1)-(2) is that it is a Hamiltonian description, which extends the structure of rigid-body dynamics into infinite-dimensional systems [19], via the use of derivatives (strains and velocities) as primary variables avoiding use of displacements and rotations (see, for example, the model of a geometrically

nonlinear flexible link in strain/momenta coordinates of [21] or the linear Euler-Bernoulli and Timoshenko beam models in [22]). Indeed, in the case of an infinitely rigid structure (i.e, $C \to 0$ and $\boldsymbol{x}_2 \to \mathbf{0}$) the Euler equations for the motion of an unconstrained rigid-body in 3D are recovered [23].

When an isotropic beam is considered, with its principal axes aligned with the local frame of reference, no elastic couplings exist and the compliance matrix is diagonal $C = \mathrm{diag}(EA, GA_2, GA_3, GJ, EI_2, EI_3)^{-1}$, with E and G denoting, here, the structure's Young and shear moduli (material properties), and A, A_2, A_3, J, I_2 and I_3 are the area, the shear areas, the torsion constant and the second moments of area of the structure's cross-section (geometrical properties). Further, if the sectional centre of mass coincides with the elastic axis[3], it is also the case that no geometrical couplings arise and the mass matrix is diagonal $M = \mathrm{diag}(\mu, \mu, \mu, i_2 + i_3, i_2, i_3)$, where μ is the mass per unit length and i_2 and i_3 are the second moments of inertia of the beam's cross-section. For such uncoupled cases, the classical linear Euler-Bernoulli beam description [24] is recovered by setting both the axial and shear stiffness $EA, GA_3 \to \infty$ and removing all quadratic terms in (1)-(2). The much-studied Timoshenko model [25], is analogously obtained, but with GA_3 finite.

E. Damping

We now show how a generalised Kelvin-Voigt model accounting for internal dissipation in the structure [25] can be considered in the geometrically-nonlinear intrinsic-beam model. In particular, consider the following modification of the constitutive relations

$$\begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{m} \end{bmatrix} = C^{-1} \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\kappa} \end{bmatrix} + C_\tau C^{-1} \frac{\partial}{\partial t} \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\kappa} \end{bmatrix} = \boldsymbol{x}_2 + C_\tau \frac{\partial \boldsymbol{x}_2}{\partial t}. \quad (20)$$

Here, $C_\tau C^{-1} \in \mathbb{S}^6_+$, where C_τ is a matrix with dimensions of time, providing time scales that characterise the damping of the viscoelastic material. Substituting the new constitutive law into (1) and (2), intrinsic beam equations with nonlinear damping are obtained

$$\begin{aligned} M\dot{\boldsymbol{x}}_1 - (\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2)' - E(\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2) + \mathcal{L}_1(\boldsymbol{x}_1) M \boldsymbol{x}_1 + \\ \mathcal{L}_2(\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2) C \boldsymbol{x}_2 = \boldsymbol{f}_e, \end{aligned} \quad (21)$$

$$C\dot{\boldsymbol{x}}_2 - \boldsymbol{x}_1' + E^\top \boldsymbol{x}_1 - \mathcal{L}_1^\top(\boldsymbol{x}_1) C \boldsymbol{x}_2 = \mathbf{0}, \quad (22)$$

where, for brevity, partial derivatives with respect to s and t are denoted by $(\bullet)'$ and $(\dot{\bullet})$, respectively.

Note that the term $C\boldsymbol{x}_2$ in the last nonlinear term of (21) and in (22) is unmodified; it accounts for strains and curvatures, and not forces or moments, and is hence unaffected by the new material model (20). The boundary conditions (5) are replaced by

$$x_{1i}(s_\partial, t) \left[\boldsymbol{x}_2(s_\partial, t) + C_\tau \dot{\boldsymbol{x}}_2(s_\partial, t) \right]_i = 0, \quad i = 1, \ldots, 6, \quad (23)$$

[3]The sectional centre of mass is the point of the beam's cross section in which the gravity resultant force is applied (i.e, depends on the density distribution of the cross section), while the elastic axis is the application point of the resultant elastic forces (i.e, the point in which applied linear forces do not generate torques). For uniform cross sections with two symmetry axes both points coincide, but they do not in U-shaped sections or conventional aircraft wing sections, for instance.

to account for the internal damping affecting the force and moment distributions.

We now describe conditions under which the damping model (20) implies energy dissipativity of the nonlinear structural equations. For the remainder of the section we assume that classical solutions (i.e, that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are absolutely continuous in space and differentiable in time) exist for $t \geq 0$. Well-posedness of (1) will be studied in future work.

Proposition 2. *Suppose that $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is a solution to the intrinsic beam equations* (21)-(23) *with damping defined as in* (20)*, and that $\boldsymbol{f}_e = \mathbf{0}$ and $C_\tau C^{-1} \in \mathbb{S}^6_+$. Then $\dot{\epsilon}(t) \leq 0, t \geq 0$.*

Proof. Differentiating (6) and using (21), (22),

$$\begin{aligned} \dot{\epsilon} &= \langle \boldsymbol{x}_1, M\dot{\boldsymbol{x}}_1 \rangle + \langle \boldsymbol{x}_2, C\dot{\boldsymbol{x}}_2 \rangle \\ &= \left\langle \boldsymbol{x}_1, (\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2)' \right\rangle + \langle \boldsymbol{x}_1, E(\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2) \rangle - \\ &\quad \langle \boldsymbol{x}_1, \mathcal{L}_1(\boldsymbol{x}_1) M \boldsymbol{x}_1 + \mathcal{L}_2(\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2) C \boldsymbol{x}_2 \rangle + \\ &\quad \langle \boldsymbol{x}_2, \boldsymbol{x}_1' \rangle - \left\langle \boldsymbol{x}_2, E^\top \boldsymbol{x}_1 \right\rangle + \left\langle \boldsymbol{x}_2, \mathcal{L}_1^\top(\boldsymbol{x}_1) C \boldsymbol{x}_2 \right\rangle. \end{aligned}$$

By definition of E, $\mathcal{L}_1$ and $\mathcal{L}_2$, it is the case that for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^6$, $\langle \boldsymbol{x}_1, E\boldsymbol{x}_2 \rangle - \left\langle \boldsymbol{x}_2, E^\top \boldsymbol{x}_1 \right\rangle = 0$, $\boldsymbol{x}_1^\top \mathcal{L}_1(\boldsymbol{x}_1) = 0$ and $\boldsymbol{x}_2^\top \mathcal{L}_1^\top(\boldsymbol{x}_1) = \boldsymbol{x}_1^\top \mathcal{L}_2(\boldsymbol{x}_2)$. Consequently,

$$\begin{aligned} \dot{\epsilon} &= \left\langle \boldsymbol{x}_1, (\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2)' \right\rangle + \langle \boldsymbol{x}_1, EC_\tau \dot{\boldsymbol{x}}_2 \rangle - \langle \boldsymbol{x}_1, \mathcal{L}_2(C_\tau \dot{\boldsymbol{x}}_2) C \boldsymbol{x}_2 \rangle \\ &\quad + \langle \boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2, \boldsymbol{x}_1' \rangle - \langle C_\tau \dot{\boldsymbol{x}}_2, \boldsymbol{x}_1' \rangle \\ &= \boldsymbol{x}_1^\top (\boldsymbol{x}_2 + C_\tau \dot{\boldsymbol{x}}_2) \big|_0^L + \langle \boldsymbol{x}_1, EC_\tau \dot{\boldsymbol{x}}_2 \rangle \\ &\quad - \langle \boldsymbol{x}_1, \mathcal{L}_2(C_\tau \dot{\boldsymbol{x}}_2) C \boldsymbol{x}_2 \rangle - \langle C_\tau \dot{\boldsymbol{x}}_2, \boldsymbol{x}_1' \rangle. \end{aligned}$$

By (23), substituting $\boldsymbol{x}_1'$ using (22), noting that the terms involving E vanish again and using the previously employed property of the $\mathcal{L}$ operators gives

$$\dot{\epsilon} = - \langle C_\tau \dot{\boldsymbol{x}}_2, C\dot{\boldsymbol{x}}_2 \rangle \leq 0. \quad (24)$$

□

We now describe a linear version of the damping equations, for which solutions are still energy dissipative, even when coupled with the nonlinear structural model. In particular, taking a first order expansion of the newly obtained damping terms ($C_\tau \dot{\boldsymbol{x}}_2 = C_d C \dot{\boldsymbol{x}}_2 \approx C_d(\boldsymbol{x}_1' - E^\top \boldsymbol{x}_1)$), the nonlinear intrinsic beam equations with linear damping read

$$\begin{aligned} M\dot{\boldsymbol{x}}_1 - \boldsymbol{x}_2' - E\boldsymbol{x}_2 + \mathcal{L}_1(\boldsymbol{x}_1) M \boldsymbol{x}_1 + \mathcal{L}_2(\boldsymbol{x}_2) C \boldsymbol{x}_2 \\ = D_0 \boldsymbol{x}_1 + D_1 \boldsymbol{x}_1' + D_2 \boldsymbol{x}_1'' + \boldsymbol{f}_e, \end{aligned} \quad (25)$$

$$C\dot{\boldsymbol{x}}_2 - \boldsymbol{x}_1' + E^\top \boldsymbol{x}_1 - \mathcal{L}_1^\top(\boldsymbol{x}_1) C \boldsymbol{x}_2 = 0, \quad (26)$$

together with boundary conditions

$$x_{1i}(s_\partial, t) \left[\boldsymbol{x}_2(s_\partial, t) + C_d(s_\partial) \left(\boldsymbol{x}_1'(s_\partial, t) - E^\top \boldsymbol{x}_1(s_\partial, t) \right) \right]_i = 0, \quad (27)$$

where $C_d = C_\tau C^{-1}$ and

$$\begin{aligned} D_0 &= -(C_d E^\top)' - EC_d E^\top, \\ D_1 &= C_d' - C_d E^\top + EC_d, D_2 = C_d. \end{aligned} \quad (28)$$

Proposition 3. *Suppose now that $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is a solution to the intrinsic beam equations with linear damping as defined in* (25)-(28)*, that $C_d \in \mathbb{S}^6_+$ and $\boldsymbol{f}_e = \mathbf{0}$. Then $\dot{\epsilon}(t) \leq 0, t \geq 0$.*

Proof. We start by using the result

$$\dot{\epsilon} = \boldsymbol{x}_1^\top \boldsymbol{x}_2\big|_0^L + \langle \boldsymbol{x}_1, D_0\boldsymbol{x}_1 + D_1\boldsymbol{x}_1' + D_2\boldsymbol{x}_1''\rangle\,, \tag{29}$$

which can be proven as in proposition 2. Now,

$$\boldsymbol{x}_1^\top \boldsymbol{x}_2\big|_0^L = -\,\boldsymbol{x}_1^\top C_d\left(\boldsymbol{x}_1' - E^\top \boldsymbol{x}_1\right)\big|_0^L\,, \tag{30}$$

$$\begin{aligned}\langle \boldsymbol{x}_1, D_0\boldsymbol{x}_1\rangle = &-\left\langle \boldsymbol{x}_1, EC_dE^\top \boldsymbol{x}_1\right\rangle + \left\langle \boldsymbol{x}_1', C_dE^\top \boldsymbol{x}_1\right\rangle + \\ &\left\langle \boldsymbol{x}_1, C_dE^\top \boldsymbol{x}_1'\right\rangle - \boldsymbol{x}_1^\top C_dE^\top \boldsymbol{x}_1\big|_0^L\,,\end{aligned} \tag{31}$$

$$\begin{aligned}\langle \boldsymbol{x}_1, D_1\boldsymbol{x}_1'\rangle = &\langle \boldsymbol{x}_1, C_d'\boldsymbol{x}_1'\rangle - \left\langle \boldsymbol{x}_1, C_dE^\top \boldsymbol{x}_1'\right\rangle + \\ &\langle \boldsymbol{x}_1, EC_d\boldsymbol{x}_1'\rangle\,,\end{aligned} \tag{32}$$

$$\begin{aligned}\langle \boldsymbol{x}_1, D_2\boldsymbol{x}_1''\rangle = &-\langle \boldsymbol{x}_1, C_d'\boldsymbol{x}_1'\rangle - \langle \boldsymbol{x}_1', C_d\boldsymbol{x}_1'\rangle + \\ &\boldsymbol{x}_1^\top C_d\boldsymbol{x}_1'\big|_0^L\,,\end{aligned} \tag{33}$$

where boundary conditions (27) have been used to get (30) and integration by parts has been applied to obtain (31) and (33). Substituting into (29) gives

$$\dot{\epsilon} = -\left\langle \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_1' \end{bmatrix}, \begin{bmatrix} 0 & -EC_d^{\frac{1}{2}} \\ 0 & C_d^{\frac{1}{2}} \end{bmatrix}\begin{bmatrix} 0 & -EC_d^{\frac{1}{2}} \\ 0 & C_d^{\frac{1}{2}} \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_1' \end{bmatrix}\right\rangle \leq 0. \tag{34}$$

□

Remark 4. *By LaSalle's Invariance Principle [26], trajectories converge to the largest invariant set in* $\{\boldsymbol{x}_1, \boldsymbol{x}_2 \mid \dot{\epsilon} = 0\}$. *Under the assumption that trajectories* $\boldsymbol{x}_1(s,t)$ *and* $\boldsymbol{x}_2(s,t)$ *are precompact, Propositions 2 and 3 indicate that solutions to* (21)-(23) *converge to trajectories given by* $\boldsymbol{x}_1' = E^\top \boldsymbol{x}_1 - \mathcal{L}_1^\top(\boldsymbol{x}_1)C\boldsymbol{x}_2$ *while solutions to* (25)-(27) *converge to* $\boldsymbol{x}_1' = E^\top \boldsymbol{x}_1$. *These are rigid-body type steady trajectories (i.e,* $\dot{\boldsymbol{x}}_1 = \dot{\boldsymbol{x}}_2 = \boldsymbol{0}$*) with constant energy* ϵ*, where the structure is allowed to have some remaining stress distribution in the first case (i.e, due to gyroscopic forces appearing from steady rotations) but limited to pure translational (and/or rotation about the longitudinal axis, where no initial curvature is present, i.e,* $\boldsymbol{\kappa}_0 = \boldsymbol{0}$*) rigid-body motions in the second. If one of the boundaries is clamped (i.e,* $\boldsymbol{x}_1(s_\partial) = \boldsymbol{0}$*), then* $\boldsymbol{x}_1 \to \boldsymbol{0}$ *as* $t \to \infty$ *in both cases.*

III. Finite dimensional approximation

A. Modal-based reduced order model

In order to construct a finite-dimensional approximation, the natural modes of the structure are employed [27]. These are the eigenfunctions $\boldsymbol{\phi}_{1j}(s)$, $\boldsymbol{\phi}_{2j}(s) : [0, L] \to \mathbb{R}^6$ of the linearised system (25) and (26) around the unloaded and undeformed configuration (i.e, $\boldsymbol{f}_e = \boldsymbol{x}_1 = \boldsymbol{x}_2 = \boldsymbol{0}$) with no damping ($D_i = 0$). Using N_m eigenfunctions (also referred to as mode shapes), approximate state variables are defined by

$$\boldsymbol{x}_1(s,t) = \boldsymbol{\phi}_{1j}(s)q_{1j}(t), \qquad \boldsymbol{x}_2(s,t) = \boldsymbol{\phi}_{2j}(s)q_{2j}(t), \tag{35}$$

where $q_{1j}(t),\ q_{2j}(t) : \mathbb{R}_+ \to \mathbb{R}$ are the temporal coefficients of the expansion (unless otherwise stated, Einstein's summation convention for index $j = 1, ..., N_m$ 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 (35) into (25) and (26), with full nonlinear terms, pre-multiplying by each eigenfunction and integrating over the spatial domain.

The resulting reduced order model is nonlinear and has the form

$$\dot{\boldsymbol{q}} = W\boldsymbol{q} + N(\boldsymbol{q})\boldsymbol{q} + \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{0} \end{bmatrix}, \tag{36}$$

where the expansion temporal coefficients are gathered in a column vector $\boldsymbol{q}(t) : \mathbb{R}_+ \to \mathbb{R}^{2N_m}$. The modes around the unloaded configuration are orthogonal [23], and, if they are normalised so that $\langle \boldsymbol{\phi}_{1i}, M\boldsymbol{\phi}_{1i}\rangle = \langle \boldsymbol{\phi}_{2i}, C\boldsymbol{\phi}_{2i}\rangle = 1$, the matrix W has the form

$$W = \begin{bmatrix} \Sigma & \Omega \\ -\Omega & 0 \end{bmatrix}, \tag{37}$$

where Ω is a diagonal matrix whose entries are the eigenvalues of the linearised system and Σ is the modal damping matrix

$$[\Sigma]_{ij} = \left\langle \boldsymbol{\phi}_{1i}, D_0\boldsymbol{\phi}_{1j} + D_1\boldsymbol{\phi}_{1j}' + D_2\boldsymbol{\phi}_{1j}''\right\rangle. \tag{38}$$

Matrix $N(\boldsymbol{q})$ is linear in $\boldsymbol{q}$,

$$N(\boldsymbol{q}) = \begin{bmatrix} -q_{1l}\Gamma_1^l & -q_{2l}\Gamma_2^l \\ q_{2l}(\Gamma_2^l)^\top & 0 \end{bmatrix}, \tag{39}$$

with constant coefficients

$$\left[\Gamma_1^l\right]_{ij} = \langle \boldsymbol{\phi}_{1i}, \mathcal{L}_1(\phi_{1j})M\boldsymbol{\phi}_{1l}\rangle\,, \tag{40}$$

$$\left[\Gamma_2^l\right]_{ij} = \langle \boldsymbol{\phi}_{1i}, \mathcal{L}_2(\phi_{2j})C\boldsymbol{\phi}_{2l}\rangle\,. \tag{41}$$

The entries of the forcing term vector are

$$[\boldsymbol{\eta}]_i = \langle \boldsymbol{\phi}_{1i}, \boldsymbol{f}_e\rangle\,, \tag{42}$$

including boundary forcing or control. Full details of the derivation are contained in [19]. The instantaneous energy of the system (6) written in modal coordinates is, simply,

$$\epsilon(t) = \frac{1}{2}\boldsymbol{q}^\top \boldsymbol{q}, \tag{43}$$

with ϵ satisfying $\dot{\epsilon} = \boldsymbol{q}_1^\top \boldsymbol{\eta} + \boldsymbol{q}_1^\top \Sigma \boldsymbol{q}_1$.

Remark 5. *Given that the linear operator of the intrinsic beam equations* (12) *is self-adjoint and positive, and under mild assumptions on the nonlinear operator* (15)*, convergence and bounds on approximation errors of finite-dimensional systems based on the Galerkin Method, as the one we introduce to get to* (36)*, can be established [28].*

Namely, for a modal-based approximation such as ours, where the solution function basis is built up using the eigenfunctions of the linear operator A, the error is bounded by the inverse square root of the first truncated eigenvalue after approximation (35)*.*

Numerical evidence on the convergence of our modal-based finite-dimensional systems has been previously addressed in [29]. This evidence suggests that the first few modes of the structure are usually enough to provide successful control, since the main nonlinear effects that drive the dynamics are well-captured. As previously shown, the fact that the modes are orthogonal allows a very good approximation to the nonlinear terms of the full PDEs (25) and (26), which are all of cross-product nature. This superior approximation

capability is clearly demonstrated when the NMPC controller is consistently able to stabilise a highly resolved system in each of the tested examples in section V-B.

B. Rotation field approximation

If the local orientation of the structure is required, the evolution equation for the quaternions (16) needs to be added to equations (25) and (26), since rotations are not part of the primary variables. Taking gravity as an example of an orientation-dependent forcing term, its contribution to the momentum balance equation (25) reads

$$\boldsymbol{f}_g(s,t) = \mu(s) \begin{bmatrix} I_3 \\ \widetilde{\boldsymbol{r}}_{cm}(s) \end{bmatrix} T(s,t)^\top \boldsymbol{g}, \tag{44}$$

where $\boldsymbol{r}_{cm}$ is the offset distance between the sectional centre of mass and the elastic axis ($\boldsymbol{r}_{cm}$ is a fixed parameter and depends on the beam's cross sectional properties), which produces a torque about the latter. The vector $\boldsymbol{g}$ is the gravity acceleration expressed in the global frame of reference.

A finite-dimensional approximation to the quaternions is required if (36) is to be employed. A consequence of using quaternions to parametrise rotations is that $\|\boldsymbol{\xi}\| = 1$ is necessary for T to be orthogonal [20, ch. 6]. Hence, it is desirable that the employed approximations are unit-norm preserving. An interpolation scheme using a Catmull-Rom spline has already been shown to accurately reproduce arbitrarily large rotations [29]. However, this approach might introduce a large number of extra state variables, easily exceeding $\dim(\boldsymbol{q})$. An alternative which might be more suitable for models constructed with a low number of modes is one that uses modal expansions for the quaternions, which results in lower-order approximate models appropriate for NMPC. Both approaches are now described.

1) Spline interpolation: The time-evolution of quaternions is tracked at certain spatial points along the 1D structure and a Catmull-Rom interpolation scheme is constructed between these points [29]

$$\boldsymbol{\xi}(s,t) = S(s)\boldsymbol{\xi}(t), \tag{45}$$

where $\boldsymbol{\xi}(t) = [\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_{N_\xi+1}^\top]^\top : \mathbb{R}_+ \to \mathbb{R}^{4(N_\xi+1)}$ gathers the set of quaternions at the nodes of a uniform spatial grid with N_ξ intervals. The matrix $S(s) \in \mathbb{R}^{4\times 4(N_\xi+1)}$ contains cubic piece-wise polynomials ensuring smooth matching ($C^1([0,L])$) of all the tracked quaternions.

With this approach, the resulting system of equations reads

$$\dot{\boldsymbol{q}} = W\boldsymbol{q} + N(\boldsymbol{q})\boldsymbol{q} + \begin{bmatrix} \boldsymbol{\eta}_g(\boldsymbol{\xi}) \\ \boldsymbol{0} \end{bmatrix}, \tag{46}$$

$$\dot{\boldsymbol{\xi}}_k = \mathcal{U}(\Phi_\omega(s_k)\boldsymbol{q}_1)\boldsymbol{\xi}_k, \quad k = 1, \ldots, N_\xi + 1, \tag{47}$$

where $\Phi_\omega(s_k) \in \mathbb{R}^{3\times N_m}$ is the horizontal concatenation of the angular velocity components of velocity modes $\boldsymbol{\phi}_{1j}$, evaluated at grid points s_k, and $\boldsymbol{\eta}_g$ in (46) reads

$$\boldsymbol{\eta}_g(\boldsymbol{\xi}) = \boldsymbol{\eta}_{g_0} + \xi_l \Gamma_g^l \boldsymbol{\xi}, \tag{48}$$

with constant coefficient matrices

$$\left[\Gamma_g^l\right]_{ij} = \left\langle \boldsymbol{\phi}_{1i}, \mu \begin{bmatrix} I_3 \\ \widetilde{\boldsymbol{r}}_{cm} \end{bmatrix} T_{jl}(S_j, S_l)^\top \boldsymbol{g} \right\rangle, \tag{49}$$

where S_j is the jth column of $S(s)$ and

$$T_{ab}(\boldsymbol{\xi}^{(a)}, \boldsymbol{\xi}^{(b)}) = -2\boldsymbol{\xi}_v^{(a)\top}\boldsymbol{\xi}_v^{(b)} I_3 + 2\boldsymbol{\xi}_v^{(a)}\boldsymbol{\xi}_v^{(b)\top} + 2\xi_0^{(a)}\widetilde{\boldsymbol{\xi}_v^{(b)}}. \tag{50}$$

The constant term $\boldsymbol{\eta}_{g0}$ can be obtained by setting $T_{jl} = I_3$ in (49).

Remark 6. *The accuracy of the spline approach depends on the refinement of the spatial grid used, with an absolute error of the order $\mathcal{O}(h_\xi^3)$ [30], where $h_\xi = L/N_\xi$ is the grid interval length. Errors are kept within (time) integration accuracy at grid points. The Nyquist criterion, based on the spatial frequency of the angular velocity modes, provides a lower bound on the number of grid points to choose for the error estimate $\mathcal{O}(h_\xi^3)$ to hold valid.*

2) Modal-based expansion: An alternative approximation, analogous to the construction of (36), involves a modal expansion of the quaternions:

$$\boldsymbol{\xi}(s,t) = \bar{\boldsymbol{\xi}}(s) + \boldsymbol{\phi}_{\xi j}(s) q_{\xi j}(t), \tag{51}$$

for chosen quaternion mode shapes $\phi_{\xi j} : [0, L] \to \mathbb{R}^4$, with coefficients $q_{\xi j} : \mathbb{R}_+ \to \mathbb{R}$, and where $\bar{\boldsymbol{\xi}}(s)$ defines the undeformed rotation field, that is, $\bar{\boldsymbol{\xi}}' = \mathcal{U}(\boldsymbol{\kappa}_0)\bar{\boldsymbol{\xi}}$.

We now define mode shapes $\boldsymbol{\phi}_{\xi j}$ to be the eigenfunctions of the linearised time evolution equation (16) around the unloaded and undeformed condition,

$$\Delta\dot{\boldsymbol{\xi}} = \mathcal{U}_\omega(\bar{\boldsymbol{\xi}})\Delta\boldsymbol{\omega}, \tag{52}$$

where

$$\mathcal{U}_\omega(\bar{\boldsymbol{\xi}}) = \frac{\partial}{\partial\omega}\left(\mathcal{U}(\boldsymbol{\omega})\bar{\boldsymbol{\xi}}\right) = \begin{bmatrix} -\bar{\boldsymbol{\xi}}_v^\top \\ \bar{\boldsymbol{\xi}}_0 I_3 + \widetilde{\bar{\boldsymbol{\xi}}}_v \end{bmatrix}. \tag{53}$$

Finally, mode shapes for the quaternions are obtained from (52) using the velocity modes $\boldsymbol{\phi}_{1j}$

$$\boldsymbol{\phi}_{\xi j} = \mathcal{U}_\omega(\bar{\boldsymbol{\xi}})\boldsymbol{\phi}_{\omega j}, \tag{54}$$

where $\boldsymbol{\phi}_{\omega j} = \Pi_\omega \boldsymbol{\phi}_{1j}$, and $\Pi_\omega = [0_3, I_3]$ is a projection operator used to retrieve the angular velocity components. However, a more favourable approximation, concerning unit-norm preservation, is obtained if an *augmented* modal expansion is used, that is, one in which $\bar{\boldsymbol{\xi}}$ is removed from (51) and employed as a shifting equilibrium mode $\boldsymbol{\phi}_{\xi 0} = \bar{\boldsymbol{\xi}}$:

$$\boldsymbol{\xi}(s,t) = \boldsymbol{\phi}_{\xi j}(s) q_{\xi j}(t), \ j = 0, 1, \ldots, N_m. \tag{55}$$

We construct a time evolution system for $\boldsymbol{q}_\xi$ applying a Galerkin projection to the time evolution equation (16) using mode shapes (54) and $\boldsymbol{\phi}_{\xi 0} = \bar{\boldsymbol{\xi}}$:

$$A_\xi \dot{\boldsymbol{q}}_\xi = U_\xi(\boldsymbol{\omega})\boldsymbol{q}_\xi, \tag{56}$$

where A_ξ and $\mathcal{U}_\xi(t)$ are matrices independent of the spatial coordinate s

$$[A_\xi]_{ij} = \left\langle \boldsymbol{\phi}_{\xi i}(s), \boldsymbol{\phi}_{\xi j}(s) \right\rangle, \tag{57}$$

$$[U_\xi(\boldsymbol{\omega})]_{ij} = \left\langle \boldsymbol{\phi}_{\xi i}(s), \mathcal{U}(\boldsymbol{\omega}(s,t))\boldsymbol{\phi}_{\xi j}(s) \right\rangle. \tag{58}$$

Here, the notation $\langle\cdot,\cdot\rangle$ is used to denote the L^2-inner product in $L^2([0,L],\mathbb{R}^4)$. Note that $U_\xi = -U_\xi^\top$ for all t, since the skew-symmetry of $\mathcal{U}$ is inherited, and $A_\xi = A_\xi^\top$. Now, we present the following unit-norm preservation property of such approximation.

Proposition 7. *Suppose that $\boldsymbol{q}_\xi$ is a solution to the time evolution equation* (56) *and initial conditions $\|\boldsymbol{\xi}(s,0)\|_2^2 = 1$ for all s, then $\int_0^L \|\boldsymbol{\xi}(s,t)\|_2^2\, ds/L = 1$, $t \geq 0$. That is, the space-average unit-norm is preserved*[4].

Proof. First, we use expansion (55) to compute

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\xi}(s,t)\|_2^2 = \boldsymbol{q}_\xi(t)^\top \Phi_\xi(s)^\top \Phi_\xi(s)\dot{\boldsymbol{q}}_\xi(t), \tag{59}$$

where $\Phi_\xi(s)$ is a matrix formed by the horizontal concatenation of all modes shapes $\boldsymbol{\phi}_{\xi j}$.

Next, using (56)

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{\xi}(s,t)\|_2^2 = \boldsymbol{q}_\xi^\top \Phi_\xi^\top \Phi_\xi A_\xi^{-1} U_\xi(\boldsymbol{\omega})\boldsymbol{q}_\xi. \tag{60}$$

Integrating (60) along space we get

$$\begin{aligned}\frac{1}{L}\int_0^L \frac{d}{dt}\|\boldsymbol{\xi}(s,t)\|_2^2\, ds &= \frac{1}{L}\int_0^L 2\boldsymbol{q}_\xi^\top \Phi_\xi^\top \Phi_\xi A_\xi^{-1} U_\xi(\boldsymbol{\omega})\boldsymbol{q}_\xi ds \\ &= \frac{2}{L}\boldsymbol{q}_\xi^\top A_\xi A_\xi^{-1} U_\xi(\boldsymbol{\omega})\boldsymbol{q}_\xi = 0,\end{aligned}$$

where the definition (57) and the skew-symmetry property of U_ξ have been used. □

With this approach, the forcing term (48) is written now in terms of the quaternion modal coefficients $\boldsymbol{\eta}_g = \boldsymbol{\eta}_{g_0} + q_{\xi l}\Gamma_g^l \boldsymbol{q}_\xi$, where

$$\left[\Gamma_g^l\right]_{ij} = \left\langle \boldsymbol{\phi}_{1i}, \mu \begin{bmatrix} I_3 \\ \widetilde{\boldsymbol{r}}_{cm} \end{bmatrix} T_{jl}(\boldsymbol{\phi}_{\xi j}, \boldsymbol{\phi}_{\xi l})^\top \boldsymbol{g} \right\rangle, \tag{61}$$

and the evolution equation for the quaternion approximation is

$$\dot{\boldsymbol{q}}_\xi = N_\xi(\boldsymbol{q}_1)\boldsymbol{q}_\xi, \tag{62}$$

where

$$N_\xi(\boldsymbol{q}_1) = q_{1l}\Gamma_\xi^l, \tag{63}$$

$$\Gamma_\xi^l = A_\xi^{-1} U_\xi(\boldsymbol{\phi}_{\omega l}). \tag{64}$$

Note that a system where expansion (51) is used can always be recovered by imposing $q_{\xi 0}(t) = 1$, achieved by zeroing the first row of $N_\xi(\boldsymbol{q}_\xi)$ in (64), which breaks the skew-symmetry and the average unit-norm preservation property is hence lost.

3) Augmented system: The nonlinear, modal-based finite-dimensional approximation to the intrinsic equations (36) coupled with the quaternion approximation is described by the following augmented system

$$\dot{\boldsymbol{q}}_a = \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix}\boldsymbol{q}_a + \begin{bmatrix} N(\boldsymbol{q}) & \eta_g(\boldsymbol{q}_\xi) \\ 0 & N_\xi(\boldsymbol{q}_1) \end{bmatrix}\boldsymbol{q}_a + \begin{bmatrix} \boldsymbol{\eta}_{g_0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}. \tag{65}$$

Here, $\boldsymbol{q}_a$ refers to the augmented state vector $\boldsymbol{q}_a = [\boldsymbol{q}_1^\top, \boldsymbol{q}_2^\top, \boldsymbol{\xi}^\top]^\top : \mathbb{R}_+ \to \mathbb{R}^{2N_m+4(N_\xi+1)}$, if the *spline* method is chosen, or $\boldsymbol{q}_a = [\boldsymbol{q}_1^\top, \boldsymbol{q}_2^\top, \boldsymbol{q}_\xi^\top]^\top : \mathbb{R}_+ \to \mathbb{R}^{3N_m+1}$, if the *modal* approximation for the quaternions is employed. Note that an equivalent matricial expression $\dot{\boldsymbol{\xi}} = N_\xi(\boldsymbol{q}_1)\boldsymbol{\xi} = q_{1l}\Gamma_\xi^l\boldsymbol{\xi}$ to (47) can always be found, meaning that both approximations have a similar (quadratic) structure and hence we express the coupled system generically by (65).

[4]Note that, to achieve $\|\boldsymbol{\xi}(s,0)\|_2^2 = 1$ with the proposed basis (55), the time coefficient corresponding to the shifting equilibrium satisfies $q_{\xi 0}(0) = \sqrt{L}$ (where the modes have been conveniently normalised so that $[A_\xi]_{ii} = 1$)

IV. Nonlinear MPC setup

A. Nonlinear optimal control problem

We follow the conventional discrete-time MPC formulation of [18] for (state and input) constrained nonlinear systems, where full state feedback and no disturbances are assumed. Specifically, we solve the following nonlinear optimal control problem over a time horizon τ_p at each sampling point t_i, where the prediction horizon is split into N intervals,

$$\begin{aligned}
\min_{\boldsymbol{q}_a^k, \boldsymbol{u}^k} \quad & \sum_{k=0}^{k=N-1} \frac{1}{2}\hat{\boldsymbol{q}}_a^{k\top} Q \hat{\boldsymbol{q}}_a^k + \frac{1}{2}\hat{\boldsymbol{u}}^{k\top} R \hat{\boldsymbol{u}}^k \\
& + \frac{1}{2}\hat{\boldsymbol{q}}_a^{N\top} P \hat{\boldsymbol{q}}_a^N && \text{(66a)} \\
\text{s.t.} \quad & \boldsymbol{q}_a^0 = \boldsymbol{q}_{a0}, && \text{(66b)} \\
& \boldsymbol{q}_a^{k+1} = \boldsymbol{f}(\boldsymbol{q}_a^k, \boldsymbol{u}^k), \qquad k = 0,\ldots,N-2, && \text{(66c)} \\
& \boldsymbol{u}_l \leq \hat{\boldsymbol{u}}^k \leq \boldsymbol{u}_u, \qquad k = 0,\ldots,N-1, && \text{(66d)} \\
& \frac{1}{2}\hat{\boldsymbol{q}}_a^{N\top} P \hat{\boldsymbol{q}}_a^N \leq a, && \text{(66e)}
\end{aligned}$$

where $Q > 0$ and $R > 0$ and the perturbations $\hat{\boldsymbol{q}}_a = \boldsymbol{q}_a - \bar{\boldsymbol{q}}_a$ and $\hat{\boldsymbol{u}} = \boldsymbol{u} - \bar{\boldsymbol{u}}$ refer to a reference state $\bar{\boldsymbol{q}}_a$ and control $\bar{\boldsymbol{u}}$. As seen, the optimal control problem is subject to initial conditions (66b) and nonlinear, discrete-time evolution equation (66c), obtained by Runge-Kutta 4$^{\text{th}}$ order integration of (65). If Q is set equal to the identity, the control can be interpreted to act via energy-shaping, since the state penalty is the perturbed energy of the system.

Two important challenges arise from the application of NMPC to nonlinear systems, describing very flexible structures, such as (65). On the one hand, the smallest system that typically results from describing the motion of a very flexible structure in three dimensions (see §V) is of the order of $\mathcal{O}(50)$. To avoid solving an NLP of impractical size (due to the fact that the numerical integration time step of (65) may be required to be very small, since it is linked to the highest frequency of the retained modes) the discrete-time system (66c) is constructed using several (typically ranging from tens to hundreds) time steps of RK4 integration to maintain the number of MPC sub-intervals to computationally tractable $N \sim \mathcal{O}(10)$. This is in contrast to typical MPC setups, where one (or very few) time steps of a chosen time-integrator [16] are employed to construct (66c).

The second challenge concerns the addition of rotations as states of our system, which are required to evaluate orientation-dependent forcing terms such as gravity. This introduces a set of linearly unstabilisable states, creating a control challenge and necessitates alternative NMPC stability results to standard approaches [18]. Linearisation of (65) around an equilibrium point, where $\boldsymbol{q}_1 = \boldsymbol{0}$, transforms the perturbation equation for the quaternions into pure integration

of the angular velocities, expressed by means of $\boldsymbol{q}_1$, causing rank deficiency of the linearised state matrix. Since this is a recurrent problem encountered every time rotations are included as part of the state-space, we formulate a specific NMPC design strategy with associated convergence results in the following section. Its performance is highlighted with numerical demonstrations in §V-B, where the controller is also shown to perform remarkably well when applied to a high-order finite-element discretisation of the full PDEs.

B. NMPC convergence for partially unstabilisable systems arising from geometrically nonlinear beam dynamics

We consider discrete time systems of the form

$$\boldsymbol{x}^{k+1} = \boldsymbol{f}(\boldsymbol{x}^k, \boldsymbol{u}^k) = A\boldsymbol{x}^k + B\boldsymbol{u}^k + \tilde{\boldsymbol{f}}(\boldsymbol{x}^k, \boldsymbol{u}^k) \tag{67}$$

where $\boldsymbol{x}^k \in \mathbb{R}^n$ and $\boldsymbol{u}^k \in \mathbb{R}^m$. The linear part of the system is governed by matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ and the function $\tilde{f} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ defines quadratic and higher-order terms of the nonlinear dynamics.

System (67) has the particular form

$$A = \begin{bmatrix} A_w & A_d \\ 0 & I_{n_2} \end{bmatrix}, \quad B = \begin{bmatrix} B_w \\ 0 \end{bmatrix}, \tag{68}$$

corresponding to a decomposition of state $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{w}^\top & \boldsymbol{d}^\top \end{bmatrix}^\top$ into stabilisable $\boldsymbol{w}$ and unstabilisable $\boldsymbol{d}$ modes. In particular $A_w \in \mathbb{R}^{n_1\times n_1}$ and $B_w \in \mathbb{R}^{n_1\times m}$ form a stabilisable pair and $\boldsymbol{w} \in \mathbb{R}^{n_1}$, while $A_d \in \mathbb{R}^{n_1\times n_2}$ and $\boldsymbol{d} \in \mathbb{R}^{n_2}$. A system with this structure is obtained after applying a staircase and a Schur decomposition to the linearised discrete-time system (66c) (constructed using several RK4 integration steps of the continuous-time nonlinear system (65)) about a reference, equilibrium point $(\bar{\boldsymbol{q}}, \bar{\boldsymbol{u}})$ of (65). Then, $\boldsymbol{x}$ in (67) defines perturbations about that reference point.

To implement a model predictive control algorithm to exploit the particular structure of (67), let $P_w \in \mathbb{S}^{n_1}_{++}, P_d \in \mathbb{S}^{n_2}_{++}, R \in \mathbb{S}^m_{++}$ and $\beta > 0$. Define a stage $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ and a terminal $V_f : \mathbb{R}^n \to \mathbb{R}_+$ cost by

$$\ell(\boldsymbol{x}, \boldsymbol{u}) := \frac{\beta}{4}\|\boldsymbol{w}\|^2_{P_w} + \|\boldsymbol{u}\|^2_R, \quad V_f(\boldsymbol{x}) := \|\boldsymbol{w}\|^2_{P_w}, \quad x = \left(\begin{smallmatrix} \boldsymbol{w} \\ \boldsymbol{d} \end{smallmatrix}\right). \tag{69}$$

The MPC cost function that we employ can now be defined.

Definition 8. *Let $N \in \mathbb{N}$ and $U \subseteq \mathbb{R}^m$ be the set of admissible control inputs. Then the MPC cost function $V_N : \mathbb{R}^n \times U^N \to \mathbb{R}$ is given by*

$$V_N(\boldsymbol{x}, \boldsymbol{u}) := \sum_{k=0}^{N-1} \ell(\boldsymbol{x}^k, \boldsymbol{u}^k) + V_f(\boldsymbol{x}^N),$$

where $\boldsymbol{x}^0 = \boldsymbol{x}$ and $\boldsymbol{x}^k = \boldsymbol{\phi}(k; \boldsymbol{x}, \boldsymbol{u})$. Here, $\boldsymbol{\phi}(k; \boldsymbol{x}, \boldsymbol{u})$ denotes the state of the system after k steps under the dynamics $\boldsymbol{x}^{k+1} = \boldsymbol{f}(\boldsymbol{x}^k, \boldsymbol{u}^k)$.

We will optimize the MPC cost function over input sequences that drive the state, in N steps, to compact subsets of the form

$$\chi(\omega, \delta) := \left\{ \boldsymbol{x} = \left(\begin{smallmatrix} \boldsymbol{w} \\ \boldsymbol{d} \end{smallmatrix}\right) \in \mathbb{R}^n : \|\boldsymbol{w}\|^2_{P_w} \le \omega,\ \|\boldsymbol{d}\|^2_{P_d} \le \delta \right\},$$

where $\omega, \delta > 0$ will be chosen later. To this end, denoting $\boldsymbol{\phi}_{N,\boldsymbol{x}}(\boldsymbol{u}) = \phi(N; \boldsymbol{x}, \boldsymbol{u}) : U^N \to \mathbb{R}^n$, we define for each $\boldsymbol{x} \in \mathbb{R}^n$ the set (possibly empty) of control sequences that can drive the system to $\chi(\omega, \delta)$ by

$$\mathcal{U}_{N,\omega,\delta}(\boldsymbol{x}) = \phi^{-1}_{N,\boldsymbol{x}}(\chi(\omega, \delta)), \qquad \boldsymbol{x} \in \mathbb{R}^n,$$

and the set of states that can reach $\chi(\omega, \delta)$ in N steps by $X_{N,\omega,\delta} = \{\boldsymbol{x} \in \mathbb{R}^n : \mathcal{U}_{N,\omega,\delta}(\boldsymbol{x}) \neq \emptyset\}$.

Assumption 9. *For any $N \in \mathbb{N}$, there exists $\omega_0, \delta_0 > 0$ such that:*

(i) (Invariance) For any $0 < \omega < \omega_0$ and $0 < \delta < \delta_0$, the sets $\chi(\omega, \delta)$ are control invariant.

(ii) (Regularity) For any $\epsilon > 0$, there exists $\eta > 0$ such that for any $\omega < \omega_0, \delta < \delta_0$ and $\boldsymbol{x} \in X_{N,\omega,\delta}$,

$$d_H\left(\mathcal{U}_{N,\omega,\delta}(\boldsymbol{x}), \mathcal{U}_{N,\omega,\delta+\eta}(\boldsymbol{x})\right) \le \epsilon.$$

Here, $d_H(\cdot,\cdot)$ is the Hausdorff distance between compact subsets.

Condition (i) states that for any $\boldsymbol{x} \in \chi(\omega, \delta)$, there exists $\boldsymbol{u} \in U$ such that $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \in \chi(\omega, \delta)$. Consequently, $\mathcal{U}_{N,\omega,\delta}(\boldsymbol{x}) \neq \emptyset$ for any $\boldsymbol{x} \in \chi(\omega, \delta)$ with $\omega < \omega_0, \delta < \delta_0$ and $X_{N,\omega,\delta} \neq \emptyset$. Hence, there exist initial states from which $\chi(\omega, \delta)$ can be reached in N steps. Condition (ii) is a mild regularity condition stating that a small relaxation in the terminal constraint corresponds to a small increase in the set of admissible control inputs.

Finally, we define the optimal cost functions

$$V^*_{N,\omega,\delta}(\boldsymbol{x}) := \min_{\boldsymbol{u} \in \mathcal{U}_{N,\omega,\delta}(\boldsymbol{x})} V_N(\boldsymbol{x}, \boldsymbol{u}), \qquad \boldsymbol{x} \in X_{N,\omega,\delta}.$$

and denote the first value of an optimising input trajectory $\boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x})$, meaning that the closed-loop MPC system evolves according to $\boldsymbol{x}^+ = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x}))$.

1) Design of weighting matrices: Design of the stage cost ℓ, terminal cost V_f, and terminal constraint set $\chi(\omega, \delta)$ relies on appropriate construction, namely conditions (70), (71), of the matrices P_w, P_d and R. The following Lemma implies that appropriate choices always exist and can be constructed by solving LMIs.

Lemma 10. *Let A_w, A_d and B_w be the system matrices from (68). Then there exist $P_w \in \mathbb{S}^{n_1}_{++}, P_d \in \mathbb{S}^{n_2}_{++}, R \in \mathbb{S}^m_{++}$, a constant $0 < \beta < 1$ and control gain matrices $K_w \in \mathbb{R}^{m\times n_1}, K_d \in \mathbb{R}^{m\times n_2}$ such that*

$$X = \begin{bmatrix} \begin{bmatrix} (1-\beta)P_w^{-1} & 0 \\ 0 & P_d^{-1} \end{bmatrix} & \begin{bmatrix} P_w^{-1}(A_w + B_wK_w)^\top & 0 \\ P_d^{-1}(A_d + B_wK_d)^\top & 0 \end{bmatrix} \\ [*] & \begin{bmatrix} P_w^{-1} & 0 \\ 0 & P_d^{-1} \end{bmatrix} \end{bmatrix} \succeq 0, \tag{70}$$

and

$$Y = \begin{bmatrix} \begin{bmatrix} \frac{\beta}{4}P_w^{-1} & 0 \\ 0 & \frac{1}{2}P_d^{-1} \end{bmatrix} & \begin{bmatrix} P_w^{-1}K_w^\top \\ P_d^{-1}K_d^\top \end{bmatrix} \\ [*] & R^{-1} \end{bmatrix} \succeq 0. \tag{71}$$

Proof. See Appendix. □

Note that (70) and (71) can be imposed as LMI conditions for fixed $0 < \beta < 1$. The importance of these conditions is

that they guarantee existence of a linear controller, namely $u = K_w \boldsymbol{w} + K_d \boldsymbol{d}$, under which the final stage cost $V_f(\boldsymbol{x}) = \|\boldsymbol{w}\|^2_{P_w}$ can be appropriately controlled.

2) Application to MPC algorithm: We are now in position to present the main result of the paper, which shows that the MPC trajectory converges to a known compact subset of state space.

Theorem 11. *Suppose that Assumption 9 holds and* (70), (71) *are satisfied. Let* $(\boldsymbol{x}^k)_{k\geq 0}$ *denote the closed-loop MPC trajectory* $\boldsymbol{x}^{k+1} = f(\boldsymbol{x}^k, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x}))$ *and let* $\epsilon > 0$. *Then there exists* $\omega > 0, \delta > 0$ *such that whenever* $\boldsymbol{x}^0 \in X_{N,\omega,\delta}$, *it follows that*

$$\boldsymbol{x}^{k_0} \in \chi(\omega^*, \delta),$$

for some $k_0 \in \mathbb{N}$, *where* $\omega^* := 4\beta^{-1}(2\delta + \epsilon)$. *Furthermore, if* $\omega^* \leq \omega$,

$$\limsup_{k\to\infty} \|\boldsymbol{w}^k\|^2_{P_w} \leq \frac{4}{\beta}(2\delta + \epsilon) + \frac{4}{\beta}\max\left\{V^*_{N,\omega,\delta}(\boldsymbol{x}) : \boldsymbol{x} \in \chi(\omega^*, \delta)\right\}$$

Proof. See Appendix. □

The practical implication of Theorem 11 is that the LMI conditions (70), (71) can be used to impose as large as possible value of β, in order to improve the final bound. It is shown in §V-B that, in practice, the corresponding value of δ may be very small (compared to β), while maintaining feasibility for a relatively large value of ω. This corresponds to widening the set of states from which the NMPC scheme will converge asymptotically. Moreover, the performance of the controller when applied to a highly resolved simulation of the full PDEs is remarkably similar to the *nominal* case (where there is no plant/model mismatch), highlighting the inherent robustness of the implementation. We note finally, that the bound on the eventual state tends to zero as $\epsilon, \delta \to 0$. That is, if a control exists which can drive the linearly uncontrollable states close to the origin, then the controllable states can also be made arbitrarily small.

C. Numerical implementation

Since the nonlinear optimal problem defined by (66) is already discretised into N sub-intervals, the multiple shooting approach of [31] is used and steps $\Delta\boldsymbol{p}$ of a Sequential Quadratic Programming (SQP) [32, ch. 18] strategy are employed to solve the nonlinear optimal control problem iteratively:

$$\min_{\Delta\boldsymbol{p}} \quad \Delta\boldsymbol{p}^\top \boldsymbol{g}_i + \frac{1}{2}\Delta\boldsymbol{p}^\top H_i \Delta\boldsymbol{p} \tag{72a}$$

$$\text{s.t. } \Delta\boldsymbol{q}_a^0 + \left(\boldsymbol{q}_a^0\right)_i = \boldsymbol{q}_{a_0}, \tag{72b}$$

$$\Delta\boldsymbol{q}_a^{k+1} + \left(\boldsymbol{q}_a^{k+1}\right)_i = \frac{d\boldsymbol{f}(\boldsymbol{q}_a^k, \boldsymbol{u}^k)}{d\boldsymbol{p}^k}\Delta\boldsymbol{p} + \boldsymbol{f}(\boldsymbol{q}_a^k, \boldsymbol{u}^k)_i, \tag{72c}$$

$$\boldsymbol{u}_l \leq (\hat{\boldsymbol{u}}^k)_i + \Delta\boldsymbol{u}^k \leq \boldsymbol{u}_u, \tag{72d}$$

$$\left|\left(\frac{1}{2}\hat{\boldsymbol{q}}_a^{N\top} P \hat{\boldsymbol{q}}_a^N\right)_i + \left(\hat{\boldsymbol{q}}_a^N\right)_i^\top P \Delta\boldsymbol{q}_a^N\right| \leq a, \tag{72e}$$

where $\boldsymbol{p}^\top = [\ldots, \boldsymbol{q}_a^k, \ldots, \boldsymbol{u}^m, \ldots]$ is the set of all optimisation parameters and the subscript i denotes the current optimisation iterate. This quadratic program is constructed using the gradient $\boldsymbol{g}_i$ and the Hessian, H_i, of the Laplacian of the constrained minimisation problem, $\mathcal{L} = \mathcal{F} + \boldsymbol{\lambda}^\top \boldsymbol{\mathcal{C}}$, where $\mathcal{F}$ is the cost function (66a) and $\boldsymbol{\mathcal{C}}(\boldsymbol{q}_a^k, \boldsymbol{u}^k) \leq 0$ is the set of constraints (66b)-(66e), with $\boldsymbol{\lambda}$ its corresponding multipliers. Note that linearisation of a positive quadratic constraint such as (66e) requires the use of absolute value in (72e) for consistency. Also, linearisation of (66c) might lead to discontinuous, unfeasible intermediate iterates. The Hessian is approximated by the BFGS update formula.

The use of multiple shooting allows speed up of the implementation, since the computational bottleneck of the methodology (time-integration of (66c) and the evaluation of its sensitivities to construct its linearised version (72c) as required by the SQP approach) can be parallelised. Furthermore, multiple shooting has been shown to improve convergence of NLPs.

In the following section we analyse this computational burden, and we show that analytically deriving the adjoint equations of the continuous-time system to evaluate the sensitivities required to construct (72c) offers a computational advantage over black-box solutions. We will also explore real-time application in §V-B, using the initial-value embedding real-time scheme of [15], where intermediate solutions of the SQP iteration (72) are fed back to the system instead of reaching convergence of (66) at each sampling time.

D. Sensitivity analysis

As seen in the previous section, it becomes paramount to evaluate the derivatives involved in the linearisation of the continuity constraints (72c) as efficiently as possible. Applying the adjoint method [33], they are evaluated by means of

$$\frac{d\boldsymbol{f}\left(\boldsymbol{q}_a^k, \boldsymbol{u}^k\right)}{d\boldsymbol{p}^k} = \begin{bmatrix}\Lambda^\top(t_{k-1}) & 0\end{bmatrix} + \begin{bmatrix}0 & \int_{t_{k-1}}^{t_k} \Lambda^\top \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{u}} dt\end{bmatrix}, \tag{73}$$

where $\boldsymbol{h}$ denotes the continuous-time system (65) and where the matrix of multipliers $\Lambda \in \mathbb{R}^{\dim(\boldsymbol{q}_a^k)\times\dim(\boldsymbol{q}_a^k)}$ is subject to the ODE and final condition

$$\dot{\Lambda} = -\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}_a}^\top \Lambda, \ \Lambda(t_k) = I. \tag{74}$$

Relevant terms to solve the adjoint system (73) are easily obtained due to the compact structure of (36):

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}_a} = \begin{bmatrix} \Sigma - q_{1l}\Gamma_1^l - q_{1j}\Gamma_1^j & \Omega - q_{2l}\Gamma_2^l - q_{2j}\Gamma_2^j & q_{\xi l}\Gamma_g^l + q_{\xi j}\Gamma_g^j \\ -\Omega + q_{2l}(\Gamma_2^\top)^l & q_{1j}(\Gamma_2^\top)^j & 0 \\ q_{\xi l}\Gamma_\xi^l & 0 & q_{1j}\Gamma_\xi^j \end{bmatrix}, \tag{75}$$

where the superscript j in the matrices Γ (differing from the usual notation Γ^l) implies a permutation of indices j and l and subsequent summation over j.

The associated cost of computing all the necessary derivatives to construct (72), which include derivatives of all final states with respect to initial states and control inputs, is roughly constant for the nonlinear equations (i.e, the sensitivity analysis cost remains proportional to the cost of one forward

simulation, independently of the problem's size). Observing the structure of (65), we identify that the computational bottleneck lays with the evaluation of the nonlinear terms $N(\boldsymbol{q})$, $N_\xi(\boldsymbol{q}_\xi)$ and $\eta_g(\boldsymbol{q}_\xi)$. These computations have a complexity, based on the total number of product operations, of the order $\mathcal{O}(N_m^3)$, which predominate over the rest of matrix-vector multiplications which scale with $\mathcal{O}(N_m^2)$. We see that solving the sensitivity analysis (73)-(74) involves the matrix multiplication $\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}_a}^\top \Lambda$ in (74), with a cost of the order $\mathcal{O}(N_m^3)$ (note that usually $\dim(\boldsymbol{u}^k) \ll N_m$). The evaluation of coefficients in (75) has also a complexity $\mathcal{O}(N_m^3)$. Therefore, we can conclude that the cost of solving the adjoint equations and computing the sensitivity analysis is proportional to solving one forward simulation, regardless of the dimension of the problem.

Most of the current available real-time optimal control packages use Automatic Differentiation, which offers machine precision sensitivities. In our particular case, due to the nonlinear terms present in (65), even the cost of the backward mode, where $3N_m+1$ sweeps are necessary, would scale with $\mathcal{O}(N_m^4)$. This reveals the competitive advantage of adjoint-based methods over AD based implementations for this kind of nonlinear systems. Besides, our adjoint equations can be derived analytically in a straightforward manner as it has been shown.

V. Numerical examples

A. Computational complexity of adjoint system

Here, the scaling of the complexity of the adjoint system and the sensitivity analysis relative to the forward simulation is assessed. In order to do so, the execution time of integrating system (65) along a simulation time of 2s of a three dimensional pendulum with gravity (such as case 3 in §V-B) is divided by the total number of time-steps taken (time integration is performed using a RK4 scheme with a time step of 0.25 times the smallest period of the modes included in the expansion, which suffices convergence criteria). Thus, an approximation of the execution time of a single time-integration step is obtained, smoothing possible fluctuations in the performance. This is repeated for a modal basis with a varying size (from 20 modes to 80). The same procedure is applied when solving the complete adjoint system (73)-(74), necessary to linearise the continuity constraints. Execution times correspond to a Matlab script running on an Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz processor.

As observed in figure 2, the execution times per RK4 iteration, of the forward simulation and the adjoint system scale with the number of modes to a power which lays within 2 and 3. As previously shown, the theoretical complexity of both systems scales with N_m^3. Execution time and computational complexity are closely related, though not fully equivalent. The former is highly dependent on the internal architecture and algorithmic functioning of the in-built functions, while the later is assumed to be roughly the number of product operations. Nonetheless, and most importantly, we note that the cost of the adjoint problem and of the forward simulation remain proportional, regardless of the size of the system which proves our size-independence claim. The adjoint problem is, at worst, five times more expensive than the forward one.

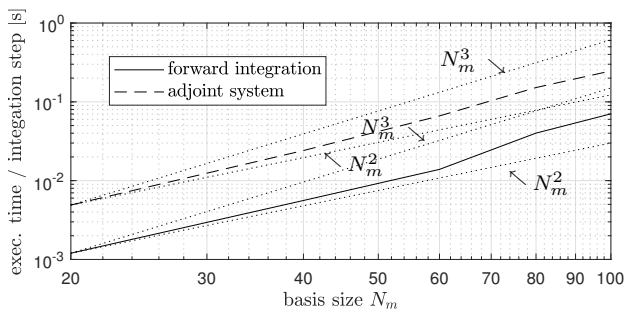


Fig. 2: Execution time per time step of the forward (solid line) and adjoint (dashed line) problems, with dotted lines showing the $+2$ and $+3$ slopes.

B. Nonlinear control of a very flexible inverted pendulum

Three different implementations of the controller are tested: (i) A linear implementation, LMPC, which uses a linearisation of (65) around the reference $\bar{\boldsymbol{q}}_a$, constructing and solving a QP (linear dynamics plus quadratic cost function) at each sampling time, (ii) the converged nonlinear controller, NMPC, and a (iii) real-time iteration, RTI-NMPC, where one SQP iteration (72) is solved at each sampling time, without shifting of the optimisation variables.

For the system plant we test two scenarios. We either use the same system as the internal model of the MPC (referred to as the *nominal* case), so we can make use of standard MPC convergence guarantees and of the results obtained in §IV. However, we also demonstrate that excellent performance is achieved when the NMPC scheme is applied to the *true* system, i.e., the full PDE. The full PDE is implemented using a high-order finite-element discretisation of the intrinsic beam equations (25)–(26), similar to that in [12], and a spline interpolation scheme where rotation information is required, since both discretisations have errors that are known to be bounded by the grid size. All the presented cases assume slight damping with $D_0 = D_1 = 0$ and $D_2 = 10^{-4}\text{diag}(1,0,0,1,1,1)$. Finally, cases 1 and 2 below use results from standard NMPC [18], since the system is linearly stabilisable around the desired equilibrium, while in the more complex case 3, which includes simulation of gravity forces, we put our novel NMPC scheme to the test.

1) 2D clamped case: For the 2D case ($1-3$ plane), a clamped, straight uniform beam of $L = 1$ is considered, with mass and compliance matrices $M = \text{diag}(1,1,1,0,8.33 \cdot 10^{-5},0)$ and $C = \text{diag}(100,\infty,10^5,\infty,0.5,\infty)^{-1}$, all with appropriate units.

The beam is initially in a straight equilibrium position (light grey line in figure 3b) and the controller is required to bring the structure to a deformed position (black line in figure 3b), described by a reference state $\bar{\boldsymbol{q}}$ in (66a). This deformed configuration is the steady solution to a follower transverse force, of magnitude $F_0 = 2$ N, applied at the tip. The controller, however, is set up to act via two degrees of

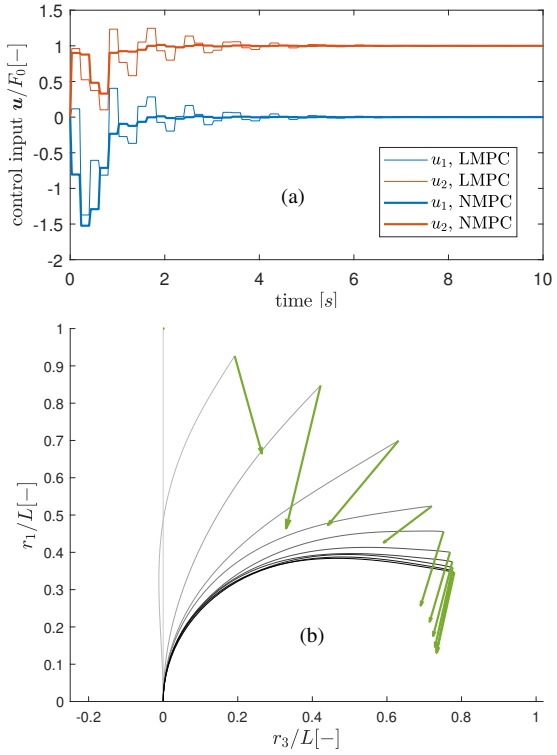


Fig. 3: (a) Control input comparison for the *nominal* case of all the linear and converged MPC controllers and (b) snapshots every 0.2 s of the control action of the NMPC controller (with resultant applied force in green).

freedom $\boldsymbol{u} = [f_1, f_3]^\top$, where f_3 and f_1 are transverse and axial follower point forces:

$$\boldsymbol{x}_2(L,t) = [f_1(t), 0, f_3(t), 0, 0, 0]^\top. \tag{76}$$

All the considered controllers use a basis with $N_m = 4$, including the first two axial and bending modes. The sampling time is $\tau_s = 0.2$ s and the prediction horizon is $\tau_p = 2$ s (0.079 and 0.79 times the period of the first bending mode, respectively), giving $\mathrm{N} = 10$ subintervals. Finally, control saturation (66d) is set as $u_p = -u_l = 2.5F_0$ and the cost function (66a) is constructed with $Q = \mathrm{blkdiag}(I_{N_m}, I_{N_m})$ and control penalty $R = 0.01I_2$.

Regarding the *nominal* case, a terminal penalty weight P has been found and a terminal region (66e) with $a = 1.98$ has been estimated following [18] MPC design for constrained nonlinear systems (since the system is linearly stabilisable around $\bar{\boldsymbol{q}}$). Attending to the obtained results, both linear and nonlinear implementations exhibit fairly similar behaviour, as seen in figure 3a, albeit with LMPC producing a notably higher oscillatory response and settling time.

It is interesting to note that both controllers exploit nonlinear buckling, to bring the structure closer to the reference condition in a shorter time (which effectively decreases the cost function (66a)): the controller applies a compression axial force, depicted by negative values of the blue lines in figure 3a which effectively causes the beam to buckle, as the increase in curvature of the lower section of the beam in figure 3b suggests.

It is also very interesting to look at the case where the controller is applied to the *true* system, shown in figure 4. At first glance, results provided by the LMPC and NMPC appear very similar to the previous case, especially for NMPC. However, for the LMPC implementation the oscillations already present in the previous case are further amplified, due to the addition of uncertainty corresponding to the lack of knowledge of the nonlinear effects. From this numerical exploration, it is implied that, firstly, the NMPC implementation appears robust to plant uncertainty, and secondly, that just a few modes (2 modes per motion) suffice to control the actual system. We also note that the RTI implementation opts for a smoother and somewhat slower response which consists of applying a transverse force progressively, maintaining the control axial input close to zero.

2) 3D clamped case: In the three-dimensional case, a similar problem has been tested. Again, a clamped, straight uniform beam of the same length with mass matrix $M = \mathrm{diag}(1, 1, 1, 1.67 \cdot 10^{-4}, 8.33 \cdot 10^{-5}, 8.33 \cdot 10^{-5})$ and compliance matrix $C = \mathrm{diag}(100, 10^5, 10^5, GJ, 0.5, 1)^{-1}$ is considered, where we will use two different values for $GJ = 0.33EI$ and $0.67EI$, where $EI = (EI_2 + EI_3)$. The deformed, target configuration is the steady solution to two transverse follower forces, of magnitude $F_0 = 2$ N each, applied at the tip. The non-symmetric equilibrium position (due to $EI_2 \neq EI_3$), shown in figure 5, will entail geometrical couplings between the bending modes in the two transverse directions and torsion. By using different values of GJ, the influence of the nonlinearities on the performance of the different controllers will be

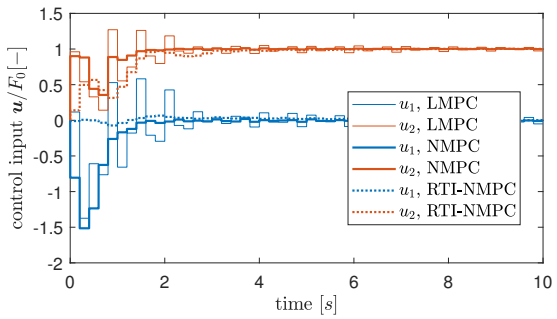


Fig. 4: Control input comparison of all the three controllers acting on the *true* system

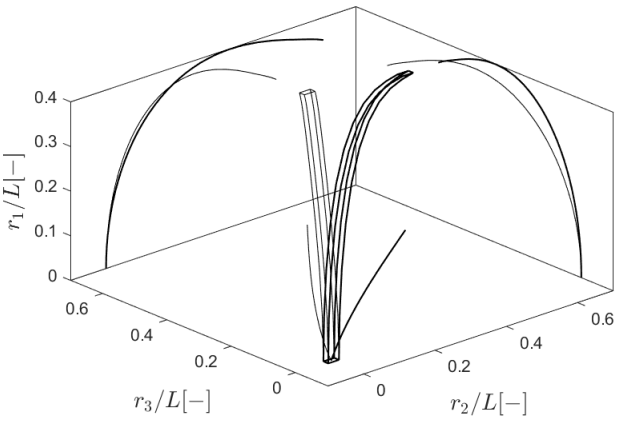


Fig. 5: Target position for the less flexible (thin stroke) and more flexible (thicker stroke) beams, where projections onto each plane have been included and where the width of the beam specifies the direction of higher stiffness.

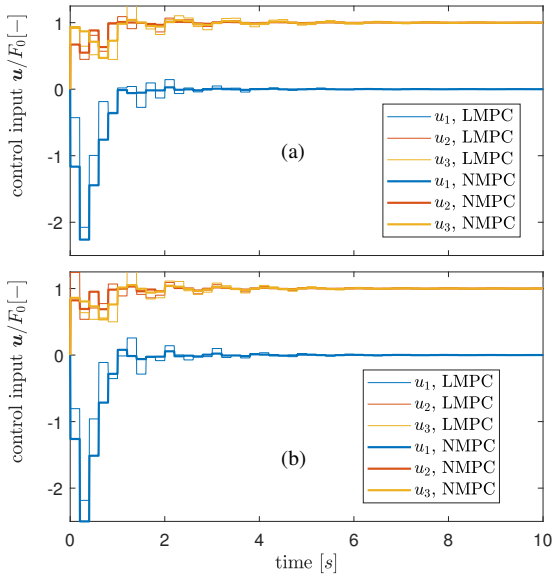


Fig. 6: Control input comparison of LMPC and NMPC for the *nominal* case for a torsion stiffness of (a) $GJ = 0.67EI$ and of (b) $GJ = 0.33EI$.

revealed. Note how in the more flexible case (thick line) in figure 5, the target solution shows an increase in twist, laying further away from the symmetric solution that $EI_2 = EI_3$ would produce.

The controller inputs are, here, three follower forces applied at the tip $\boldsymbol{u} = [f_1, f_2, f_3]^\top$:

$$\boldsymbol{x}_2(L,t) = [f_1(t), f_2(t), f_3(t), 0, 0, 0]^\top. \tag{77}$$

The controller parameters (prediction horizon and intervals) are the same as in the 2D case except for the number of modes (now $N_m = 8$): two torsion modes and two bending modes corresponding to the other transverse direction are added. Similarly to the previous 2D case, the system is fully stabilisable around $\bar{\boldsymbol{q}}$, and a much smaller terminal region of $a = 0.03$ is estimated to satisfy the necessary conditions [18].

It is in this 3D case that the nonlinear effects become more relevant when compared to the 2D one, since the follower forces at the tip are very sensitive to the overall twist of the beam. This is observed in figure 6, which shows how the LMPC controller becomes more oscillatory and starts to diverge from the NMPC response, particularly the transverse force inputs, as the beam becomes more flexible in torsion (Fig. 6b), eventually leading to failure for values of $GJ \lessapprox 0.167EI$, where the NMPC still shows convergence. The NMPC shows more effective control, even reaching saturation on the control input u_1 for the more flexible beam case.

Similarly, we include simulations of the more flexible case ($GJ = 0.33EI$) where the three different MPC implementations are set to control the *true* system. The corresponding obtained control inputs are gathered in figure 7. The difference between the results on the nominal system and on the *true* system is larger in the 3D scenario, where even the more robust NMPC shows higher oscillations throughout most of the simulation, however it successfully stabilises the system. The RTI-NMPC implementation follows the same tendency as in the 2D case and achieves stabilisation, producing a smooth solution by increasing gradually the control input until the target configuration is reached.

3) 3D case on a moving base with gravity effects: Finally, a case including gravity is presented, and hence, the rotation field needs to be simulated. It is in this case where the bespoke MPC setup described in §IV is required. The problem is now a disturbance rejection one, where an inverted pendulum is to be stabilised via control of the moving base using follower force inputs in every direction $\boldsymbol{u} = [f_1, f_2, f_3]^\top$

$$\boldsymbol{x}_2(0,t) = -[f_1(t), f_2(t), f_3(t), 0, 0, 0]^\top, \tag{78}$$

where the minus sign arises from the distinct sign convention between internal ($\boldsymbol{x}_2$ variable) and external forces (input) for $s = 0$. The system is perturbed with an initial rigid-body angular velocity of $\sqrt{2}\,\mathrm{rad/s}$ with equal components in both transverse directions, with the similar aim as in the previous 3D case to trigger nonlinear behaviour.

The pendulum is a straight beam with $L = 1$, with free boundary conditions on both ends except that the structure is prevented from spinning around its longitudinal axis only at the base (twisting deformations are still allowed). The beam has a mass matrix $M = \mathrm{diag}(1, 1, 1, 9 \cdot 10^{-5}, 6 \cdot 10^{-5}, 3 \cdot 10^{-5})$ and compliance matrix $C = \mathrm{diag}(500, 5 \cdot 10^4, 5 \cdot 10^4, 0.05, 2, 1)^{-1}$, with appropriate units.

To set up the MPC controller we use a sampling period of $\tau_s = 0.2$ s and a prediction time horizon of $\tau_p = 4$ s (0.12 and 9.8 times the period of the first bending mode in the 3-direction), which results in $\mathrm{N} = 20$ intervals. The controller's model uses 13 mode shapes which include two axial, four

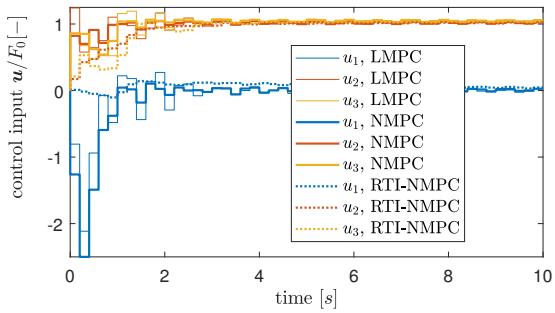


Fig. 7: Control input comparison of the three controllers on the *true* system for the flexible beam case.

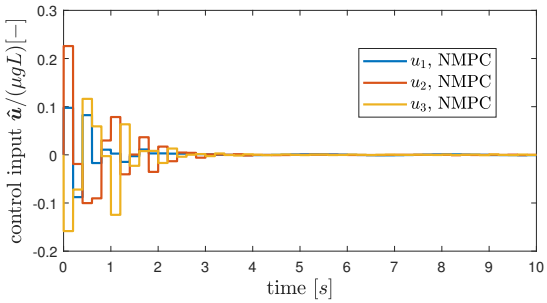


Fig. 8: Control input in the *nominal* case using the present NMPC design for $\delta = 0.0001$.

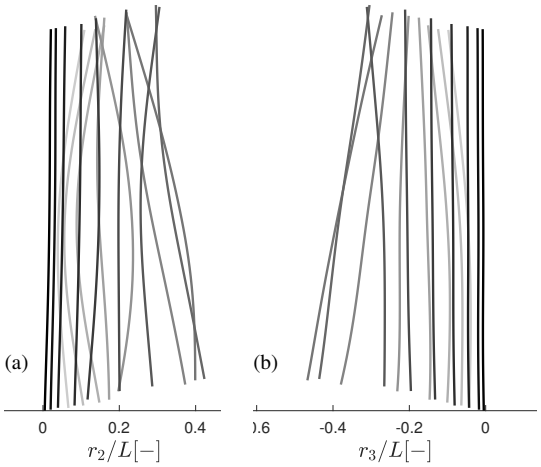


Fig. 9: Fifteen snapshots for $t \in [0.1, 6]$ s, with logarithmic spacing in time, of the controlled pendulum in the (a) 1–2 and the (b) 1–3 planes, with shade darkening for increasing time.

bending (two for each direction), two torsion and the five rigid-body modes corresponding to the described configuration.

For a first test case, we use the spline approach to parametrise rotations described in §III-B using a grid with two intervals within the internal model. This results in a state space with 36 states: 13 q_1–states, 8 q_2–states, 12 quaternion components and 3 extra states which correspond to the base displacement in the global reference

$$\dot{\boldsymbol{r}}_b = T(\boldsymbol{\xi}_1)\Phi_{1v}(0)\boldsymbol{q}_1, \tag{79}$$

where Φ_{1v} is a matrix constructed by horizontally concatenating the three first components (translational velocity) of the velocity mode shapes $\boldsymbol{\phi}_1$. As discussed previously, linearisation of (65) gives unstabilisable modes (in this case 10), which have components in both the quaternion field, the four bending q_2–states and the three displacements of the base.

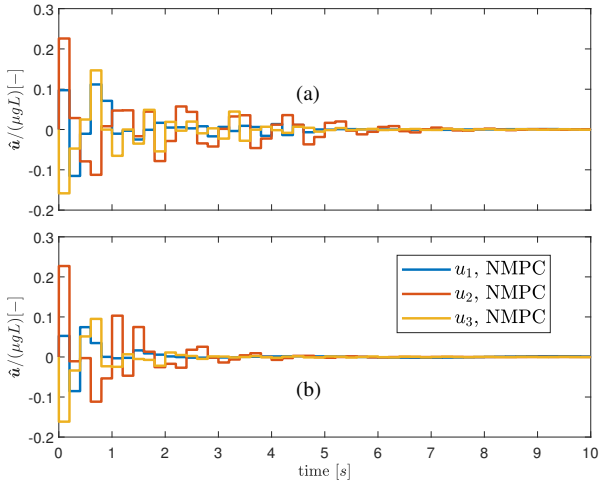


Fig. 10: Control input for both the (a) spline and the (b) modal finite-dimensional approximations for the quaternions applied to the *true* system

This means that the structure can be brought to a state where velocities are zero (i.e, equilibrium) but with remaining deformed configurations due to the presence of gravitational

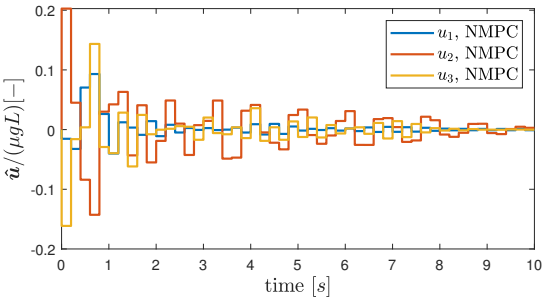


Fig. 11: Control input on the *true* system using the present NMPC design for $\delta = 0.0001$ and constraint on vertical base displacement.

forces and with the base at a location different from the origin. However, we want to shrink the set of all these possible undesired equilibrium points, that is, the control objective is to stabilise the system to a point where velocities are zero, and the remaining stress configuration is minimum, also returning the base of the pendulum as close as possible to the origin. To achieve this, we make use of the NMPC machinery described in § IV.

Initially, we have to solve the LMI system (70) and (71) for a fixed β. We are interested in the highest value of β possible, since it is linked with desirable decay rates, and we obtain for this particular system, a value of $\beta = 0.0092$, together with weight matrices P_w and P_d with $1.05 \leq \sigma(P_w) \leq 66.12$, $1.25 \leq \sigma(P_d) \leq 22.06$, and $R = 10^{-4}\text{diag}(1.35, 1.05, 1.23)$.

Next, by Lemma A.2, the existence of (possibly very small) constants δ and ω which lead to decay of the MPC cost function is assured. We estimate values of δ and ω for which the condition (85) of Lemma A.2 is satisfied, at least for a large cloud of points within the corresponding balls that δ and ω define. We are interested in attaining the lowest possible values for δ and the highest possible for ω (i.e., to minimise the effect of unstabilisable modes and widen the set of stable modes from where convergence is assured), achieving very reasonable values of $\delta = 0.0001$ and $\omega = 0.178$. These

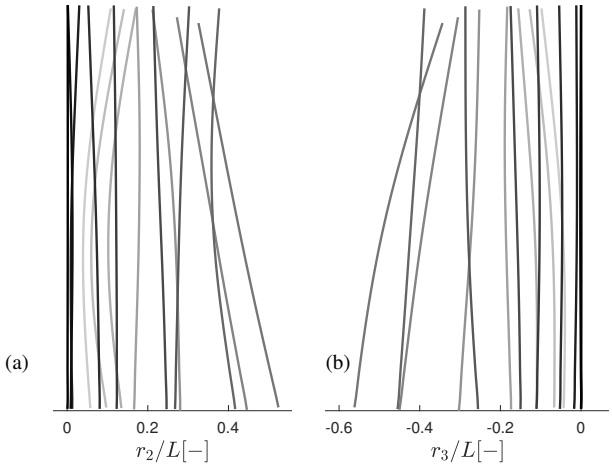


Fig. 12: Fifteen snapshots for $t \in [0.1, 10]$ s, with logarithmic spacing in time, of the controlled pendulum in the (a) 1–2 and the (b) 1–3 planes, with shade darkening with time for the vertical displacement constrained case.

terminal constraints are imposed as in (66e), however splitting the constraint into one for the stabilisable modes and a second for the unstabilisable ones.

Very good performance is achieved, as suggested by the obtained control input trajectories shown in figure 8, with stabilisation achieved efficiently and inputs close to zero for $t \geq 4$. Snapshots showing how the control achieves stabilisation of the structure are provided in figure 9.

We now show how even when applied to the *true* system (i.e. a resolved numerical simulation of the PDE), stabilisation is achieved using both the modal and the spline approaches. The system constructed with the modal approach for the quaternions has exactly the same structure, as represented by (65), resulting in the same number of stabilisable modes as the spline approximation model. Although the same $\beta = 0.0092$ is obtained for this new system, different parameters result from the MPC design: $\delta = 0.0001$ is chosen to be the same, however we need a smaller $\omega = 0.068$ and weight matrices P_w and P_d have $1.06 \leq \sigma(P_w) \leq 139.87$, $1.25 \leq \sigma(P_d) \leq 69.00$, and $R = 10^{-4}\text{diag}(2.74, 1.72, 2.17)$. Indeed, both internal models achieve stabilisation showing only a slight increase in oscillatory response due to the uncertainty arising from the plant/model mismatch. This gives evidence that the modal approach performs as efficiently as the spline method also when applied to the *true* system.

We go even further to exploit MPC's potential to place constraints on the states to simulate a case where we constrain the vertical displacement of the base to be zero, emulating the scenario where the pendulum has to be controlled maintaining the base in a horizontal 2D plane. Employing the internal model equipped with the spline approximation and using the previously obtained MPC parameters applied to the *true* system, the obtained results are again very satisfactory. Despite a slower stabilisation of the pendulum as observed in figure 11 due to the extra constraints that the control inputs must satisfy, we see that the vertical displacement of the base in figure 12 is effectively suppressed.

VI. Conclusions

An NMPC strategy to control very flexible structures has been successfully demonstrated. The use of a modal-based finite-dimensional approximation to the intrinsic beam equations allows us to retain geometrically nonlinear phenomena characteristic of structures that undergo large deformations. This nonlinear, but low-frequency behaviour, can be easily captured in our formulation by using a model constructed with just a few modes, which is of great significance since this can be attained at a low computational cost. Through an extensive numerical investigation it is demonstrated that an NMPC scheme underpinned by such minimal finite-dimensional descriptions can successfully stabilise the highly-resolved PDE, providing evidence that the error produced in the modal approximation does not hinder the convergence and stability properties of the scheme. Proving rigorous bounds on this error will be the focus of future research.

We have shown that since the PDE class considered in this paper contains Euler-Bernoulli and Timoshenko beams, our NMPC results apply to a wide class of flexible structure models. Furthermore, damping models which have been typically expressed under the Euler-Bernoulli or Timoshenko beam hypotheses are translated into our more general formulation.

To integrate rotations within the intrinsic description, we have introduced two parametrisations that yield computationally efficient, yet alternative, approaches. These have been shown to underpin successful NMPC control strategies. Properties of the two quaternion parametrisations, which ensure that the nature of the associated rotation transformation is preserved, have been also investigated. The use of a continuous-time adjoint-based sensitivity analysis, exploiting the compact structure of the resulting finite-dimensional approximation, has also been shown to provide a computational advantage over standard sensitivity analysis strategies.

For these reasons, the proposed NMPC framework offers, for the first time, a viable strategy for real-time nonlinear control of very flexible structures.

Appendix

The purpose of this section is to prove Theorem 11, which requires a series of preliminary Lemmas.

Proof of Lemma 10

Since $[A_w, B_w]$ are stabilisable, there exists $0 < \beta < 1, P_w \in \mathbb{S}^{n_1}_{++}$ and $K_w \in \mathbb{R}^{m\times n_1}$ such that $(A_w + B_wK_w)^\top P_w(A_w + B_wK_w) - P_w \preceq -\beta P_w$. Now, by a Schur complement argument $X \succeq 0$ is equivalent to

$$\begin{bmatrix} \hat{A}_w^\top P_w \hat{A}_w & \hat{A}_w^\top P_w \hat{A}_d \\ [*] & \hat{A}_d^\top P_w \hat{A}_d \end{bmatrix} \preceq \begin{bmatrix} (1-\beta)P_w & 0 \\ 0 & P_d \end{bmatrix},$$

where $\hat{A}_w = A_w + B_wK_w$ and $\hat{A}_d = A_d + B_wK_d$. Consequently, for β, P_w, K_w as above and fixed K_d, it is clear that there exists $P_d \in \mathbb{S}^{n_2}_{++}$ sufficiently large and $R \in \mathbb{S}^m_{++}$ sufficiently small such that $X \succeq 0$ and $Y \succeq 0$. □

Lemma A.1. *Suppose that* (70) *and* (71) *hold. Let* $\hat{A}_w$ *and* $\hat{A}_d$ *be as above and let* $K = \begin{bmatrix} K_w & K_d \end{bmatrix}$. *Then for any* $\boldsymbol{x} = \left(\begin{smallmatrix} \boldsymbol{w} \\ \boldsymbol{d} \end{smallmatrix}\right) \in \mathbb{R}^n$

$$\|\hat{A}_w\boldsymbol{w} + \hat{A}_d\boldsymbol{d}\|^2_{P_w} - \|\boldsymbol{w}\|^2_{P_w} \leq -\beta\|\boldsymbol{w}\|^2_{P_w} + \|\boldsymbol{d}\|^2_{P_d}. \quad (80)$$

and

$$\ell(\boldsymbol{x}, K\boldsymbol{x}) \leq \frac{\beta}{2}\|\boldsymbol{w}\|^2_{P_w} + \frac{1}{2}\|\boldsymbol{d}\|^2_{P_d}. \quad (81)$$

Proof. Follows from a Schur complement argument. □

Lemma A.2. *Suppose* (70) *and* (71) *hold. Let* $\eta > 0$. *There exists* $\omega > 0$ *such that for any* $0 < \delta < \frac{\omega\beta}{3}$ *the following is true: for any* $\boldsymbol{x} \in \chi(\omega, \delta)$, *if* $\boldsymbol{x}^+ = (A + BK)\boldsymbol{x} + \tilde{\boldsymbol{f}}(\boldsymbol{x}, K\boldsymbol{x})$ *is one step of the nonlinear dynamics* (67) *under the linear control* $\boldsymbol{u} = K\boldsymbol{x}$, *then*

$$V_f(\boldsymbol{x}^+) - V_f(\boldsymbol{x}) + \ell(\boldsymbol{x}, K\boldsymbol{x}) \leq 2\|\boldsymbol{d}\|^2_{P_d} \quad (82)$$

and $\boldsymbol{x}^+ \in \chi(\omega, \delta + \eta)$.

Proof. Letting $\tilde{\boldsymbol{f}} = \left[\begin{smallmatrix} \tilde{\boldsymbol{f}}_w \\ \tilde{\boldsymbol{f}}_d \end{smallmatrix}\right]$ and $\boldsymbol{x}^+ = \left[\begin{smallmatrix} \boldsymbol{w}^+ \\ \boldsymbol{d}^+ \end{smallmatrix}\right]$ we have that

$$\begin{aligned} \|\boldsymbol{w}^+\|^2_{P_w} = \|\hat{A}_w\boldsymbol{w} + \hat{A}_d\boldsymbol{d}\|^2_{P_w} + \|\tilde{\boldsymbol{f}}_w\|^2_{P_w} \\ + 2\boldsymbol{x}^\top(A + BK)^\top \Pi_w^\top P_w \tilde{\boldsymbol{f}}_w(\boldsymbol{x}, K\boldsymbol{x}), \end{aligned} \quad (83)$$

where $\Pi_w = [I_{n_1}, 0_{n_1}]$ gives the stabilisable part $\boldsymbol{w}$ of $\boldsymbol{x}$ and

$$\|\boldsymbol{d}^+\|_{P_d}^2 = \|\boldsymbol{d}\|_{P_d}^2 + \|\tilde{\boldsymbol{f}}_d(\boldsymbol{x}, K\boldsymbol{x})\|_{P_d}^2 + 2\boldsymbol{d}^\top P_d \tilde{\boldsymbol{f}}_d(\boldsymbol{x}, K\boldsymbol{x}). \tag{84}$$

Now, since $\tilde{\boldsymbol{f}}$ contains higher-than-linear terms, there exist sufficiently small $\omega, \delta > 0$ such that both

$$\begin{aligned}\|\tilde{\boldsymbol{f}}_w(\boldsymbol{x}, K\boldsymbol{x})\|_{P_w}^2 + 2\boldsymbol{x}^\top (A+BK)^\top \Pi_w^\top P_w \tilde{\boldsymbol{f}}_w(\boldsymbol{x}, K\boldsymbol{x}) \\ \leq \frac{\beta}{2}\|\boldsymbol{w}\|_{P_w}^2 + \frac{1}{2}\|\boldsymbol{d}\|_{P_d}^2\end{aligned} \tag{85}$$

and

$$\|\tilde{\boldsymbol{f}}_d(\boldsymbol{x}, K\boldsymbol{x})\|_{P_d}^2 + 2\boldsymbol{d}^\top P_d \tilde{\boldsymbol{f}}_d(\boldsymbol{x}, K\boldsymbol{x}) \leq \eta, \tag{86}$$

hold for any $\boldsymbol{x} \in \chi(\omega, \delta)$.

Now, using (80) from Lemma A.1 gives

$$\begin{aligned}V_f(\boldsymbol{x}^+) - V_f(\boldsymbol{x}) &= \|\hat{A}_w \boldsymbol{w} + \hat{A}_d \boldsymbol{d}\|_{P_w}^2 - \|w\|_{P_w}^2 + \|\tilde{\boldsymbol{f}_w}\|_{P_w}^2 \\ &\quad + 2\boldsymbol{x}^\top (A+BK)^\top \Pi_w^\top P_w \tilde{\boldsymbol{f}}_w(\boldsymbol{x}, K\boldsymbol{x}) \\ &\leq -\frac{\beta}{2}\|\boldsymbol{w}\|_{P_w}^2 + \frac{3}{2}\|\boldsymbol{d}\|_{P_d}^2.\end{aligned}$$

Equation (82) then follows from (81) of Lemma A.1.

We now show that $\boldsymbol{x}^+ \in \chi(\omega, \delta+\eta)$, given that $\boldsymbol{x} \in \chi(\omega, \delta)$. From (84) and (86),

$$\|\boldsymbol{d}^+\|_{P_d}^2 \leq \|\boldsymbol{d}\|_{P_d}^2 + \epsilon \leq \delta + \eta.$$

Finally, using (83), (85) and (80),

$$\begin{aligned}\|\boldsymbol{w}^+\|_{P_w}^2 = V_f(\boldsymbol{x}^+) &\leq \left(1 - \frac{\beta}{2}\right)\|\boldsymbol{w}\|_{P_w}^2 + \frac{3}{2}\|\boldsymbol{d}\|_{P_d}^2 \\ &\leq \left(1 - \frac{\beta}{2}\right)\omega + \frac{3\delta}{2} \leq \omega\end{aligned}$$

where the final inequality holds with the additional assumption $\delta \leq \frac{\beta\omega}{3}$. □

Lemma A.3. *Suppose that Assumption 9 holds and* (70)*,* (71) *are satisfied. For any $\epsilon > 0$, there exist $0 < \omega < \omega_0$ and $0 < \delta < \delta_0$ such that for any $\boldsymbol{x} \in X_{N,\omega,\delta}$, we have $\boldsymbol{x}^+ = f(\boldsymbol{x}, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x})) \in X_{N,\omega,\delta}$ and*

$$V_{N,\omega,\delta}^*(\boldsymbol{x}^+) \leq V_{N,\omega,\delta}^*(\boldsymbol{x}) - \frac{\beta}{4}\|\boldsymbol{w}\|_{P_w}^2 + 2\delta + \epsilon.$$

Proof. Let $\epsilon > 0$. First note that since V_N is continuously differentiable and $\chi(\omega_0, \delta_0)$ is compact, there exists $L_V > 0$ such that

$$|V_N(\boldsymbol{y}, \boldsymbol{u}) - V(\boldsymbol{y}, \boldsymbol{v})| \leq L_V \|\boldsymbol{u} - \boldsymbol{v}\|_{\boldsymbol{u}^{N-1}} \tag{87}$$

for $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{u}^{N-1}, \boldsymbol{y} \in \chi(\omega_0, \delta_0)$. By Assumption 9 there exists $\eta > 0$ such that

$$d_H\left(\mathcal{U}_{N,\omega,\delta}(\boldsymbol{y}), \mathcal{U}_{N,\omega,\delta+\eta}(\boldsymbol{y})\right) \leq \frac{\epsilon}{L_V}, \tag{88}$$

for any $\omega < \omega_0, \delta < \delta_0$ and and $\boldsymbol{y} \in X_{N,\omega,\delta}$.

Next, pick $\omega < \omega_0$ and $\delta < \omega\beta/3$ such that the consequences of Lemma A.2 are true for $\eta > 0$. Let $\boldsymbol{x} \in X_{N,\omega,\delta}$ and let $\boldsymbol{u}^* = (\boldsymbol{u}_*^0, \ldots, \boldsymbol{u}_*^{N-1})$ be an optimizing sequence for $V_{N,\omega,\delta}^*(\boldsymbol{x})$. Let $\boldsymbol{x}_*^k = \boldsymbol{\phi}(k; \boldsymbol{x}, \boldsymbol{u}^*)$ be the associated state trajectory. Then $\boldsymbol{x}_*^N \in \chi(\omega, \delta)$ and, by the assumption of control invariance,

$$\boldsymbol{x}_*^1 = f(\boldsymbol{x}, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x})) \in X_{N,\omega,\delta}.$$

Now, since $\boldsymbol{x}_*^N \in \chi(\omega, \delta)$, Lemma A.2 implies that $\boldsymbol{f}(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N) \in \chi(\omega, \delta+\eta)$ and

$$V_f(f(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N)) - V_f(\boldsymbol{x}_*^N) + \ell(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N) \leq 2\|\boldsymbol{d}_*^N\|_{P_d}^2. \tag{89}$$

Furthermore, since $\boldsymbol{f}(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N) \in \chi(\omega, \delta+\eta)$, it follows that the input sequence $\tilde{\boldsymbol{u}} = (\boldsymbol{u}_*^1, \ldots, \boldsymbol{u}_*^{N-1}, K\boldsymbol{x}_*^N)$ satisfies

$$\tilde{\boldsymbol{u}} \in \mathcal{U}_{N,\omega,\delta+\eta}(\boldsymbol{x}_*^1).$$

By (88), there exists $\boldsymbol{v} \in \mathcal{U}_{N,\omega,\delta}(\boldsymbol{x}_*^1)$ such that $\|\tilde{\boldsymbol{u}} - \boldsymbol{v}\|_{\boldsymbol{u}^{N-1}} \leq \epsilon/L_V$. By (87),

$$|V_N(\boldsymbol{x}_*^1, \boldsymbol{v}) - V_N(\boldsymbol{x}_*^1, \tilde{\boldsymbol{u}})| \leq \epsilon.$$

By optimality and the above inequality,

$$V_{N,\omega,\delta}^*(\boldsymbol{x}_*^1) \leq V_N(\boldsymbol{x}_*^1, \boldsymbol{v}) \leq V_N(\boldsymbol{x}_*^1, \tilde{\boldsymbol{u}}) + \epsilon.$$

Using the definitions of $\tilde{\boldsymbol{u}}$ and $V_{N,\omega,\delta}^*(\boldsymbol{x})$,

$$\begin{aligned}V_N(\boldsymbol{x}_*^1, \tilde{\boldsymbol{u}}) &\leq \sum_{k=1}^{N-1} \ell(\boldsymbol{x}_*^k, \boldsymbol{u}_*^k) + \ell(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N) \\ &\quad + V_f(f(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N)) \\ &= \left[\sum_{k=0}^{N-1} \ell(\boldsymbol{x}_*^k, \boldsymbol{u}_*^k) + V_f(\boldsymbol{x}_*^N)\right] \\ &\quad - \ell(\boldsymbol{x}, \boldsymbol{u}_*^0) - V_f(\boldsymbol{x}_*^N) \\ &\quad + V_f(f(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N)) + \ell(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N) \\ &= V_{N,\omega,\delta}^*(\boldsymbol{x}) - \ell(\boldsymbol{x}, \boldsymbol{u}_*^0) - V_f(\boldsymbol{x}_*^N) \\ &\quad + V_f(\boldsymbol{f}(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N)) + \ell(\boldsymbol{x}_*^N, K\boldsymbol{x}_*^N) \\ \text{(by (89))} &\leq V_{N,\omega,\delta}^*(\boldsymbol{x}) - \ell(\boldsymbol{x}, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x})) + 2\|\boldsymbol{d}_*^N\|_{P_d}^2 \\ (\boldsymbol{x}_*^N \in \chi(\omega, \delta)) &\leq V_{N,\omega,\delta}^*(\boldsymbol{x}) - \ell(\boldsymbol{x}, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x})) + 2\delta \\ \text{(by def. of } \ell) &\leq V_{N,\omega,\delta}^*(\boldsymbol{x}) - \frac{\beta}{4}\|\boldsymbol{w}\|_{P_w}^2 + 2\delta.\end{aligned}$$

Hence, since $\boldsymbol{x}_*^1 = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\kappa}_{N,\omega,\delta}(\boldsymbol{x})) = \boldsymbol{x}^+$ it follows that

$$V_{N,\omega,\delta}^*(\boldsymbol{x}^+) \leq V_{N,\omega,\delta}^*(\boldsymbol{x}) - \frac{\beta}{4}\|\boldsymbol{w}\|_{P_w}^2 + 2\delta + \epsilon.$$

□

Proof of Theorem 11

By Lemma A.3, there exist $\omega < \omega_0, \delta < \delta_0$ such that for any $k \in \mathbb{N}$,

$$V_{N,\omega,\delta}^*(\boldsymbol{x}^k) \leq V_{N,\omega,\delta}^*(\boldsymbol{x}^0) - \frac{\beta}{4}\sum_{i=0}^{k-1} \|\boldsymbol{w}^i\|_{P_w}^2 + k(2\delta + \epsilon).$$

Now, suppose that there exists $\rho > 0$ such that $\|\boldsymbol{w}^k\|_{P_w}^2 \geq (4/\beta)(2\delta + \epsilon) + \rho$ for all $k \in \mathbb{N}$. Then

$$0 \leq V_{N,\omega,\delta}^*(\boldsymbol{x}^k) \leq V_{N,\omega,\delta}^*(\boldsymbol{x}^0) - \rho k, \qquad k \in \mathbb{N},$$

which is a contradiction. Hence, there exists $k_0 \in \mathbb{N}$ such that $\|\boldsymbol{w}^{k_0}\|_{P_w}^2 \leq \frac{4}{\beta}(2\delta + \epsilon)$. The eventual bound follows from estimating the maximum increase in $\|w^k\|_{P_w}$ after the first time its value is less than $(4/\beta)(2\delta + \epsilon)$. □

ACKNOWLEDGMENT

Marc Artola is member of the Innovative Training Network ConFlex. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 765579.

REFERENCES

[1] J. C. Simo, "A Finite Strain Beam Formulation. The Three-Dimensional Dynamic Problem," *Comput. Methods Appl. Mech. Eng*., vol. 49, no. 1, pp. 55–70, 1985.

[2] D. H. Hodges, "Geometrically Exact , Intrinsic Theory for Dynamics of Curved and Twisted Anisotropic Beams," *AIAA J*., vol. 41, no. 6, pp. 1131–1137, 2003.

[3] S. Feng and H. Wu, "Hybrid robust boundary and fuzzy control for disturbance attenuation of nonlinear coupled ode-beam systems with application to a flexible spacecraft," *IEEE Transactions on Fuzzy Syst*., vol. 25, pp. 1293–1305, Oct 2017.

[4] W. He and S. S. Ge, "Dynamic modeling and vibration control of a flexible satellite," *IEEE Transactions on Aerosp. Electron. Syst*., vol. 51, pp. 1422–1431, April 2015.

[5] M. Mileti, D. Strzer, A. Arnold, and A. Kugi, "Stability of an euler-bernoulli beam with a nonlinear dynamic feedback system," *IEEE Transactions on Autom. Control*., vol. 61, pp. 2782–2795, Oct 2016.

[6] S. E. Khadem and A. A. Pirmohammadi, "Analytical development of dynamic equations of motion for a three-dimensional flexible link manipulator with revolute and prismatic joints," *IEEE Transactions on Syst. Man, Cybern. Part B*, vol. 33, pp. 237–249, April 2003.

[7] H. Ueno, Y. Xu, and T. Yoshida, "Modeling and control strategy of a 3-d flexible space robot," in *IEEE/RSJ International Workshop on Intelligent Robots and Systems '91*, pp. 978–983 vol.2, Nov 1991.

[8] A. Macchelli, "Stabilisation of a nonlinear flexible beam in port-hamiltonian form," *IFAC Proc. Vol*., vol. 46, no. 23, pp. 412 – 417, 2013. 9th IFAC Symposium on Nonlinear Control Systems.

[9] NASA, "Investigation of the Helios Prototype Aircraft Mishap," tech. rep., NASA, 2004.

[10] NASA, "Helios Prototype Flying Wing," 2009. Available online: https://www.nasa.gov/centers/dryden/news/ResearchUpdate/Helios/ [Accessed on 23/11/2018].

[11] E. F. Camacho and C. Bordons Alba, *Model Predictive Control*. Springer-Verlag, 2 ed., 2007.

[12] R. Palacios, J. Murua, and R. Cook, "Structural and Aerodynamic Models in Nonlinear Flight Dynamics of Very Flexible Aircraft," *AIAA J*., vol. 48, no. 11, pp. 2648–2659, 2010.

[13] S. Haghighat, H. H. T. Liu, and J. R. R. A. Martins, "Model-Predictive Gust Load Alleviation Controller for a Highly Flexible Aircraft," *J. Guid. Control. Dyn*., vol. 35, no. 6, pp. 1751–1766, 2012.

[14] Y. Wang, A. Wynn, and R. Palacios, "Nonlinear Aeroelastic Control of Very Flexible Aircraft Using Model Updating," *J. Aircr*., vol. 55, no. 4, pp. 1551–1563, 2018.

[15] M. Diehl, H. Bock, and J. Schlöder, "A real-time iteration scheme for nonlinear optimization in optimal feedback control," *SIAM J. on Control. Optim*., vol. 43, no. 5, pp. 1714–1736, 2005.

[16] R. Verschueren, G. Frison, D. Kouzopis, N. van Duijkeren, A. Zanelli, R. Quirynen, and M. Diehl, "Towards a modular software package for embedded optimization," *IFAC Pap. OnLine*, vol. 51, no. 20, pp. 374–380, 2018.

[17] J. Andersson, J. Åkesson, and M. Diehl, "Dynamic optimization with CasADi," in *51st IEEE Conference on Decision and Control (CDC)*, pp. 681–686, 2012.

[18] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, *Model Predictive Control: Theory, Computation and Design*. Nob Hill Publishing, 2nd ed., 2018.

[19] A. Wynn, Y. Wang, R. Palacios, and P. J. Goulart, "An energy-preserving description of nonlinear beam vibrations in modal coordinates," *J. Sound Vib*., vol. 332, no. 21, pp. 5543 – 5558, 2013.

[20] A. Hanson, *Visualizing quaternions*. Elsevier, 2005.

[21] A. Macchelli, C. Melchiorri, and S. Stramigioli, "Port-Based Modeling of a Flexible Link," *IEEE Transactions on Robotics*, vol. 23, no. 4, pp. 650–660, 2007.

[22] B. Jacob and H. Zwart, *Linear Port-Hamiltonian Systems on Infinite-dimensional Spaces*. Springer, 2012. 10.1007/978-3-0348-0399-1.

[23] R. Palacios, "Nonlinear normal modes in an intrinsic theory of anisotropic beams," *J. Sound Vib*., vol. 330, no. 8, pp. 1772–1792, 2011.

[24] D. Karagiannis and V. Radisavljevic-Gajic, "Sliding Mode Boundary Control of an Euler-Bernoulli Beam Subject to Disturbances," *IEEE Transactions on Autom. Control*., vol. 63, no. 10, pp. 3442–3448, 2018.

[25] T. Endo, M. Sasaki, and F. Matsuno, "Contact-Force Control of a Flexible Timoshenko Arm," *IEEE Transactions on Autom. Control*., vol. 62, no. 2, pp. 1004–1009, 2017.

[26] J. LaSalle and S. Lefschetz, *Stability by Lyapunov's Direct Method,*. Academic Press, 1961.

[27] 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), Apr. 2011.

[28] S. E. Zhelezovskii and A. D. Lyashko, "Error estimates of the galerkin method for quasilinear hyperbolic equations," *Differ. Equations*, vol. 37, no. 7, pp. 988–997, 2001.

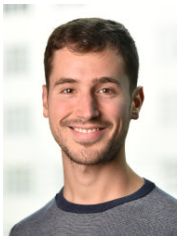
[29] M. Artola, A. Wynn, and R. Palacios, "A Nonlinear Modal-Based Framework for Low Computational Cost Optimal Control of 3D Very Flexible Structures," in *2019 European Control Conference*, June 2019.

[30] R. G. Keys, "Cubic Convolution Interpolation for Digital Image Processing," *IEEE Transactions on Acoust. Speech, Signal Process*., vol. 29, no. 6, pp. 1153–1160, 1981.

[31] H. Bock and P. Krämer-Eis, "A Multiple Shooting Method for Numerical Computation of Open and Closed Loop Controls in Nonlinear Systems," *IFAC Proc. Vol*., vol. 17, no. 2, pp. 411–415, 1984.

[32] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 2nd ed., 2006.

[33] Y. Cao, S. Li, L. Petzold, and R. Serban, "Adjoint Sensitivity Analysis for Differential-Algebraic Equations: the Adjoint DAE System and its Numerical Solution," *J. Sci. Comput*., vol. 24, no. 3, pp. 1076–1089, 2003.



Marc Artola received a Bachelor's Degree in Aerospace Technologies Engineering and a Master's Degree in Aerospace Engineering at Universitat Politècnica de Catalunya. In 2016, he joined Imperial College as a student of the MSc in Advanced Computational Methods for aeronautics and currently working towards the Ph.D. degree in aeronautics. His research interests include nonlinear aeroservoelasticity, optimal control and predictive control.

Dr. Andrew Wynn received an MMath degree in mathematics, in 2005, and the D.Phil. degree in pure mathematics, in 2009, from the University of Oxford. He held Postdoctoral positions (2009-2012) in the Dpt. of Mathematics, University College London, and the Dpt. of Aeronautics, Imperial College London, in which he is currently Senior Lecturer in Control Engineering. His research interests include operator theory, distributed parameter systems, fluid dynamics and flow control.

Prof. Rafael Palacios is Aeronautical Engineer from Universidad Politécnica, Madrid (1998), and Ph.D. in Aerospace Engineering from the University of Michigan (2005). He is AE of journals JFS, JAircraft and JPAS, FRAeS and Associate Fellow of the AIAA. He is the Prof. of Computational Aeroelasticity at the Dpt. of Aeronautics, Imperial College London and leads the Load Control and Aeroelastics lab. His research interests include aeroelastic design, aerodynamic load control, and fluid-structure interaction problems.