Dynamical reduced basis methods for Hamiltonian systems

Dynamical reduced basis methods for Hamiltonian systems

Cecilia Pagliantini[1]

Abstract

We consider model order reduction of parameterized Hamiltonian systems describing nondissipative phenomena, like wave-type and transport dominated problems. The development of reduced basis methods for such models is challenged by two main factors: the rich geometric structure encoding the physical and stability properties of the dynamics and its *local* low-rank nature. To address these aspects, we propose a nonlinear structure-preserving model reduction where the reduced phase space evolves in time. In the spirit of dynamical low-rank approximation, the reduced dynamics is obtained by a symplectic projection of the Hamiltonian vector field onto the tangent space of the approximation manifold at each reduced state. A priori error estimates are established in terms of the projection error of the full model solution onto the reduced manifold. For the temporal discretization of the reduced dynamics we employ splitting techniques. The reduced basis satisfies an evolution equation on the manifold of symplectic and orthogonal rectangular matrices having one dimension equal to the size of the full model. We recast the problem on the tangent space of the matrix manifold and develop intrinsic temporal integrators based on Lie group techniques together with explicit Runge–Kutta (RK) schemes. The resulting methods are shown to converge with the order of the RK integrator and their computational complexity depends only linearly on the dimension of the full model, provided the evaluation of the reduced flow velocity has a comparable cost.

Mathematics Subject Classification 37N30 · 65P10 · 15A24 · 78M34

1 Introduction

Hamiltonian mechanics is a cornerstone of physics and has provided the mathematical foundation for the equations of motion of systems that describe conservative processes.

✉ Cecilia Pagliantini
c.pagliantini@tue.nl

1 Centre for Analysis, Scientific computing and Applications, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands

Hamiltonian systems can be viewed as dynamical extension of the first law of thermodynamics. In this work, we consider parameterized finite-dimensional canonical Hamiltonian systems: these can model energy-conserving nondissipative flows or can ensue from the numerical discretization of partial differential equations derived from action principles. Many relevant models in mathematical physics can be written as Hamiltonian systems, and find application in, for example, classical mechanics, quantum dynamics, population and epidemics dynamics. Furthermore, partial differential equations that can be derived from action principles include Maxwell's equations, Schrödinger's equation, Korteweg–de Vries and the wave equation, compressible and incompressible Euler equations, Vlasov–Poisson and Vlasov–Maxwell equations.

Our target problem is as follows. Let $\mathcal{T} := (t_0, T]$ be a temporal interval and let $\mathcal{V}_{2N}$ be a $2N$-dimensional vector space. Let $\Gamma \subset \mathbb{R}^d$, with $d \geq 1$, be a compact set of parameters. For each $\eta \in \Gamma$, we consider the initial value problem: For $u_0(\eta) \in \mathcal{V}_{2N}$, find $u(\cdot, \eta) \in C^1(\mathcal{T}, \mathcal{V}_{2N})$ such that

$$\begin{cases} \partial_t u(t, \eta) = \mathcal{X}_{\mathcal{H}}(u(t, \eta), \eta), & \text{for } t \in \mathcal{T}, \\ u(t_0, \eta) = u_0(\eta), \end{cases} \tag{1.1}$$

where $\mathcal{X}_{\mathcal{H}}(u, \eta) \in \mathcal{V}_{2N}$ is the Hamiltonian vector field at time $t \in \mathcal{T}$, and $C^1(\mathcal{T}, \mathcal{V}_{2N})$ denotes continuous differentiable functions in time taking values in $\mathcal{V}_{2N}$. Numerical simulations of systems like (1.1) can become prohibitively expensive, in terms of computational cost, if the number $2N$ of degrees of freedom is large. In the context of long-time and many-query simulations, this often leads to unmanageable demands on computational resources. Model order reduction aims at alleviating this computational burden by replacing the original high-dimensional problem with a low-dimensional, efficient model that is fast to solve but that approximates well the underlying full-order dynamics. When dealing with Hamiltonian systems additional difficulties are encountered to ensure that the geometric structure of the phase space, the stability and the conservation properties of the original system are not hindered during the reduction. The main goal of this work is to develop and analyze structure-preserving model order reduction methods for the efficient, accurate, and physically consistent approximation of high-dimensional parametric Hamiltonian systems.

Within model order reduction techniques, projection-based reduced basis methods (RBM) consist in building, during a computationally intensive offline phase, a reduced basis from a proper orthogonal decomposition of a set of high-fidelity simulations (referred to as snapshots) at sampled values of time and parameters. A reduced dynamics is then obtained via projection of the full model onto the lower dimension space spanned by the reduced basis. Projection-based RBM for Hamiltonian systems tailored to preserve the geometric structure of the dynamics were developed in [21] and [6] using a variational Lagrangian formulation of the problem, in [2,4,30] for canonically symplectic dynamical systems, and in [16] to deal with Hamiltonian problems whose phase space is endowed with a state-dependent Poisson manifold structure. Although the aforementioned approaches can provide robust and efficient reduced models, they might require a sufficiently large approximation space to achieve even moderate accuracy. This can be ascribed to the fact that nondissipative phenomena, like advection and wave-type problems, do not possess a *global* low-rank structure,

and are therefore characterized by slowly decaying Kolmogorov widths, as highlighted in [11]. Hence, local reduced spaces seem to provide a more effective instrument to deal with this kind of dynamical systems.

In this work we propose a nonlinear projection-based model order reduction of parameterized Hamiltonian systems where the reduced basis is dynamically evolving in time. The idea is to consider a modal decomposition of the approximate solution to (1.1) of the form

$$u(t, \eta) \approx \sum_{i=1}^{2n} U_i(t) Z_i(t, \eta), \qquad n \ll N, \quad \forall t \in \mathcal{T}, \, \eta \in \Gamma, \qquad (1.2)$$

where the reduced basis $\{U_i\}_{i=1}^{2n} \subset \mathbb{R}^{2N}$, and the expansion coefficients $\{Z_i\}_{i=1}^{2n} \subset \mathbb{R}$ can both change in time. The approximate reduced flow is then generated by the velocity field resulting from the projection of the vector field $\mathcal{X}_{\mathcal{H}}$ in (1.1) into the tangent space of the reduced space at the current state. By imposing that the evolving reduced space spanned by $\{U_i\}_{i=1}^{2n}$ is a symplectic manifold at every time the continuous reduced dynamics preserves the geometric structure of the full model.

Low-rank approximations based on a modal decomposition of the approximate solution with dynamically evolving modes similar to (1.2), have been widely studied in quantum mechanics in the multiconfiguration time-dependent Hartree (MCTDH) method, see e.g. [23]. In the finite dimensional setting, a similar approach, known as dynamical low-rank approximation [20], provides a low-rank factorization updating technique to efficiently compute approximations of time-dependent large data matrices, by projecting the matrix time derivative onto the tangent space of the low-rank matrix manifold. For the discretization of time-dependent stochastic PDEs, Sapsis and Lermusiaux proposed in [31] the so-called dynamically orthogonal (DO) scheme, where the deterministic approximation space adapts over time by evolving according to the differential operator describing the stochastic problem. A connection between dynamical low-rank approximations and DO methods was established in [29]. Further, a geometric perspective on the relation between dynamical low-rank approximation, DO field equations and model order reduction in the context of time-dependent matrices has been investigated in [14]. To the best of our knowledge, the only work to address structure-preserving dynamical low-rank approximations is [28], where the authors develop a DO discretization of stochastic PDEs possessing a symplectic Hamiltonian structure. The method proposed in [28] consists in recasting the continuous PDE into the complex setting and then applying a dynamical low-rank strategy to derive field equations for the evolution of the stochastic modal decomposition of the approximate solution. The approach we propose for the nonlinear model order reduction of problem (1.1) adopts a geometric perspective similar to [14] and yields an evolution equation for the reduced solution analogous to [28], although we do not resort to a reformulation of the evolution problem in a complex framework.

Concerning the temporal discretization of the reduced dynamics describing the evolution of the approximate solution (1.2), the low-dimensional system for the expansion coefficients $\{Z_i\}_{i=1}^{2n}$ is Hamiltonian and can be approximated using standard symplectic integrators. On the other hand, the development of numerical schemes for the

evolution of the reduced basis is more involved as two major challenges need to be addressed: (i) a structure-preserving approximation requires that the discrete evolution remains on the manifold of symplectic and (semi-)orthogonal rectangular matrices; (ii) since the reduced basis forms a matrix with one dimension equal to the size of the full model, the effectiveness of the model reduction might be thwarted by the computational cost associated with the numerical solution of the corresponding evolution equation. Various methods have been proposed in the literature to solve differential equations on manifolds, see e.g. [15, Chapter IV]. Most notably projection methods apply a conventional discretization scheme and, after each time step, a "correction" is made by projecting the updated approximate solution to the constrained manifold. Alternatively, methods based on the use of local parameterizations of the manifold, so-called *intrinsic*, are well-developed in the context of differential equations on Lie groups, *cf.* [15, Sect. IV.8]. The idea is to recast the evolution equation in the corresponding Lie algebra, which is a linear space, and to then recover an approximate solution in the Lie group via local coordinate maps. Instrinsic methods possess excellent structure-preserving properties provided the local coordinate map can be computed exactly. However, they usually require a considerable computational cost associated with the evaluation of the coordinate map and its inverse at every time step (possibly at every stage within each step).

We propose and analyze two structure-preserving temporal approximations and show that their computational complexity scales linearly with the dimension of the full model, under the assumption that the velocity field of the reduced flow can be evaluated at a comparable cost. The first algorithm we propose is a Runge–Kutta Munthe–Kaas (RK-MK) method [24], and we rely on the action on the orthosymplectic matrix manifold by the quadratic Lie group of unitary matrices. By exploiting the structure of our dynamical low-rank approximation and the properties of the local coordinate map supplied by the Cayley transform, we prove the computational efficiency of this algorithm with respect to the dimension of the high-fidelity model. However, a polynomial dependence on the number of stages of the RK temporal integrator might yield high computational costs in the presence of full models of moderate dimension. To overcome this issue, we propose a discretization scheme based on the use of retraction maps to recast the local evolution of the reduced basis on the tangent space of the matrix manifold at the current state, inspired by the works [9,10] on intrinsic temporal integrators for orthogonal flows.

The remainder of the paper is organized as follows. In Sect. 2 the geometric structure underlying the dynamics of Hamiltonian systems is presented, and the concept of orthosymplectic basis spanning the approximate phase space is introduced. In Sect. 3 we describe the properties of linear symplectic maps needed to guarantee that the geometric structure of the full dynamics is inherited by the reduced problem. Subsequently, in Sect. 4 we develop and analyze a dynamical low-rank approximation strategy resulting in dynamical systems for the reduced orthosymplectic basis and the corresponding expansion coefficients in (1.2). In Sect. 5 efficient and structure-preserving temporal integrators for the reduced basis evolution problem are derived. Section 6 concerns a numerical test where the proposed method is compared to a global reduced basis approach. We present some concluding remarks and open questions in Sect. 7.

2 Hamiltonian dynamics on symplectic manifolds

The phase space of Hamiltonian dynamical systems is endowed with a differential Poisson manifold structure which underpins the physical properties of the system. Most prominently, Poisson structures encode a family of conserved quantities that, by Noether's theorem, are related to symmetries of the Hamiltonian. Here we focus on dynamical systems whose phase space has a global Poisson structure that is canonical and nondegenerate, namely symplectic.

Definition 2.1 (*Symplectic vector space*) Let $\mathcal{V}_{2N}$ be a $2N$-dimensional real vector space. A skew-symmetric bilinear form $\omega : \mathcal{V}_{2N} \times \mathcal{V}_{2N} \to \mathbb{R}$ is *symplectic* if it is nondegenerate, i.e., if $\omega(u, v) = 0$, for any $v \in \mathcal{V}_{2N}$, then $u = 0$. The map ω is called a *linear symplectic structure* on $\mathcal{V}_{2N}$, and $(\mathcal{V}_{2N}, \omega)$ is called a *symplectic vector space*.

On a finite $2N$-dimensional smooth manifold $\mathcal{V}_{2N}$, let ω be a 2-form, that is, for any $p \in \mathcal{V}_{2N}$, the map $\omega_p : T_p\mathcal{V}_{2N} \times T_p\mathcal{V}_{2N} \to \mathbb{R}$ is skew-symmetric and bilinear on the tangent space to $\mathcal{V}_{2N}$ at p, and it varies smoothly in p. The 2-form ω is a *symplectic structure* if it is closed and ω_p is symplectic for all $p \in \mathcal{V}_{2N}$, in the sense of Definition 2.1. A manifold $\mathcal{V}_{2N}$ endowed with a symplectic structure ω is called a *symplectic manifold* and denoted by $(\mathcal{V}_{2N}, \omega)$. The algebraic structure of a symplectic manifold $(\mathcal{V}_{2N}, \omega)$ can be characterized through the definition of a bracket: Let $d\mathcal{F}$ be the 1-form given by the exterior derivative of a given smooth function $\mathcal{F}$. Then, for all $\mathcal{F}, \mathcal{G} \in C^\infty(\mathcal{V}_{2N})$,

$$\{\mathcal{F}, \mathcal{G}\}_{2N} := {}_{T^*\mathcal{V}_{2N}}\langle d\mathcal{F}, \mathcal{J}_{2N}\, d\mathcal{G}\rangle_{T\mathcal{V}_{2N}} = \omega(\mathcal{J}_{2N}\, d\mathcal{F}, \mathcal{J}_{2N}\, d\mathcal{G}), \qquad (2.1)$$

where ${}_{T^*\mathcal{V}_{2N}}\langle\cdot, \cdot\rangle_{T\mathcal{V}_{2N}}$ denotes the duality pairing between the cotangent and the tangent bundle. The function $\mathcal{J}_{2N} : T^*\mathcal{V}_{2N} \to T\mathcal{V}_{2N}$ is a contravariant 2-tensor on the manifold $\mathcal{V}_{2N}$, commonly referred to as *Poisson tensor*. The space $C^\infty(\mathcal{V}_{2N})$ of real-valued smooth functions over the manifold $(\mathcal{V}_{2N}, \{\cdot, \cdot\}_{2N})$, together with the bracket $\{\cdot, \cdot\}_{2N}$, forms a Lie algebra [1, Proposition 3.3.17].

To any function $\mathcal{H} \in C^\infty(\mathcal{V}_{2N})$, the symplectic form ω allows to associate a vector field $\mathcal{X}_\mathcal{H} \in T\mathcal{V}_{2N}$, called *Hamiltonian vector field*, via the relation

$$d\mathcal{H} = i_{\mathcal{X}_\mathcal{H}}\omega, \qquad (2.2)$$

where i denotes the contraction operator. Since ω is nondegenerate, $\mathcal{X}_\mathcal{H} \in T\mathcal{V}_{2N}$ is unique. Any vector field $\mathcal{X}_\mathcal{H}$ on a manifold $\mathcal{V}_{2N}$ determines a phase flow, namely a one-parameter group of diffeomorphisms $\Phi^t_{\mathcal{X}_\mathcal{H}} : \mathcal{V}_{2N} \to \mathcal{V}_{2N}$ satisfying $d_t\Phi^t_{\mathcal{X}_\mathcal{H}}(u) = \mathcal{X}_\mathcal{H}(\Phi^t_{\mathcal{X}_\mathcal{H}}(u))$ for all $t \in \mathcal{T}$ and $u \in \mathcal{V}_{2N}$, with $\Phi^0_{\mathcal{X}_\mathcal{H}}(u) = u$. The flow of a Hamiltonian vector field satisfies $(\Phi^t_{\mathcal{X}_\mathcal{H}})^*\omega = \omega$, for each $t \in \mathcal{T}$, that is $\Phi^t_{\mathcal{X}_\mathcal{H}}$ is a symplectic diffeomorphism (symplectomorphism) on its domain.

Definition 2.2 (*Symplectic map*) Let $(\mathcal{V}_{2N}, \{\cdot, \cdot\}_{2N})$ and $(\mathcal{V}_{2n}, \{\cdot, \cdot\}_{2n})$ be symplectic manifolds of finite dimension $2N$ and $2n$ respectively, with $n \leq N$. A smooth map $\Psi : (\mathcal{V}_{2N}, \{\cdot, \cdot\}_{2N}) \to (\mathcal{V}_{2n}, \{\cdot, \cdot\}_{2n})$ is called *symplectic* if it satisfies

$$\Psi^*\{\mathcal{F}, \mathcal{G}\}_{2n} = \{\Psi^*\mathcal{F}, \Psi^*\mathcal{G}\}_{2N}, \qquad \forall\, \mathcal{F}, \mathcal{G} \in C^\infty(\mathcal{V}_{2n}).$$

In addition to possessing a symplectic phase flow, Hamiltonian dynamics is characterized by the existence of differential invariants, and symmetry-related conservation laws.

Definition 2.3 (*Invariants of motion*) A function $\mathcal{I} \in C^\infty(\mathcal{V}_{2N})$ is an *invariant of motion* of the dynamical system (2.2), if $\{\mathcal{I}, \mathcal{H}\}_{2N}(u) = 0$ for all $u \in \mathcal{V}_{2N}$. Consequently, $\mathcal{I}$ is constant along the orbits of $\mathcal{X}_\mathcal{H}$.

The Hamiltonian, if time-independent, is an invariant of motion. A particular subset of the invariants of motion of a dynamical system is given by the *Casimir invariants*, smooth functions $\mathcal{C}$ on $\mathcal{V}_{2N}$ that $\{\cdot, \cdot\}_{2N}$-commute with every other functions, i.e. $\{\mathcal{C}, \mathcal{F}\}_{2N} = 0$ for all $\mathcal{F} \in C^\infty(\mathcal{V}_{2N})$. Since Casimir invariants are associated with the center of the Lie algebra $(C^\infty(\mathcal{V}_{2N}), \{\cdot, \cdot\}_{2N})$, symplectic manifolds only possess trivial Casimir invariants.

Resorting to a coordinate system, the canonical structure on a symplectic manifold can be characterized by canonical charts whose existence is postulated in [1, Proposition 3.3.21].

Definition 2.4 Let $(\mathcal{V}_{2N}, \{\cdot, \cdot\}_{2N})$ be a symplectic manifold and (U, ψ) a cotangent coordinate chart $\psi(u) = (q^1(u), \ldots, q^N(u), p_1(u), \ldots, p_N(u))$, for all $u \in U$. Then (U, ψ) is a *symplectic canonical chart* if and only if $\{q^i, q^j\}_{2N} = \{p_i, p_j\}_{2N} = 0$, and $\{q^i, p_j\}_{2N} = \delta_{i,j}$ on U for all $i, j = 1, \ldots, N$.

In the local canonical coordinates introduced in Definition 2.4, the vector bundle map $\mathcal{J}_{2N}$, defined in (2.1), takes the canonical symplectic form

$$J_{2N} := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix} : T^*\mathcal{V}_N \times T^*\mathcal{V}_N \longrightarrow T\mathcal{V}_{2N},$$

where Id and 0 denote the identity and zero map, respectively. Symplectic canonical charts on a symplectic vector space allow to identify a *Kälher structure*, namely a compatible combination of a scalar product and symplectic form, as follows. On a symplectic vector space $(\mathcal{V}_{2N}, \omega)$, the operator $J_{2N}^\top$ is an *almost complex structure*, that is a linear map on $\mathcal{V}_{2N}$ such that $J_{2N}^\top \circ J_{2N}^\top = -\mathrm{Id}$. Furthermore, $J_{2N}^\top$ is *compatible* with the symplectic structure ω, namely, for any $u, v \in \mathcal{V}_{2N}$, $u \neq 0$, it holds

$$\omega(J_{2N}^\top u, J_{2N}^\top v) = \omega(u, v), \qquad \text{and} \qquad \omega(u, J_{2N}^\top u) > 0.$$

A symplectic form ω on a vector space $\mathcal{V}_{2N}$ together with a compatible positive almost complex structure $J_{2N}^\top$ determines an inner product on $\mathcal{V}_{2N}$, given by

$$(u, v) := \omega\left(u, J_{2N}^\top v\right), \quad \forall u, v \in \mathcal{V}_{2N}. \tag{2.3}$$

A symplectic basis on $(\mathcal{V}_{2N}, \omega)$ is an orthonormal basis for the compatible inner product (2.3), and we refer to it as *orthosymplectic*. A subspace $\mathcal{U}$ of a symplectic vector space $(\mathcal{V}_{2N}, \omega)$ is called *Lagrangian* if it coincides with its symplectic complement in $\mathcal{V}_{2N}$, namely the set of $u \in \mathcal{V}_{2N}$ such that $\omega(u, v) = 0$ for all $v \in \mathcal{U}$. As a

consequence of the fact that any basis of a Lagrangian subspace of a symplectic vector space can be extended to a symplectic basis, every symplectic vector space admits an orthosymplectic basis, *cf.* for example [5, Sect. 1.2].

With the definitions introduced hitherto, we can recast the dynamical system (1.1) on a symplectic vector space $(\mathcal{V}_{2N}, \omega)$ as a Hamiltonian initial value problem. For each $\eta \in \Gamma$, and for $u_0(\eta) \in \mathcal{V}_{2N}$, find $u(\cdot, \eta) \in C^1(\mathcal{T}, \mathcal{V}_{2N})$ such that

$$\begin{cases} \partial_t u(t, \eta) = J_{2N} \nabla_u \mathcal{H}(u(t, \eta); \eta), & \text{for } t \in \mathcal{T}, \\ u(t_0, \eta) = u_0(\eta), \end{cases} \tag{2.4}$$

where $\mathcal{H}(\cdot, \eta) \in C^\infty(\mathcal{V}_{2N})$ is the Hamiltonian function, and ∇_u denotes the gradient with respect to the variable u. The well-posedness of (2.4) is guaranteed by assuming that, for any fixed $\eta \in \Gamma$, the operator $\mathcal{X}_\mathcal{H} : \mathcal{V}_{2N} \times \Gamma \to \mathbb{R}$ defined as $\mathcal{X}_\mathcal{H}(u, \eta) := J_{2N} \nabla_u \mathcal{H}(u; \eta)$ is Lipschitz continuous in u uniformly in $t \in \mathcal{T}$ in a suitable norm.

3 Orthosymplectic matrices

In order to construct surrogate models preserving the physical and geometric properties of the original Hamiltonian dynamics we build approximation spaces of reduced dimension endowed with the same geometric structure of the full model. To this aim, the reduced space is constructed as the span of suitable symplectic and orthonormal time-dependent bases, so that the reduced space inherits the geometric structure of the original dynamical system. In this Section we describe the properties of linear symplectic maps between finite dimensional symplectic vector spaces.

Analogously to [1, p. 168], we can easily extend the characterization of symplectic linear maps to the case of vector spaces of different dimension as in the following result.

Lemma 3.1 *Let $(\mathcal{V}_{2N}, \omega)$ and $(\mathcal{V}_{2n}, \omega)$ be symplectic vector spaces of finite dimension $2N$ and $2n$, respectively, with $N \geq n$. A linear map $M_+ : (\mathcal{V}_{2N}, \omega) \to (\mathcal{V}_{2n}, \omega)$ is symplectic, in the sense of Definition 2.2, if and only if the corresponding matrix representation $M_+ \in \mathbb{R}^{2n \times 2N}$ satisfies $M_+ J_{2N} M_+^\top = J_{2n}$.*

We define *symplectic right inverse* of the symplectic matrix $M_+ \in \mathbb{R}^{2n \times 2N}$ the matrix $M = J_{2N} M_+^\top J_{2n}^\top \in \mathbb{R}^{2N \times 2n}$. It can be easily verified that $M_+ M = I_{2n}$, and that $M : (\mathcal{V}_{2n}, \omega) \to (\mathcal{V}_{2N}, \omega)$ is the adjoint operator with respect to the symplectic form ω, i.e. $\omega(M_+ u, y) = \omega(u, My)$ for any $u \in \mathcal{V}_{2N}$, and $y \in \mathcal{V}_{2n}$. Furthermore, the symplectic condition $M_+ J_{2N} M_+^\top = J_{2n}$ is equivalent to $M^\top J_{2N} M = J_{2n}$. Owing to this equivalence, with a small abuse of notation, we will say that $M \in \mathbb{R}^{2N \times 2n}$ is symplectic if it belongs to the space

$$\text{Sp}\left(2n, \mathbb{R}^{2N}\right) := \left\{ L \in \mathbb{R}^{2N \times 2n} : L^\top J_{2N} L = J_{2n} \right\}.$$

Definition 3.2 A matrix $M \in \mathbb{R}^{2N \times 2n}$ is called *orthosymplectic* if it belongs to the space

$$\mathcal{U}\left(2n, \mathbb{R}^{2N}\right) := \mathrm{St}\left(2n, \mathbb{R}^{2N}\right) \cap \mathrm{Sp}\left(2n, \mathbb{R}^{2N}\right),$$

where $\mathrm{St}(2n, \mathbb{R}^{2N}) := \{M \in \mathbb{R}^{2N \times 2n} : M^\top M = I_{2n}\}$ is the Stiefel manifold.

Orthosymplectic rectangular matrices can be characterized as follows.

Lemma 3.3 *Let $M_+ \in \mathbb{R}^{2n \times 2N}$ be symplectic and let $M \in \mathbb{R}^{2N \times 2n}$ be its symplectic inverse. Then, $M_+ M_+^\top = I_{2n}$ if and only if $M = M_+^\top$.*

Proof Let $M = [A \mid B]$ with $A, B \in \mathbb{R}^{2N \times n}$. The (semi-)orthogonality and symplecticity of M_+ give $A^\top A = B^\top B = I_n$ and $A^\top J_{2N} B = I_n$. These conditions imply that the column vectors of A and $J_{2N} B$ have unit norm and are pairwise parallel, hence $A = J_{2N} B$. Therefore, $M = [A \mid J_{2N}^\top A]$ with $A^\top A = I_n$ and $A^\top J_{2N} A = 0_n$. The definition of symplectic inverse yields $M_+^\top = J_{2N}^\top M J_{2n} = J_{2N}^\top [A \mid J_{2N}^\top A] J_{2n} = [J_{2N}^\top A \mid -A] J_{2n} = [A \mid J_{2N}^\top A] = M$.
    Conversely, the symplecticity of M_+ implies $M_+ M_+^\top = M_+ J_{2N} M_+^\top J_{2n}^\top = I_{2n}$. $\square$

In order to design numerical methods for evolution problems on the manifold $\mathcal{U}(2n, \mathbb{R}^{2N})$ of orthosymplectic rectangular matrices, we will need to characterize its tangent space. To this aim we introduce the vector space $\mathfrak{so}(2n)$ of skew-symmetric $2n \times 2n$ real matrices $\mathfrak{so}(2n) := \{M \in \mathbb{R}^{2n \times 2n} : M^\top + M = 0_{2n}\}$, and the vector space $\mathfrak{sp}(2n)$ of Hamiltonian $2n \times 2n$ real matrices, namely $\mathfrak{sp}(2n) := \{M \in \mathbb{R}^{2n \times 2n} : M J_{2n} + J_{2n} M^\top = 0_{2n}\}$. Throughout, if not otherwise specified, we will denote with $G_{2n} := \mathcal{U}(2n)$ the Lie group of orthosymplectic $2n \times 2n$ matrices and with $\mathfrak{g}_{2n}$ the corresponding Lie algebra $\mathfrak{g}_{2n} := \mathfrak{so}(2n) \cap \mathfrak{sp}(2n)$, with bracket given by the matrix commutator $\mathrm{ad}_M(L) = [M, L] := ML - LM$, for any $M, L \in \mathfrak{g}_{2n}$.

4 Orthosymplectic dynamical reduced basis method

Assume we want to solve the parameterized Hamiltonian problem (2.4) at $p \in \mathbb{N}$ samples of the parameter $\{\eta_j\}_{j=1}^p =: \Gamma_h \subset \mathbb{R}^{pd}$. To simplify the notation we take $d = 1$, namely we assume that the parameter η is a scalar quantity, for vector-valued η the derivation henceforth applies *mutatis mutandis*. Then, the Hamiltonian system (2.4) can be recast as a set of ordinary differential equations in a $2N \times p$ matrix unknown. Let $\eta_h \in \mathbb{R}^p$ denote the vector of sampled parameters, the evolution problem reads: For $\mathcal{R}_0(\eta_h) := \left[u_0(\eta_1) | \dots | u_0(\eta_p)\right] \in \mathbb{R}^{2N \times p}$, find $\mathcal{R} \in C^1(\mathcal{T}, \mathbb{R}^{2N \times p})$ such that

$$\begin{cases} \dot{\mathcal{R}}(t) = \mathcal{X}_{\mathcal{H}}(\mathcal{R}(t), \eta_h), & \text{for } t \in \mathcal{T}, \\ \mathcal{R}(t_0) = \mathcal{R}_0(\eta_h). \end{cases} \tag{4.1}$$

Let $n \ll N$, to characterize the reduced solution manifold we consider an approximation of the solution of (4.1) of the form

$$\mathcal{R}(t) \approx R(t) = \sum_{i=1}^{2n} \mathbf{U}_i(t)\mathbf{Z}_i(t, \eta_h) = U(t)Z(t)^\top, \qquad (4.2)$$

where $U = \left[\mathbf{U}_1|\ldots|\mathbf{U}_{2n}\right] \in \mathbb{R}^{2N \times 2n}$, and $Z \in \mathbb{R}^{p \times 2n}$ is such that $Z_{j,i}(t) = \mathbf{Z}_i(t, \eta_j)$ for $i = 1, \ldots, 2n$, and $j = 1, \ldots, p$. Since we aim at a structure-preserving model order reduction of (4.1), we impose that the basis $U(t)$ is orthosymplectic at all $t \in \mathcal{T}$, in analogy with the symplectic reduction techniques employing globally defined reduced spaces. Here, since U is changing in time, this means that we constrain its evolution to the manifold $\mathcal{U}(2n, \mathbb{R}^{2N})$ from Definition 3.2. With this in mind, the reduced solution is sought in the reduced space defined as

$$\mathcal{M}_{2n}^{\text{spl}} := \left\{ R \in \mathbb{R}^{2N \times p} : R = UZ^\top \text{ with } U \in \mathcal{M}, Z \in V^{p \times 2n} \right\}, \qquad (4.3)$$

where

$$\mathcal{M} := \mathcal{U}\left(2n, \mathbb{R}^{2N}\right) = \left\{ U \in \mathbb{R}^{2N \times 2n} : U^\top U = I_{2n}, U^\top J_{2N} U = J_{2n} \right\},$$
$$V^{p \times 2n} := \left\{ Z \in \mathbb{R}^{p \times 2n} : \text{rank}(Z^\top Z + J_{2n}^\top Z^\top Z J_{2n}) = 2n \right\}. \qquad (4.4)$$

Note that (4.3) is a smooth manifold of dimension $2(N + p)n - 2n^2$, as follows from the characterization of the tangent space given in Proposition 4.1. The characterization of the reduced manifold (4.3) is analogous to [28, Definition 6.2]. Let $C \in \mathbb{R}^{2n \times 2n}$ denote the correlation matrix $C := Z^\top Z$. The full-rank condition in (4.4),

$$\text{rank}(C + J_{2n}^\top C J_{2n}) = 2n, \qquad (4.5)$$

guarantees that, for Z fixed, if $UZ^\top = WZ^\top$ with $U, W \in \mathcal{M}$, then $U = W$. If the full-rank condition (4.5) is satisfied, then the number p of samples of the parameter $\eta \in \Gamma$ satisfies $p \geq n$. This means that, for a fixed p, a too large reduced basis might lead to a violation of the full rank condition, which would entail a rank-deficient evolution problem for the coefficient matrix $Z \in \mathbb{R}^{p \times 2n}$. This is related to the problem of overapproximation in dynamical low-rank techniques, see [20, Sect. 5.3]. Observe also that if $p \geq 2n$ and $\text{rank}(Z) = 2n$ then the full rank condition (4.5) is always satisfied. In general, the elements of $\mathcal{M}_{2n}^{\text{spl}}$ might not have full rank $2n$: for any $R \in \mathcal{M}_{2n}^{\text{spl}}$ it holds $\text{rank}(Z) \leq \text{rank}(R) \leq \min\{2n, p\}$.

The decomposition $UZ^\top$ of matrices in $\mathcal{M}_{2n}^{\text{spl}}$ is not unique: the map $\phi : (U, Z) \in \mathcal{M} \times V^{p \times 2n} \mapsto R = UZ^\top \in \mathcal{M}_{2n}^{\text{spl}}$ is surjective but not injective. In particular, $(\mathcal{M} \times V^{p \times 2n}, \mathcal{M}_{2n}^{\text{spl}}, \phi, \mathcal{U}(2n))$ is a fiber bundle with fibers given by the group of unitary matrices $\mathcal{U}(2n)$, and $\mathcal{M}_{2n}^{\text{spl}}$ is isomorphic to $(\mathcal{M}/\mathcal{U}(2n)) \times V^{p \times 2n}$. Indeed, let

$U_1 \in \mathcal{M}$ and $Z_1 \in V^{p \times 2n}$, then, for any arbitrary $A \in \mathcal{U}(2n)$, it holds $U_2 := U_1 A \in \mathcal{M}$, $Z_2 := Z_1 A \in V^{p \times 2n}$, and $U_1 Z_1^\top = U_2 Z_2^\top$.

In dynamically orthogonal approximations [31] a characterization of the reduced solution is obtained by fixing a gauge constraint in the tangent space of the reduced solution manifold. For the manifold $\mathcal{M}_{2n}^{\mathrm{spl}}$ the tangent space at $R \in \mathcal{M}_{2n}^{\mathrm{spl}}$ is defined as the set of $X \in \mathbb{R}^{2N \times p}$ such that there exists a differentiable path $\gamma : (-\varepsilon, \varepsilon) \subset \mathcal{T} \to \mathbb{R}^{2N \times p}$ with $\gamma(0) = R$, $\dot{\gamma}(0) = X$. The tangent vector at $U(t) Z^\top(t) \in \mathcal{M}_{2n}^{\mathrm{spl}}$ is of the form $X = \dot{U} Z^\top + U \dot{Z}^\top$, where $\dot{U}$ and $\dot{Z}$ denote the time derivatives of $U(t)$ and $Z(t)$, respectively. Taking the derivative of the orthogonality constraint on U yields $\dot{U}^\top U + U^\top \dot{U} = 0$. Analogously, the symplecticity constraint gives $\dot{U}^\top J_{2N} U + U^\top J_{2N} \dot{U} = 0$ which is equivalent to $\dot{U}^\top U J_{2n} + J_{2n} U^\top \dot{U} = 0$ owing to the fact that $U \in \mathrm{Sp}(2n, \mathbb{R}^{2N})$. Therefore, the tangent space of $\mathcal{M}_{2n}^{\mathrm{spl}}$ at $U Z^\top$ is defined as

$$
\begin{aligned}
T_{UZ^\top} \mathcal{M}_{2n}^{\mathrm{spl}} = \{ X \in \mathbb{R}^{2N \times p} : \ & X = X_U Z^\top + U X_Z^\top \ \text{with} \ X_Z \in \mathbb{R}^{p \times 2n}, \\
& X_U \in \mathbb{R}^{2N \times 2n}, \ X_U^\top U \in \mathfrak{g}_{2n} \}.
\end{aligned}
\tag{4.6}
$$

However, this parameterization is not unique. Indeed, let $S \in \mathfrak{g}_{2n}$ be arbitrary: if $X_U^\top U \in \mathfrak{g}_{2n}$ then the matrix $(X_U + US)^\top U$ belongs to $\mathfrak{g}_{2n}$, and the pairs (X_U, X_Z) and $(X_U + US, X_Z + ZS)$ identify the same tangent vector $X := X_U Z^\top + U X_Z^\top$. We fix the parameterization of the tangent space as follows.

Proposition 4.1 *The tangent space of $\mathcal{M}_{2n}^{\mathrm{spl}}$ at $U Z^\top$ defined in* (4.6) *is uniquely parameterized by the space $H_{(U,Z)} := H_U \times \mathbb{R}^{p \times 2n}$, where*

$$
H_U := \left\{ X_U \in \mathbb{R}^{2N \times 2n} : \ X_U^\top U = 0, \ X_U J_{2n} = J_{2N} X_U \right\}.
\tag{4.7}
$$

This means that the map

$$
\begin{aligned}
\Psi : \quad H_{(U,Z)} \ & \longrightarrow \ T_{UZ^\top} \mathcal{M}_{2n}^{\mathrm{spl}} \\
(X_U, X_Z) \ & \longmapsto \ X_U Z^\top + U X_Z^\top,
\end{aligned}
$$

is a bijection.

Proof We first observe that, if $(X_U, X_Z) \in H_{(U,Z)}$ then $X_U^\top U \in \mathfrak{g}_{2n}$ is trivially satisfied, and hence $X_U Z^\top + U X_Z^\top \in T_{UZ^\top} \mathcal{M}_{2n}^{\mathrm{spl}}$.

To show that the map Ψ is injective, we take $X = 0 \in T_{UZ^\top} \mathcal{M}_{2n}^{\mathrm{spl}}$. By the definition of the tangent space (4.6), the zero vector admits the representation $0 = X_U Z^\top + U X_Z^\top$ with $U^\top X_U = 0$. This implies $0 = U^\top (X_U Z^\top + U X_Z^\top) = X_Z^\top$. Hence, $X_U Z^\top = 0$ and

$$
\begin{aligned}
0 &= X_U Z^\top Z + J_{2N} X_U Z^\top Z J_{2n}^\top \\
&= X_U Z^\top Z + J_{2N}^\top X_U J_{2n} J_{2n} Z^\top Z J_{2n}^\top \\
&= X_U \left(Z^\top Z + J_{2n} Z^\top Z J_{2n}^\top \right),
\end{aligned}
$$

which implies $X_U = 0$ in view of the full-rank condition (4.5).

For the surjectivity of Ψ we show that

$$\forall X \in T_{UZ^\top}\mathcal{M}_{2n}^{\mathrm{spl}} \quad \exists (X_U, X_Z) \in H_{(U,Z)} \quad \text{such that} \quad X = X_U Z^\top + U X_Z^\top.$$

Any $X \in T_{UZ^\top}\mathcal{M}_{2n}^{\mathrm{spl}}$ can be written as $X = \dot{U}Z^\top + U\dot{Z}^\top$ where $\dot{Z} \in \mathbb{R}^{p \times 2n}$ and $\dot{U} \in \mathbb{R}^{2N \times 2n}$ satisfies $\dot{U}^\top U \in \mathfrak{g}_{2n}$. Hence, the tangent vector X can be recast as

$$X = \dot{U}Z^\top + U\dot{Z}^\top = U\left(\dot{Z}^\top + U^\top \dot{U}Z^\top\right) + \left(\left(I_{2N} - UU^\top\right)\dot{U}\right)Z^\top.$$

We need to show that the pair (X_U, X_Z), defined as $X_U := (I_{2N} - UU^\top)\dot{U}$ and $X_Z := \dot{Z} + Z\dot{U}^\top U$, belongs to the space $H_{(U,Z)}$. From the orthogonality of U it easily follows that

$$U^\top X_U = U^\top\left(I_{2N} - UU^\top\right)\dot{U} = U^\top \dot{U} - U^\top \dot{U} = 0.$$

To prove that $X_U = J_{2N}^\top X_U J_{2n}$, we introduce the matrix $S := Z^\top Z + J_{2n}Z^\top Z J_{2n}^\top \in \mathbb{R}^{2n \times 2n}$ for which it holds $SJ_{2n} = J_{2n}S$. We then show the equivalent condition $X_U S J_{2n}^\top = J_{2N}^\top X_U S$. First, we add to X_U the zero term $(I_{2N} - UU^\top)U(\dot{Z}^\top Z + J_{2n}\dot{Z}^\top Z J_{2n}^\top)$, and use the symplectic constraint on U and its temporal derivative to get

$$\begin{aligned} X_U &= \left(I_{2N} - UU^\top\right)\dot{U} = \left(I_{2N} - UU^\top\right)\dot{U}SS^{-1} \\ &= \left(I_{2N} - UU^\top\right)\left(U\left(\dot{Z}^\top Z + J_{2n}\dot{Z}^\top Z J_{2n}^\top\right) + \dot{U}\left(Z^\top Z + J_{2n}Z^\top Z J_{2n}^\top\right)\right)S^{-1} \\ &= \left(I_{2N} - UU^\top\right)\left(XZ + J_{2N}XZJ_{2n}^\top\right)S^{-1}. \end{aligned}$$

Then, using the commutativity of the symplectic unit J_{2N} and the projection onto the orthogonal complement to the space spanned by U, i.e. $(I_{2N} - UU^\top)J_{2N} = J_{2N}(I_{2N} - UU^\top)$, results in

$$\begin{aligned} X_U S J_{2n}^\top &= \left(I_{2N} - UU^\top\right)\left(XZ + J_{2N}XZJ_{2n}^\top\right)J_{2n}^\top \\ &= J_{2N}^\top\left(I_{2N} - UU^\top\right)J_{2N}\left(XZJ_{2n}^\top + J_{2N}^\top XZ\right) = J_{2N}^\top X_U S. \end{aligned}$$

$\square$

Remark 4.2 Proposition 4.1 provides a *connection* on the fiber bundle $(\mathcal{M} \times V^{p \times 2n}, \mathcal{M}_{2n}^{\mathrm{spl}}, \phi, \mathcal{U}(2n))$ via the smooth splitting $T_{UZ^\top}\mathcal{M}_{2n}^{\mathrm{spl}} = V_{(U,Z)} \oplus H_{(U,Z)}$, for any $UZ^\top \in \mathcal{M}_{2n}^{\mathrm{spl}}$. The factor $V_{(U,Z)}$, the *vertical space*, is the subspace of $T_{UZ^\top}\mathcal{M}_{2n}^{\mathrm{spl}}$ that consists of all vectors tangent to the fiber of $UZ^\top$, while the space $H_{(U,Z)} := H_U \times \mathbb{R}^{p \times 2n}$, with H_U defined in (4.7), is a *horizontal space*. This decomposition into the subset of directions tangent to the fiber and its complementary space

provides a unique parameterization of the tangent space. We refer the reader to e.g. [13] and [19, Chapter 2], for further details on the topic.

Owing to Proposition 4.1, the tangent space of $\mathcal{M}_{2n}^{\text{spl}}$ can be characterized as

$$
\begin{aligned}
T_{UZ^\top}\mathcal{M}_{2n}^{\text{spl}} = \{X \in \mathbb{R}^{2N \times p} : \ & X = X_U Z^\top + U X_Z^\top \ \text{with} \ X_Z \in \mathbb{R}^{p \times 2n}, \\
& X_U \in \mathbb{R}^{2N \times 2n}, \ X_U^\top U = 0, \ X_U J_{2n} = J_{2N} X_U\},
\end{aligned}
$$

Henceforth, we consider $\mathcal{M}$ endowed with the metric induced by the ambient space $\mathbb{C}^{2N \times 2n}$, namely the Frobenius inner product $\langle A, B \rangle := \text{tr}(A^\mathsf{H} B)$, where A^H denotes the conjugate transpose of the complex matrix A, and we will denote with $\|\cdot\|$ the Frobenius norm. Note that, on simple Lie algebras, the Frobenius inner product is a multiple of the Killing form.

4.1 Dynamical low-rank symplectic variational principle

For any fixed $\eta \in \Gamma$, the vector field $\mathcal{X}_\mathcal{H}$ in (2.4) at time t belongs to $T_{u(t)}\mathcal{V}_{2N}$. Taking the cue from dynamical low-rank approximations [20], we derive a dynamical system on the reduced space $\mathcal{M}_{2n}^{\text{spl}}$ via projection of the velocity field $\mathcal{X}_\mathcal{H}$ of the full dynamical system (4.1) onto the tangent space of $\mathcal{M}_{2n}^{\text{spl}}$ at the current state. The reduced dynamical system is therefore optimal in the sense that the resulting vector field is the best dynamic approximation of $\mathcal{X}_\mathcal{H}$, in the Frobenius norm, at every point on the manifold $\mathcal{V}_{2N}$. To preserve the geometric structure of the full dynamics we construct a projection which is symplectic for each value of the parameter $\eta_j \in \Gamma_h$, with $1 \le j \le p$. To this aim, let us introduce on the symplectic vector space $(\mathcal{V}_{2N}, \omega)$ the family of skew-symmetric bilinear forms $\omega_j : \mathbb{R}^{2N \times p} \times \mathbb{R}^{2N \times p} \to \mathbb{R}$ defined as

$$
\omega_j(a, b) := \omega(a_j, b_j), \qquad 1 \le j \le p, \tag{4.8}
$$

where $a_j \in \mathbb{R}^{2N}$ denotes the j-th column of the matrix $a \in \mathbb{R}^{2N \times p}$, and similarly for $b_j \in \mathbb{R}^{2N}$.

Proposition 4.3 *Let $T_R \mathcal{M}_{2n}^{\text{spl}}$ be the tangent space of the symplectic reduced manifold $\mathcal{M}_{2n}^{\text{spl}}$, defined in (4.3), at a given $R := UZ^\top \in \mathcal{M}_{2n}^{\text{spl}}$. Let $S := Z^\top Z + J_{2n} Z^\top Z J_{2n} \in \mathbb{R}^{2n \times 2n}$. Then, the map*

$$
\begin{aligned}
\Pi_{T_R \mathcal{M}_{2n}^{\text{spl}}} : \mathbb{R}^{2N \times p} &\longrightarrow T_R \mathcal{M}_{2n}^{\text{spl}} \\
w &\longmapsto (I_{2N} - UU^\top)(wZ + J_{2N}wZJ_{2n}^\top)S^{-1}Z^\top + UU^\top w,
\end{aligned}
$$

is a symplectic projection, in the sense that

$$
\sum_{j=1}^{p} \omega_j\big(w - \Pi_{T_R \mathcal{M}_{2n}^{\text{spl}}} w, y\big) = 0, \qquad \forall\, y \in T_R \mathcal{M}_{2n}^{\text{spl}},
$$

where ω_j is defined in (4.8).

Proof Let $X_U(w) := (I_{2N} - UU^\top)(wZ + J_{2N}wZJ_{2n}^\top)(Z^\top Z + J_{2n}Z^\top ZJ_{2n}^\top)^{-1}$ and $X_Z(w) = w^\top U$. Using a reasoning analogous to the one in the proof of Proposition 4.1, it can be shown that $(X_U, X_Z) \in H_{(U,Z)}$. Moreover, by means of the identification $TT_R\mathcal{M}_{2n}^{\text{spl}} \cong T_R\mathcal{M}_{2n}^{\text{spl}}$, we prove that $\Pi := \Pi_{T_R\mathcal{M}_{2n}^{\text{spl}}}$ is a projection. It can be easily verified that $X_Z(\Pi w) = (\Pi w)^\top U = X_Z(w)$. Furthermore, let $F_w := wZ + J_{2N}wZJ_{2n}^\top \in \mathbb{R}^{2N \times 2n}$, then

$$
\begin{aligned}
X_U(\Pi w) &= \left(I_{2N} - UU^\top\right)\left(\left(I_{2N} - UU^\top\right)F_w S^{-1} Z^\top Z\right.\\
&\quad \left. + J_{2N}\left(I_{2N} - UU^\top\right)F_w S^{-1} Z^\top Z J_{2n}^\top\right)S^{-1}\\
&= X_U(w)Z^\top Z S^{-1} + J_{2N} X_U(w) Z^\top Z J_{2n}^\top S^{-1}.
\end{aligned}
$$

Since $X_U(w)J_{2n} = J_{2N}X_U(w)$, it follows that $X_U(\Pi w) = X_U(w)$.

Assume we have fixed a parameter $\eta_j \in \Gamma$ so that $p = 1$. Let $v := w_j \in \mathbb{R}^{2N}$ be the j-th column of the matrix $w \in \mathbb{R}^{2N \times p}$ and, hence, $\Pi v \in \mathbb{R}^{2N}$. We want to show that $\omega(v - \Pi v, y) = 0$ for all $y \in T_R\mathcal{M}_{2n}^{\text{spl}}$. By the characterization of the tangent space from Proposition 4.1, any $y \in T_R\mathcal{M}_{2n}^{\text{spl}}$ is of the form $y = Y_U Z^\top + U Y_Z^\top$ where $Y_Z \in \mathbb{R}^{1 \times 2n}$ and $Y_U \in H_U$. Therefore,

$$
\omega(v - \Pi v, y) = \omega\left(v - \Pi v, Y_U Z^\top\right) + \omega\left(v, U Y_Z^\top\right) - \omega\left(X_U Z^\top + U X_Z^\top, U Y_Z^\top\right),
$$

where $X_U = X_U(v)$ and $X_Z = X_Z(v)$, but henceforth we omit the dependence on v. Using the definition of X_Z and the symplecticity of the basis U the last term becomes

$$
\begin{aligned}
\omega\left(U X_Z^\top, U Y_Z^\top\right) &= \omega\left(UU^\top v, U Y_Z^\top\right) = \omega\left(v, J_{2N}^\top U J_{2n} U^\top U Y_Z^\top\right)\\
&= \omega\left(v, J_{2N}^\top U J_{2n} Y_Z^\top\right) = \omega\left(v, U Y_Z^\top\right).
\end{aligned}
$$

Moreover, it can be easily checked that $\omega(X_U Z^\top, U Y_Z^\top) = 0$ by definition of X_U and by the orthosymplecticity of U. Hence, the only non-trivial terms are $\omega(v - \Pi v, y) = \omega(v, Y_U Z^\top) - \omega(\Pi v, Y_U Z^\top)$. Any $Y_U \in H_U$ can be written as $Y_U = \frac{1}{2}(Y_U + J_{2N}^\top Y_U J_{2n})$; thereby

$$
\begin{aligned}
\omega(v - \Pi v, 2y) &= \omega\left(v, Y_U Z^\top + J_{2N}^\top Y_U J_{2n} Z^\top\right)\\
&\quad - \omega\left(X_U Z^\top + U X_Z^\top, Y_U Z^\top + J_{2N}^\top Y_U J_{2n} Z^\top\right) =: T_1 - T_2.
\end{aligned}
$$

We need to prove that T_1 and T_2 coincide. Let $M_i \in \mathbb{R}^{2N}$ denote the i-th column vector of a given matrix $M \in \mathbb{R}^{2N \times 2n}$. The properties of the symplectic canonical form ω yield

$$T_1 = \omega\left(v, \sum_{i=1}^{2n}(Y_U)_i Z_i\right) + \omega\left(J_{2N}v, \sum_{i=1}^{2n}(Y_U)_i(J_{2n}Z^\top)_i\right)$$

$$= \sum_{i=1}^{2n}\omega\big(v,(Y_U)_i\big)Z_i + \sum_{i=1}^{2n}\omega\left(J_{2N}v,(Y_U)_i\right)\left(J_{2n}Z^\top\right)_i$$

$$= \sum_{i=1}^{2n}\omega\left(vZ_i + J_{2N}v\big(ZJ_{2n}^\top\big)_i, (Y_U)_i\right).$$

To deal with the term T_2 first observe that $\omega(UX_Z^\top, Y_U Z^\top) = 0$ since $Y_U^\top U = 0$. Moreover, using once more the fact that $Y_U \in H_U$ results in

$$T_2 = \omega\left(X_U Z^\top, Y_U Z^\top\right) + \omega\left(X_U J_{2n} Z^\top, Y_U J_{2n} Z^\top\right)$$

$$= \sum_{i,j=1}^{2n}\omega\big((X_U)_j Z_j, (Y_U)_i Z_i\big) + \omega\left((X_U)_j\big(J_{2n}Z^\top\big)_j, (Y_U)_i\big(J_{2n}Z^\top\big)_i\right)$$

$$= \sum_{i,j=1}^{2n}\omega\left((X_U)_j, (Y_U)_i\right)\left(Z_j Z_i + \big(J_{2n}Z^\top\big)_j\big(ZJ_{2n}^\top\big)_i\right).$$

The result follows by definition of $X_U(v)$. □

Remark 4.4 Owing to the inner product structure (2.3), the projection operator from Proposition 4.3 is orthogonal in the Frobenius norm since

$$\sum_{j=1}^{p}\omega_j\left(w - \Pi_{T_R\mathcal{M}_{2n}^{\mathrm{spl}}}w, y\right) = \langle w - \Pi_{T_R\mathcal{M}_{2n}^{\mathrm{spl}}}w, J_{2N}y\rangle = 0, \qquad \forall\, y \in T_R\mathcal{M}_{2n}^{\mathrm{spl}}.$$

This means that the projection gives the best low-rank approximation of the velocity vector, and hence the reduced dynamics is associated with the flow field ensuing from the best approximation in the tangent space to the reduced manifold.

To compute the initial condition of the reduced problem, we perform the complex SVD of $\mathcal{R}_0(\eta_h) \in \mathbb{R}^{2N\times p}$ truncated at the n-th mode. Then the initial value $U_0 \in \mathcal{M}$ is obtained from the resulting unitary matrix of left singular vectors of $\mathcal{R}_0(\eta_h)$ by exploiting the isomorphism between $\mathcal{M}$ and $\mathrm{St}(n, \mathbb{C}^N)$, *cf.* Lemma 4.8. The expansion coefficients matrix is initialized as $Z_0 = \mathcal{R}_0(\eta_h)^\top U_0$. Therefore, the dynamical system for the approximate reduced solution (4.2) reads: Find $R \in C^1(\mathcal{T}, \mathcal{M}_{2n}^{\mathrm{spl}})$ such that

$$\begin{cases} \dot{R}(t) = \Pi_{T_R\mathcal{M}_{2n}^{\mathrm{spl}}}\mathcal{X}_{\mathcal{H}}(R(t), \eta_h), & \text{for } t \in \mathcal{T}, \\ R(t_0) = U_0 Z_0^\top. \end{cases} \tag{4.9}$$

For any $1 \le j \le p$ and $t \in \mathcal{T}$, let $Z_j(t) \in \mathbb{R}^{1\times 2n}$ be the j-th row of the matrix $Z(t) \in V^{p\times 2n}$, and let $Y(t) := [Y_1|\dots|Y_p] \in \mathbb{R}^{2N\times p}$ where $Y_j := \nabla_{UZ_j^\top}\mathcal{H}(UZ_j^\top, \eta_j) \in$

$\mathbb{R}^{2N \times 1}$, and $\nabla_{U Z_j^\top}$ denotes the gradient with respect to $U Z_j^\top$. Using the decomposition $R = U Z^\top$ in (4.3), we can now derive from (4.9) evolution equations for U and Z: Given $\mathcal{R}_0(\eta_h) \in \mathbb{R}^{2N \times p}$, find $(U, Z) \in C^1(\mathcal{T}, \mathcal{M}) \times C^1(\mathcal{T}, V^{p \times 2n})$ such that

$$\begin{cases} \dot{Z}_j(t) = J_{2n} \nabla_{Z_j} \mathcal{H}(U Z_j^\top, \eta_j), & t \in \mathcal{T}, \ 1 \le j \le p, \\ \dot{U}(t) = (I_{2N} - U U^\top)(J_{2N} Y Z - Y Z J_{2n}^\top) S^{-1}, & t \in \mathcal{T}, \\ U(t_0) Z(t_0)^\top = U_0 Z_0^\top. \end{cases} \tag{4.10}$$

The reduced problem (4.10) is analogous to the system derived in [28, Proposition 6.9]. The evolution equations for the coefficients Z form a system of p equations in $2n$ unknowns and correspond to the Galerkin projection onto the space spanned by the columns of U, as obtained with a standard reduced basis method. Here, however, the projection is changing over time as the reduced basis U is evolving. For U fixed, the flow map characterizing the evolution of each Z_j, for $1 \le j \le p$, is a symplectomorphism (*cf.* Definition 2.2), i.e. the dynamics is canonically Hamiltonian. The evolution problem satisfied by the basis U is a matrix equation in $2N \times 2n$ unknowns on the manifold of orthosymplectic rectangular matrices introduced in Definition 3.2, as shown in the following result.

Proposition 4.5 *If $U(t_0) \in \mathcal{M}$ then $U(t) \in \mathbb{R}^{2N \times 2n}$ solution of (4.10) satisfies $U(t) \in \mathcal{M}$ for all $t \in \mathcal{T}$.*

Proof We first show that, for any matrix $W(t) \in \mathbb{R}^{2N \times 2n}$, if $W(t_0) \in \mathcal{M}$ and $\dot{W} \in H_W$, with H_W defined in (4.7), then $W(t) \in \mathcal{M}$ for any $t > t_0$. The condition $\dot{W}^\top W = 0$ implies $d_t(W^\top(t) W(t)) = \dot{W}^\top W + W^\top \dot{W} = 0$, hence $W^\top(t) W(t) = W^\top(t_0) W(t_0) = I_{2n}$ by the assumption on the initial condition. Moreover, the condition $\dot{W} = J_{2N}^\top \dot{W} J_{2n}$ together with the dynamical orthogonality $\dot{W}^\top W = 0$ results in $d_t(W^\top(t) J_{2N} W(t)) = \dot{W}^\top J_{2N} W + W^\top J_{2N} \dot{W} = J_{2n}^\top \dot{W}^\top W + W^\top \dot{W} J_{2n}^\top = 0$. Hence, the symplectic constraint on the initial condition yields $W^\top(t) J_{2N} W(t) = W^\top(t_0) J_{2N} W(t_0) = J_{2n}$.

Owing to the reasoning above, we only need to verify that the solution of (4.10) satisfies $\dot{U} \in H_U$. The dynamical orthogonal condition $\dot{U}^\top U = 0$ is trivially satisfied. Moreover, since $S J_{2n} = J_{2n} S$, the constraint $\dot{U} = J_{2N}^\top \dot{U} J_{2n}$ is satisfied if $\dot{U} S J_{2n}^\top = J_{2N}^\top \dot{U} S$. One can easily show that $A := J_{2N} Y Z - Y Z J_{2n}^\top = J_{2N} A J_{2n}^\top$. Therefore, $\dot{U} S J_{2n}^\top = (I_{2N} - U U^\top) A J_{2n}^\top = J_{2N}^\top (I_{2N} - U U^\top) J_{2N} A J_{2n}^\top = J_{2N}^\top \dot{U} S$. $\qquad \square$

Remark 4.6 Observe that the dynamical reduced basis technique proposed in the previous Section can be extended to more general Hamiltonian systems endowed with a degenerate constant Poisson structure. The idea is to proceed as in [16, Sect. 3] by splitting the dynamics into the evolution on a symplectic submanifold of the phase space and the trivial evolution of the Casimir invariants. The symplectic dynamical model order reduction developed in Sect. 4 can then be performed on the symplectic component of the dynamics.

4.2 Conservation properties of the reduced dynamics

The velocity field of the reduced flow (4.9) is the symplectic projection of the full model velocity onto the tangent space of the reduced manifold. For any fixed parameter $\eta_j \in \Gamma_h$, let $\mathcal{H}_j := \mathcal{H}(\cdot, \eta_j)$. In view of Proposition 4.3, the reduced solution $R \in C^1(\mathcal{T}, \mathcal{M}_{2n}^{\mathrm{spl}})$ satisfies the symplectic variational principle

$$\sum_{j=1}^{p} \omega_j \big(\dot{R} - J_{2N} \nabla \mathcal{H}_j(R), y \big) = 0, \qquad \forall \, y \in T_R \mathcal{M}_{2n}^{\mathrm{spl}}.$$

This implies that the Hamiltonian $\mathcal{H}$ is a conserved quantity of the continuous reduced problem (4.10). Indeed,

$$\sum_{j=1}^{p} \frac{d}{dt} \mathcal{H}_j(R(t)) = \sum_{j=1}^{p} \big(\nabla_{R_j} \mathcal{H}_j(R), \dot{R}_j \big) = \sum_{j=1}^{p} \omega \big(J_{2N} \nabla_{R_j} \mathcal{H}_j(R), \dot{R}_j \big)$$

$$= \sum_{j=1}^{p} \omega_j(\dot{R}, \dot{R}) = 0.$$

Therefore, if $\mathcal{R}_0(\eta_h) \in \mathrm{span}\{U_0\}$ then the Hamiltonian is preserved,

$$\sum_{j=1}^{p} \big(\mathcal{H}_j(\mathcal{R}(t)) - \mathcal{H}_j(R(t)) \big) = \sum_{j=1}^{p} \big(\mathcal{H}_j(\mathcal{R}_0) - \mathcal{H}_j(R(t_0)) \big)$$

$$= \sum_{j=1}^{p} \Big(\mathcal{H}_j(\mathcal{R}_0) - \mathcal{H}_j(U_0 U_0^\top \mathcal{R}_0) \Big).$$

To deal with the other invariants of motion, let us assume for simplicity that $p = 1$. Since the linear map $\mathbb{R}^{2N} \to \mathrm{span}\{U(t)\}$ associated with the reduced basis at any time $t \in \mathcal{T}$ cannot be symplectic, the invariants of motion of the full and reduced model cannot be in one-to-one correspondence. Nevertheless, a result analogous to [16, Lemma 3.9] holds.

Lemma 4.7 *Let $\pi_{+,t}^*$ be the pullback of the linear map associated with the reduced basis $U^\top(t)$ at time $t \in \mathcal{T}$. Assume that $\mathcal{H} \in \mathrm{Im}(\pi_{+,t}^*)$ for any $t \in \mathcal{T}$. Then, $\mathcal{I}(t) \in C^\infty(\mathbb{R}^{2n})$ is an invariant of $\Phi_{X_{\pi_{+,t}^* \mathcal{H}}}^t$ if and only if $(\pi_{+,t}^* \mathcal{I})(t) \in C^\infty(\mathbb{R}^{2N})$ is an invariant of $\Phi_{X_\mathcal{H}}^t$ in $\mathrm{Im}(\pi_{+,t}^*)$.*

4.3 Convergence estimates with respect to the best low-rank approximation

In order to derive error estimates for the reduced solution of problem (4.9), we extend to our setting the error analysis of [14, Sect. 5] which shows that the error committed by the dynamical approximation with respect to the best low-rank approximation is

bounded by the projection error of the full model solution onto the reduced manifold of low-rank matrices. To this aim, we resort to the isomorphism between the reduced symplectic manifold $\mathcal{M}_{2n}^{\mathrm{spl}}$ defined in (4.3) and the manifold $\mathcal{M}_n$ of rank-n complex matrices, already established in [28, Lemma 6.1]. Then, we derive the dynamical orthogonal approximation of the resulting problem in the complex setting and prove that it is isomorphic to the solution of the reduced Hamiltonian system (4.9). The differentiability properties of orthogonal projections onto smooth embedded manifolds and the trivial extension to complex matrices of the curvature bounds in [14] allows to derive an error estimate.

Let $\mathfrak{L}(\Omega)$ denote the set of functions with values in the vector space Ω, and let $\mathfrak{F} : \mathfrak{L}(\mathbb{R}^{2N \times p}) \to \mathfrak{L}(\mathbb{C}^{N \times p})$ be the isomorphism

$$R(\cdot) = \begin{pmatrix} R_q(\cdot) \\ R_p(\cdot) \end{pmatrix} \longmapsto \mathfrak{F}(R)(\cdot) = R_q(\cdot) + i R_p(\cdot). \tag{4.11}$$

Then, problem (4.1) can be recast in the complex setting as: For $\mathcal{R}_0(\eta_h) \in \mathbb{R}^{2N \times p}$, find $\mathcal{C} \in C^1(\mathcal{T}, \mathbb{C}^{N \times p})$ such that

$$\begin{cases} \dot{\mathcal{C}}(t) = \mathfrak{F}(\mathcal{X}_{\mathcal{H}})(\mathcal{C}(t), \eta_h) =: \widehat{\mathcal{X}}_{\mathcal{H}}(\mathcal{C}(t), \eta_h), & \text{for } t \in \mathcal{T}, \\ \mathcal{C}(t_0) = \mathfrak{F}(\mathcal{R}_0)(\eta_h). \end{cases} \tag{4.12}$$

Similarly to dynamically orthogonal approximations we consider the manifold of rank-n complex matrices $\mathcal{M}_n := \{C \in \mathbb{C}^{N \times p} : \mathrm{rank}(C) = n\}$. Any $C \in \mathcal{M}_n$ can be decomposed, up to unitary $n \times n$ transformations, as $C = WY^\top$ where $W \in \mathrm{St}(n, \mathbb{C}^N) = \{M \in \mathbb{C}^{N \times n} : M^{\mathsf{H}}M = I_n\}$, and $Y \in \mathcal{V}^{p \times n} := \{M \in \mathbb{C}^{p \times n} : \mathrm{rank}(M) = n\}$. Analogously to [28, Lemma 6.1] one can establish the following result.

Lemma 4.8 *The manifolds $\mathcal{M}_n$ and $\mathcal{M}_{2n}^{\mathrm{spl}}$ are isomorphic via the map*

$$(U, Z) \in \mathcal{M} \times V^{p \times 2n} \longmapsto \left(\mathfrak{F}(A), \mathfrak{F}(Z^\top)^\top \right) \in \mathrm{St}\left(n, \mathbb{C}^N\right) \times \mathcal{V}^{p \times n}, \tag{4.13}$$

where $\mathfrak{F}$ is defined in (4.11) and $A \in \mathbb{R}^{2N \times n}$ is such that $U = [A \mid J_{2N}^\top A]$ in view of Lemma 3.3.

For $C(t_0) \in \mathcal{M}_n$ associated with $R(t_0) \in \mathcal{M}_{2n}^{\mathrm{spl}}$ via the map (4.13), we can therefore derive the DO dynamical system: find $C \in C^1(\mathcal{T}, \mathcal{M}_n)$ such that

$$\dot{C}(t) = \Pi_{T_C \mathcal{M}_n} \widehat{\mathcal{X}}_{\mathcal{H}}(C(t), \eta_h), \qquad \text{for } t \in \mathcal{T}, \tag{4.14}$$

where $\Pi_{T_C \mathcal{M}_n}$ is the projection onto the tangent space of $\mathcal{M}_n$ at $C = WY^\top$, defined as

$$T_C \mathcal{M}_n = \Big\{ X \in \mathbb{C}^{N \times p} : X = X_W Y^\top + W X_Y^\top \text{ with } X_Y \in \mathbb{C}^{p \times n},$$
$$X_W \in \mathbb{C}^{N \times n}, X_W^{\mathsf{H}} W + W^{\mathsf{H}} X_W = 0 \Big\}.$$

The so-called dynamically orthogonal condition $X_W^H W = 0$, allows to uniquely parameterize the tangent space $T_C \mathcal{M}_n$ by imposing that the complex reduced basis evolves orthogonally to itself.

Let M^* indicate the complex conjugate of a given matrix M. The projection onto the tangent space of $\mathcal{M}_n$ can be characterized as in the following result.

Lemma 4.9 *At every $C = W Y^\top \in \mathcal{M}_n$, the map*

$$\Pi_{T_C \mathcal{M}_n} : \mathbb{C}^{N \times p} \longrightarrow T_C \mathcal{M}_n$$
$$w \longmapsto (I_N - W W^H) w \, Y^* (Y^\top Y^*)^{-1} Y^\top + W W^H w, \quad (4.15)$$

is the $\|\cdot\|$-orthogonal projection onto the tangent space of $\mathcal{M}_n$ at C.

Proof The result can be derived similarly to the proof of [14, Proposition 7] by minimizing the convex functional $\mathfrak{J}(X_W, X_Y) := \frac{1}{2} \| w - X_W Y^\top - W X_Y^\top \|^2$ under the constraint $X_W^H W = 0$. $\qquad\square$

Using the expression (4.15) for the projection onto the tangent space of $\mathcal{M}_n$, we can derive from (4.14) evolution equations for the terms W and Y: Given $C_0 = \Pi_{\mathcal{M}_n} \mathcal{C}(t_0) \in \mathbb{C}^{N \times p}$ orthogonal projection onto $\mathcal{M}_n$, find $(W, Y) \in C^1(\mathcal{T}, \mathrm{St}(n, \mathbb{C}^N)) \times C^1(\mathcal{T}, \mathcal{V}^{p \times n})$ such that

$$\begin{cases} \dot{Y}^*(t) = \widehat{\mathcal{X}}_{\mathcal{H}}^H (W Y^\top, \eta_h) W, & t \in \mathcal{T}, \\ \dot{W}^*(t) = (I_N - W^* W^\top) \widehat{\mathcal{X}}_{\mathcal{H}}^* (W Y^\top, \eta_h) Y (Y^H Y)^{-1}, & t \in \mathcal{T}. \end{cases} \quad (4.16)$$

Proposition 4.10 *Under the assumption of well-posedness, problem (4.9) is equivalent to problem (4.14).*

Proof The proof easily follows from algebraic manipulations of the field equations (4.10) and (4.16) and from the definition of the isomorphism (4.13). $\qquad\square$

In view of Proposition 4.10, we can revert to the error estimate established in [14].

Theorem 4.11 ([14, Theorem 32]). *Let $\mathcal{C} \in C^1(\mathcal{T}, \mathbb{C}^{N \times p})$ denote the exact solution of (4.12) and let $C \in C^1(\mathcal{T}, \mathcal{M}_n)$ be the solution of (4.14) at time $t \in \mathcal{T}$. Assume that no crossing of the singular values of $\mathcal{C}$ occurs, namely*

$$\sigma_n(\mathcal{C}(t)) > \sigma_{n+1}(\mathcal{C}(t)), \qquad \forall t \in \mathcal{T}.$$

Let $\Pi_{\mathcal{M}_n}$ be the $\|\cdot\|$-orthogonal projection onto $\mathcal{M}_n$. Then, at any time $t \in \mathcal{T}$, the error between the approximate solution $C(t)$ and the best rank-n approximation of $\mathcal{C}(t)$ can be bounded as

$$\| C(t) - \Pi_{\mathcal{M}_n} \mathcal{C}(t) \| \le \int_{\mathcal{T}} \nu \| \mathcal{C}(s) - \Pi_{\mathcal{M}_n} \mathcal{C}(s) \| e^{\mu(t-s)} \, ds,$$

where $\mu \in \mathbb{R}$ and $\nu \in \mathbb{R}$ are defined as

$$\mu := L_{\mathcal{X}} + 2 \sup_{t \in \mathcal{T}} \frac{\| \mathcal{X}_{\mathcal{H}}(\mathcal{C}(t), \eta_h) \|}{\sigma_n(\mathcal{C}(t))}, \qquad \nu := L_{\mathcal{X}} + \frac{\| \mathcal{X}_{\mathcal{H}}(\mathcal{C}(s), \eta_h) \|}{\sigma_n(\mathcal{C}(s)) - \sigma_{n+1}(\mathcal{C}(s))},$$

and $L_\mathcal{X} \in \mathbb{R}$ denotes the Lipschitz continuity constant of $\mathcal{X}_\mathcal{H}$.

The remainder of this work pertains to numerical methods for the temporal discretization of the reduced dynamics (4.10). Since we consider splitting techniques, see e.g. [15, Sect. II.5], the evolution problems for the expansion coefficients and for the reduced basis are examined separately. The coefficients $Z(t) \in V^{p \times 2n}$ of the expansion (4.2) satisfy a Hamiltonian dynamical system (4.10) in the reduced symplectic manifold of dimension $2n$ spanned by the evolving orthosymplectic basis $U(t) \in \mathcal{M}$. The numerical approximation of the evolution equation for $Z(t)$ can, thus, be performed using symplectic integrators, *cf.* [15, Sect. VI]. Observe that the use of standard splitting techniques might require the approximate reduced solution, at a given time step, to be projected into the space spanned by the updated basis. This might cause an error in the conservation of the invariants due to the projection step, that, however, can be controlled under sufficiently small time steps. In principle, exact conservation can be guaranteed if the evolution of the reduced basis evolves smoothly at the interface of temporal interval (or temporal subintervals associated with the splitting), or, in other words, if the splitting is synchronous and the two systems are concurrently advanced in time. We postpone to future work the investigation and the numerical study of splitting methods that exactly preserve the Hamiltonian.

5 Numerical methods for the evolution of the reduced basis

Contrary to global projection-based model order reduction, dynamical reduced basis methods eschew the standard online-offline paradigm. The construction and evolution of the local reduced basis (4.10) does not require queries of the high-fidelity model so that the method does not incur a computationally expensive offline phase. However, the evolution of the reduced basis entails the solution of a matrix equation in which one dimension equals the size of the full model. Numerical methods for the solution of (4.10) will have arithmetic complexity $\min\{C_\mathcal{R}, C_\mathcal{F}\}$ where $C_\mathcal{F}$ is the computational cost required to evaluate the velocity field of (4.10), and $C_\mathcal{R}$ denotes the cost associated with all other operations. Assume that the cost to evaluate the Hamiltonian at the reduced solution has order $O(\alpha(N))$. Then, a standard algorithm for the evaluation of the right hand side of (4.10) will have arithmetic complexity $C_\mathcal{F} = O(\alpha(N)) + O(Nn^2) + O(Npn) + O(n^3)$, where the last two terms are associated with the computation of YZ, and the inversion of $C + J_{2n}^\top C J_{2n}$, respectively. This Section focuses on the development of structure-preserving numerical methods for the solution of (4.10) such that $C_\mathcal{R}$ is at most linear in N. The efficient treatment of the nonlinear terms is out of the scope of the present study and will be the subject of future investigations on structure-preserving hyper-reduction techniques.

To simplify the notation, we recast (4.10) as: For $Q \in \mathcal{M}$, find $U \in C^1(\mathcal{T}, \mathbb{R}^{2N \times 2n})$ such that

$$\begin{cases} \dot{U}(t) = \mathcal{F}(U(t)), & \text{for } t \in \mathcal{T}, \\ U(t_0) = Q, \end{cases} \tag{5.1}$$

where, for any fixed $t \in \mathcal{T}$,

$$\mathcal{F}(U) := \left(I_{2N} - UU^\top\right)\left(J_{2N}YZ - YZJ_{2n}^\top\right)\left(Z^\top Z + J_{2n}^\top Z^\top ZJ_{2n}\right)^{-1}. \quad (5.2)$$

Observe that $\mathcal{F} : U \in \mathcal{M} \mapsto \mathcal{F}(U) \in H_U \subset T_U\mathcal{M}$, where H_U is defined as in (4.7), and $T_U\mathcal{M} = \{V \in \mathbb{R}^{2N \times 2n} : U^\top V \in \mathfrak{g}_{2n}\}$. In a temporal splitting perspective, we assume that the matrix $Z(t) \in V^{p \times 2n}$ is given at each time instant $t \in \mathcal{T}$. Owing to Proposition 4.5, if $Q \in \mathcal{M}$, then $U(t) \in \mathcal{M}$ for all $t \in \mathcal{T}$. Then, the goal is to develop an efficient numerical scheme such that the discretization of (5.1) yields an approximate flow map with trajectories belonging to $\mathcal{M}$.

We propose two intrinsic numerical methods for the solution of the differential equation (5.1) within the class of numerical methods based on local charts on manifolds [15, Sect. IV.5]. The analyticity and the favorable computational properties of the Cayley transform, *cf.* Proposition 5.2 and [17], makes it our choice as coordinate map on the orthosymplectic matrix manifold.

5.1 Cayley transform as coordinate map

Orthosymplectic square matrices form a subgroup $\mathcal{U}(2N)$ of a quadratic Lie group. We can, therefore, use the Cayley transform to induce a local parameterization of the Lie group $\mathcal{U}(2N)$ near the identity, with the corresponding Lie algebra as parameter space. The following results extend to orthosymplectic matrices the properties of the Cayley transform presented in e.g. [15, Sect. IV.8.3].

Lemma 5.1 *Let $\mathcal{G}_{2N}$ be the group of orthosymplectic square matrices and let $\mathfrak{g}_{2N}$ be the corresponding Lie algebra. Let* cay $: \mathfrak{g}_{2N} \to \mathbb{R}^{2N \times 2N}$ *be the Cayley transform defined as*

$$\mathrm{cay}(\Omega) = \left(I - \frac{\Omega}{2}\right)^{-1}\left(I + \frac{\Omega}{2}\right), \quad \forall\, \Omega \in \mathfrak{g}_{2N}. \quad (5.3)$$

Then,

(i) cay *maps the Lie algebra $\mathfrak{g}_{2N}$ into the Lie group $\mathcal{G}_{2N}$.*
(ii) cay *is a diffeomorphism in a neighborhood of the zero matrix $0 \in \mathfrak{g}_{2N}$. The differential of* cay *at $\Omega \in \mathfrak{g}_{2N}$ is the map* $\mathrm{dcay}_\Omega : T_\Omega\mathfrak{g}_{2N} \cong \mathfrak{g}_{2N} \to T_{\mathrm{cay}(\Omega)}\mathcal{G}_{2N}$,

$$\mathrm{dcay}_\Omega(A) = \left(I - \frac{\Omega}{2}\right)^{-1} A \left(I + \frac{\Omega}{2}\right)^{-1}, \quad \forall\, A \in \mathfrak{g}_{2N},$$

and its inverse is

$$\mathrm{dcay}_\Omega^{-1}(A) = \left(I - \frac{\Omega}{2}\right) A \left(I + \frac{\Omega}{2}\right), \quad \forall\, A \in T_{\mathrm{cay}(\Omega)}\mathcal{G}_{2N}. \quad (5.4)$$

(iii) [12, Theorem 3] *Let $\sigma(A)$ denote the spectrum of $A \in \mathbb{R}^{2N \times 2N}$. If $\Omega \in C^1(\mathbb{R}, \mathfrak{g}_{2N})$ then $A := \mathrm{cay}(\Omega) \in C^1(\mathbb{R}, \mathcal{G}_{2N})$. Conversely, if $A \in C^1(\mathbb{R}, \mathcal{G}_{2N})$*

and $-1 \notin \bigcup_{t \in \mathbb{R}} \sigma(A(t))$ *then there exists a unique* $\Omega \in C^1(\mathbb{R}, \mathfrak{g}_{2N})$ *such that* $\Omega = \mathrm{cay}^{-1}(A) = 2(A - I_{2N})(A + I_{2N})^{-1}$.

Proof Let $\Omega \in \mathfrak{g}_{2N}$ and let $\overline{\Omega} := \Omega/2$. Since Ω is skew-symmetric then $I - \overline{\Omega}$ is invertible.

(i) The Cayley transform defined in (5.3) can be recast as

$$\begin{aligned}
\mathrm{cay}(\Omega) &= -(I - \overline{\Omega})^{-1}(-2I + (I - \overline{\Omega})) = 2(I - \overline{\Omega})^{-1} - I \\
&= -(-2I + (I - \overline{\Omega}))(I - \overline{\Omega})^{-1} = (I + \overline{\Omega})(I - \overline{\Omega})^{-1}.
\end{aligned} \tag{5.5}$$

Then, using (5.5) and the skew-symmetry of $\Omega \in \mathfrak{g}_{2N}$ results in

$$\begin{aligned}
\mathrm{cay}(\Omega)^\top \mathrm{cay}(\Omega) &= (I - \overline{\Omega})^{-\top}(I + \overline{\Omega}^\top \overline{\Omega})(I - \overline{\Omega})^{-1} \\
&= (I - \overline{\Omega})^{-\top}(I - \overline{\Omega} - \overline{\Omega}^\top + \overline{\Omega}^\top \overline{\Omega})(I - \overline{\Omega})^{-1} = I.
\end{aligned}$$

Moreover, $\mathrm{cay}(\Omega) J_{2N} = J_{2N} \mathrm{cay}(\Omega)$ since

$$\begin{aligned}
\mathrm{cay}(\Omega) J_{2N} &= (I + \overline{\Omega})(-J_{2N} + \overline{\Omega} J_{2N})^{-1} = (I + \overline{\Omega})(-J_{2N} + J_{2N} \overline{\Omega})^{-1} \\
&= (I + \overline{\Omega}) J_{2N} (I - \overline{\Omega})^{-1} = (J_{2N} - J_{2N} \overline{\Omega}^\top)(I - \overline{\Omega})^{-1} \\
&= (J_{2N} + J_{2N} \overline{\Omega})(I - \overline{\Omega})^{-1} = J_{2N} \mathrm{cay}(\Omega).
\end{aligned}$$

(ii) The map cay (5.3) has non-zero derivative at $0 \in \mathfrak{g}_{2N}$. Therefore, by the inverse function theorem, it is a diffeomorphism in a neighborhood of $0 \in \mathfrak{g}_{2N}$. Standard rules of calculus yield the expression (5.4), *cf.* [15, Sect. IV.8.3, Lemma 8.8]. □

The factor $1/2$ in the definition (5.3) of the Cayley transform is arbitrary and has been introduced to guarantee that $\mathrm{dcay}_0 = I_{2N}$, which will be used in Sect. 5.3 for the construction of retraction maps.

To derive computationally efficient numerical schemes for the solution of the basis evolution equation (5.1) we exploit the properties of analytic functions evaluated at the product of rectangular matrices.

Proposition 5.2 *Let* $\Omega \in \mathfrak{g}_{2N}$ *and* $Y \in \mathbb{R}^{2N \times r}$. *If* Ω *has rank* $k \le 2N$, *then* $\mathrm{cay}(\Omega)Y$ *can be evaluated with computational complexity of order* $O(Nrk) + O(k^2 r) + O(k^3)$.

Proof Since Ω has rank k it admits the splitting $\Omega = \alpha \beta^\top$ for some $\alpha, \beta \in \mathbb{R}^{2N \times k}$. To evaluate the Cayley transform in a computationally efficient way we exploit the properties of analytic functions of low-rank matrices. More in details, let $f(z) := z^{-1}(\mathrm{cay}(z) - 1)$ for any $z \in \mathbb{C}$. The function f has a removable pole at $z = 0$. Its analytic extension reads,

$$f(z) = \sum_{m=0}^{\infty} 2^{-m} z^m.$$

For any $m \in \mathbb{N} \setminus \{0\}$ it holds $\Omega^m = (\alpha\beta^\top)^m = \alpha(\beta^\top\alpha)^{m-1}\beta^\top$. Hence,

$$
\begin{aligned}
\text{cay}(\Omega) &= I_{2N} + \sum_{m=1}^{\infty} 2^{1-m}\Omega^m = I_{2N} + \sum_{m=1}^{\infty} 2^{1-m}\alpha(\beta^\top\alpha)^{m-1}\beta^\top \\
&= I_{2N} + \alpha f(\beta^\top\alpha)\beta^\top.
\end{aligned}
$$

The cost to compute $A := \beta^\top\alpha \in \mathbb{R}^{k\times k}$ is $O(Nk^2)$. Moreover,

$$
\text{cay}(\Omega)Y = (I_N + \alpha f(\beta^\top\alpha)\beta^\top)Y = Y + \alpha(\beta^\top\alpha)^{-1}\big(\text{cay}(\beta^\top\alpha) - I_k\big)\beta^\top Y.
$$

The evaluation of $f(A) = A^{-1}(\text{cay}(A) - I_k) \in \mathbb{R}^{k\times k}$ requires $O(k^3)$ operations. Finally, the matrix multiplications $\alpha f(A)\beta^\top Y$ can be performed in $O(Nrk) + O(k^2 r)$ operations.

The approach suggested hitherto is clearly not unique. The invertibility of the matrix A is ensured under the condition that the low-rank factors α and β are full rank. Although a low-rank decomposition with full rank factors is achievable [8, Proposition 4], one could alternatively envision the use of Woodbury matrix identity [33] to compute the matrix inverse appearing in the definition (5.3) of the Cayley transform. This yields the formula

$$
\text{cay}(\Omega)Y = Y + \frac{1}{2}\alpha\big(\text{cay}(\beta^\top\alpha) + I_k\big)\beta^\top Y = Y - \alpha\left(\frac{1}{2}\beta^\top\alpha - I_k\right)^{-1}\beta^\top Y,
$$

which can also be evaluated in $O(Nrk) + O(k^2 r) + O(k^3)$ operations. $\qquad\square$

5.2 Numerical integrators based on Lie groups acting on manifolds

In this Section we propose a numerical scheme for the solution of (5.1) based on Lie group methods, *cf.* [18]. The idea is to consider $\mathcal{M}$ as a manifold acted upon by the Lie group $\mathcal{G}_{2N} = \mathcal{U}(2N)$ of square orthosymplectic matrices. Then, since the local structure in a neighbourhood of any point of $\mathcal{G}_{2N}$ can be described by the corresponding Lie algebra $\mathfrak{g}_{2N}$, a local coordinate map is employed to derive a differential equation on $\mathfrak{g}_{2N}$. Since Lie algebras are linear spaces, using Runge–Kutta methods to solve the equation on $\mathfrak{g}_{2N}$ allows to derive discrete trajectories that remain on the Lie algebra. This approach falls within the class of numerical integration schemes based on canonical coordinates of the first kind, also known as Runge–Kutta Munthe-Kaas (RK-MK) methods [24–27].

Proposition 5.3 *The evolution equation (5.1) with arbitrary $\mathcal{F} : \mathcal{M} \to T\mathcal{M}$ is equivalent to the problem: For $Q \in \mathcal{M}$, find $U \in C^1(\mathcal{T}, \mathcal{M})$ such that*

$$
\begin{cases} \dot{U}(t) = \mathcal{L}(U(t))\, U(t), & \text{for } t \in \mathcal{T}, \\ U(t_0) = Q, \end{cases} \tag{5.6}
$$

with $\mathcal{L} : \mathcal{M} \rightarrow \mathfrak{g}_{2N}$ defined as

$$\mathcal{L}(U) := \frac{1}{2}\left(\mathcal{S}(U) + J_{2N}^{\top}\mathcal{S}(U)J_{2N}\right), \tag{5.7}$$

where $\mathcal{S}(U) := (I_{2N} - UU^{\top})\mathcal{F}(U)U^{\top} - U\mathcal{F}(U)^{\top}$. Furthermore, if $\mathcal{F}(U) \in H_U$, for any $U \in \mathcal{M}$, then,

$$\mathcal{L}(U) = \mathcal{F}(U)U^{\top} - U\mathcal{F}(U)^{\top}. \tag{5.8}$$

Proof Let us consider, at each time $t \in \mathcal{T}$, an orthosymplectic extension $Y(t) \in \mathbb{R}^{2N \times 2N}$ of $U(t)$ by the matrix $W(t) \in \mathbb{R}^{2N \times 2(N-n)}$, such that $Y(t) = [U(t) \mid W(t)] \in \mathcal{U}(2N)$. Since Y is orthosymplectic by construction, it holds

$$0 = \frac{d}{dt}\left(Y^{\top}Y\right) = \dot{Y}^{\top}Y + Y^{\top}\dot{Y}, \qquad \implies \quad \dot{Y} = -Y\dot{Y}^{\top}Y,$$

$$0 = \frac{d}{dt}\left(Y^{\top}J_{2N}Y\right) = \dot{Y}^{\top}J_{2N}Y + Y^{\top}J_{2N}\dot{Y}, \quad \implies \quad \dot{Y} = -J_{2N}^{\top}Y\dot{Y}^{\top}J_{2N}Y.$$

It follows that $\dot{Y}(t) = \mathcal{A}(Y, \dot{Y})Y(t)$, for all $t \in \mathcal{T}$, with

$$\mathcal{A}(Y, \dot{Y}) := -\frac{1}{2}\left(Y\dot{Y}^{\top} + J_{2N}^{\top}Y\dot{Y}^{\top}J_{2N}\right) \in \mathfrak{g}_{2N},$$

and $Y\dot{Y}^{\top} = U\dot{U}^{\top} + W\dot{W}^{\top}$. Expressing $\mathcal{A}(Y, \dot{Y})$ explicitly in terms of U and W, and using the evolution equation satisfied by U, yields

$$\mathcal{A}(Y, \dot{Y}) = -\frac{1}{2}\left(U\mathcal{F}(U)^{\top} + J_{2N}^{\top}U\mathcal{F}(U)^{\top}J_{2N} + W\dot{W}^{\top} + J_{2N}^{\top}W\dot{W}^{\top}J_{2N}\right). \tag{5.9}$$

Moreover, since $\mathcal{A}$ is skew-symmetric, it holds

$$\dot{W}W^{\top} + W\dot{W}^{\top} = -U\mathcal{F}(U)^{\top} - \mathcal{F}(U)U^{\top}. \tag{5.10}$$

If $W \in \mathbb{R}^{2N \times 2(N-n)}$ is such that $\dot{W}W^{\top} = -U\mathcal{F}(U)^{\top}(I_{2N} - UU^{\top})$, then (5.10) it satisfied, owing to the fact that $\mathcal{F}(U) \in T_U\mathcal{M}$. Substituting this expression in (5.9) yields expression (5.7) with $\mathcal{L}(U) = \mathcal{A}(Y, \dot{Y})$.

Finally, if $\mathcal{F}(U)$ belongs to H_U then $U^{\top}\mathcal{F}(U) = 0$. Substituting in (5.7) and using the fact that $J_{2N}^{\top}\mathcal{F}(U)U^{\top}J_{2N} = \mathcal{F}(U)U^{\top}$ yields (5.8). □

Once we have recast (5.1) into the equivalent problem (5.6), the idea is to derive an evolution equation on the Lie algebra $\mathfrak{g}_{2N}$ via a coordinate map. A coordinate map of the first kind is a smooth function $\psi : \mathfrak{g}_{2N} \rightarrow \mathcal{G}_{2N}$ such that $\psi(0) = \text{Id} \in \mathcal{G}_{2N}$ and $d\psi_0 = \text{Id}$, where $d\psi : \mathfrak{g}_{2N} \times \mathfrak{g}_{2N} \rightarrow \mathfrak{g}_{2N}$ is the right trivialized tangent of ψ defined as

$$\frac{d}{dt}\psi(A(t)) = d\psi_{A(t)}(\dot{A}(t))\psi(A(t)), \qquad \forall A : \mathbb{R} \rightarrow \mathfrak{g}_{2N}. \tag{5.11}$$

For sufficiently small $t \geq t_0$, the solution of (5.6) is given by $U(t) = \psi(\Omega(t))U(t_0)$ where $\Omega(t) \in \mathfrak{g}_{2N}$ satisfies

$$\begin{cases} \dot{\Omega}(t) = \mathrm{d}\psi_{\Omega(t)}^{-1}\big(\mathcal{L}(U(t))\big), & \text{for } t \in \mathcal{T}, \\ \Omega(t_0) = 0. \end{cases} \tag{5.12}$$

Problem (5.12) can be solved using traditional RK methods. Let $(b_i, a_{i,j})$, for $i = 1, \ldots, s$ and $j = 1, \ldots, s$, be the coefficients of the Butcher tableau describing an s-stage explicit RK method. Then, the numerical approximation of (5.12) in the interval $(t^m, t^{m+1}]$ is performed as in Algorithm 1.

Algorithm 1 Explicit RK-MK scheme in $(t^m, t^{m+1}]$

Input: $U_m \in \mathcal{M}, \{b_i\}_{i=1}^s, \{a_{i,j}\}_{i,j=1}^s$
1: $\Omega_m^1 = 0, U_m^1 = U_m$
2: **for** $i = 2, \ldots, s$ **do**
3: $\quad \Omega_m^i = \Delta t \sum_{j=1}^{i-1} a_{i,j} \, \mathrm{d}\psi_{\Omega_m^j}^{-1}\big(\mathcal{L}(U_m^j)\big),$
4: $\quad U_m^i = \psi(\Omega_m^i)\, U_m,$
5: **end for**
6: $\Omega_{m+1} = \Delta t \sum_{i=1}^s b_i \, \mathrm{d}\psi_{\Omega_m^i}^{-1}\big(\mathcal{L}(U_m^i)\big),$
7: **return** $U_{m+1} = \psi(\Omega_{m+1})\, U_m \in \mathcal{M}$

As anticipated in Sect. 5.1, we resort to the Cayley transform as coordinate map in Algorithm 1. The use of the Cayley transform in the solution of matrix differential equations on Lie groups was proposed in [12,17,22]. Analogously to [12, Theorem 5], it can be shown that the invertibility of cay and dcay is guaranteed if $U(t) \in \mathcal{M}$ solution of (5.6) satisfies $-1 \notin \bigcup_{t \in \mathcal{T}} \sigma(U(t))$. Note that choosing a sufficiently small time step for the temporal integrator can prevent the numerical solution from having an eigenvalue close to -1, for some $t \in \mathcal{T}$. Alternatively, restarting procedures of the Algorithm 1 can be implemented similarly to [12, pp. 323, 324].

The computational cost of Algorithm 1 with $\psi = \text{cay}$ is assessed in the following result.

Proposition 5.4 *Consider the evolution problem (5.12) on a fixed temporal interval $(t^m, t^{m+1}] \subset \mathcal{T}$. Assume that the problem is solved with Algorithm 1 where the coordinate map ψ is given by the Cayley transform cay defined in (5.3). Then, the computational complexity of the resulting scheme is of order $O(Nn^2s^2) + O(n^3s^4) + C_{\mathcal{F}}$, where $C_{\mathcal{F}}$ is the complexity of the algorithm to compute $\mathcal{F}(U)$ in (5.2) at any given $U \in \mathcal{M}$.*

Proof We need to assess the computational cost of two operations in Algorithm 1: the evaluation of the map $\Lambda_i := \text{dcay}_{\Omega_m^i}^{-1}(\mathcal{L}(U_m^i))$ and the computation of $U_m^i = \text{cay}(\Omega_m^i)U_m$, for any $i = 2, \ldots, s$ and with $U_m \in \mathcal{M}$. First we prove that $\text{rank}(\Lambda_i) \leq$

$4n$. Observe that each term $\{\mathcal{L}(U_m^i)\}_{i=1}^s$, with $\mathcal{L}$ defined in (5.8), can be written as $\mathcal{L}(U_m^i) = \gamma_i \delta_i^\top$ where

$$\gamma_i := \left[\mathcal{F}(U_m^i) \mid -U_m^i\right] \in \mathbb{R}^{2N \times 4n}, \qquad \delta_i := \left[U_m^i \mid \mathcal{F}(U_m^i)\right] \in \mathbb{R}^{2N \times 4n}. \quad (5.13)$$

For $i = 1$, $\Omega_m^1 = 0$ and, hence, $\Lambda_1 = \mathcal{L}(U_m^1) = \gamma_1 \delta_1^\top$ owing to (5.13). Using definition (5.4), it holds

$$\Lambda_i = \text{dcay}_{\Omega_m^i}^{-1}(\gamma_i \delta_i^\top) = \left(I_{2N} - \frac{\Omega_m^i}{2}\right)\gamma_i \delta_i^\top \left(I_{2N} + \frac{\Omega_m^i}{2}\right) =: e_i f_i^\top, \quad \forall i \geq 2,$$

where $e_i, f_i \in \mathbb{R}^{2N \times 4n}$ are defined as

$$e_i := \left(I_{2N} - \frac{\Omega_m^i}{2}\right)\gamma_i, \qquad f_i := \left(I_{2N} + \frac{(\Omega_m^i)^\top}{2}\right)\delta_i.$$

Using Line 3 of Algorithm 1, the rank of Ω_m^i can be bounded as

$$\text{rank}(\Omega_m^i) = \text{rank}\left(\Delta t \sum_{j=1}^{i-1} a_{i,j}\Lambda_i\right) \leq \sum_{j=1}^{i-1} \text{rank}(\Lambda_i) \leq 4n(i-1),$$

and similarly $\text{rank}(\Omega_{m+1}) \leq 4ns$. Since the cost to compute each factor e_i, f_i is $O(Nn \, \text{rank}(\Omega_m^i))$, the computation of all Λ_i, for $1 \leq i \leq s$, requires $O(Nn^2 s^2)$ operations. Furthermore, in view of Proposition 5.2, each U_m^i can be computed with $O(Nn^2 i) + O(n^3 i^3)$ operations. Summing over the number s of stages of the Runge–Kutta scheme, the computational complexity of Algorithm 1 becomes $O(Nn^2 s^2) + O(n^3 s^4)$. □

In principle one can solve the evolution equation (5.12) on the Lie algebra $\mathfrak{g}_{2N}$ using the matrix exponential as coordinate map instead of the Cayley transform, in the spirit of [26]. However, there is no significant gain in terms of computational cost, as shown in details in Appendix A.

Although the computational complexity of Algorithm 1 is linear in the full dimension, it presents a suboptimal dependence on the number s of stages of the RK scheme. However, in practical implementations, the computational complexity of Proposition 5.4 might prove to be pessimistic in s, and might be mitigated with techniques that exploit the structure of the operators involved.

In the following Section we improve the efficiency of the numerical approximation of (5.1) by developing a scheme which is structure-preserving and has a computational cost $O(Nn^2 s)$, namely only linear in the dimension N of the full model and in the number s of RK stages.

5.3 Tangent methods on the orthosymplectic matrix manifold

In this Section we derive a tangent method based on retraction maps for the numerical solution of the reduced basis evolution problem (5.1). The idea of tangent methods is presented in [9, Sect. 2] and consists in expressing any $U(t) \in \mathcal{M}$ in a neighborhood of a given $Q \in \mathcal{M}$, via a smooth local map $\mathcal{R}_Q : T_Q\mathcal{M} \to \mathcal{M}$, as

$$U(t) = \mathcal{R}_Q(V(t)), \qquad V(t) \in T_Q\mathcal{M}. \tag{5.14}$$

Let $\mathcal{R}_Q$ be the restriction of a smooth map $\mathcal{R}$ to the fiber $T_Q\mathcal{M}$ of the tangent bundle. Assume that $\mathcal{R}_Q$ is defined in some open ball around $0 \in T_Q\mathcal{M}$, and $\mathcal{R}_Q(V) = Q$ if and only if $V \equiv 0 \in T_Q\mathcal{M}$. Moreover, let $\mathrm{d}\mathcal{R}_Q : TT_Q\mathcal{M} \cong T_Q\mathcal{M} \times T_Q\mathcal{M} \longrightarrow T\mathcal{M}$ be the (right trivialized) tangent of the map $\mathcal{R}_Q$, *cf.* definition (5.11). Let us fix the first argument of $\mathrm{d}\mathcal{R}_Q$ so that, for any $U, V \in \mathcal{M}$, the tangent map $\mathrm{d}\mathcal{R}_Q\big|_U : T_Q\mathcal{M} \to T_{\mathcal{R}_Q(U)}\mathcal{M}$ is defined as $\mathrm{d}\mathcal{R}_Q\big|_U (V) = \mathrm{d}\mathcal{R}_Q(U, V)$. Assume that the local rigidity condition $\mathrm{d}\mathcal{R}_Q\big|_0 = \mathrm{Id}_{T_Q\mathcal{M}}$ is satisfied. Under these assumptions, $\mathcal{R}$ is a retraction and, instead of solving the evolution problem (5.1) for U, one can derive the local behavior of U in a neighborhood of Q by evolving $V(t)$ in (5.14) in the tangent space of $\mathcal{M}$ at Q. Indeed, using (5.1) we can derive an evolution equation for $V(t)$ as

$$\dot{U}(t) = \mathrm{d}\mathcal{R}_Q\big|_{V(t)} (\dot{V}(t)) = \mathcal{F}\big(\mathcal{R}_Q(V(t))\big).$$

By the continuity of V and the local rigidity condition, the map $\mathrm{d}\mathcal{R}_Q\big|_{V(t)}$ is invertible for sufficiently small t (i.e., $V(t)$ sufficiently close to $0 \in T_Q\mathcal{M}$) and hence

$$\dot{V}(t) = \left(\mathrm{d}\mathcal{R}_Q\big|_{V(t)}\right)^{-1} \mathcal{F}\big(\mathcal{R}_Q(V(t))\big). \tag{5.15}$$

Since the initial condition is $U(t_0) = Q$ it holds $V(t_0) = 0 \in T_Q\mathcal{M}$.

This strategy allows to solve the ODE (5.15) on the tangent space $T\mathcal{M}$, which is a linear space, with a standard temporal integrator and then recover the approximate solution on the manifold $\mathcal{M}$ via the retraction map as in (5.14). If the retraction map can be computed exactly, this approach yields, by construction, a structure-preserving discretization. The key issue here is to build a suitable smooth retraction $\mathcal{R} : T\mathcal{M} \to \mathcal{M}$ such that its evaluation and the computation of the inverse of its tangent map can be performed exactly at a computational cost that depends only linearly on the dimension of the full model.

In order to locally solve the evolution problem (5.15) on the tangent space to the manifold $\mathcal{M}$ at a point $Q \in \mathcal{M}$ we follow a similar approach to the one proposed in [10] for the solution of differential equations on the Stiefel manifold. Observe that, for any $Q \in \mathcal{M}$, the velocity field $\mathcal{F}(Q)$ in (5.2), which describes the flow of the reduced basis on the manifold $\mathcal{M}$, belongs to the space H_Q defined in (4.7). We thus construct a retraction $\mathcal{R}_Q : H_Q \to \mathcal{M}$ as composition of three functions: a linear map Υ_Q from the space H_Q to the Lie algebra $\mathfrak{g}_{2N}$ associated with the Lie group $\mathcal{G}_{2N}$ acting on the

manifold $\mathcal{M}$, the Cayley transform (5.3) as coordinate map from the Lie algebra to the Lie group and the group action $\Lambda : \mathcal{G}_{2N} \times \mathcal{M} \to \mathcal{M}$,

$$\Lambda(G, Q) = \Lambda_Q(G) = G Q, \qquad \Lambda_Q : \mathcal{G}_{2N} \longrightarrow \mathcal{M},$$

that we take to be the matrix multiplication. This is summarized in the diagram below,
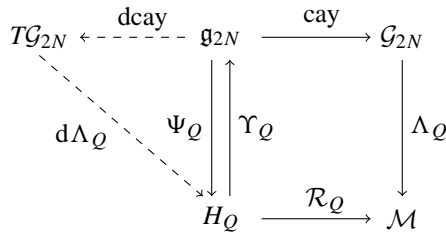
In more details, we take Υ_Q to be, for each $Q \in \mathcal{M}$, the linear map $\Upsilon_Q : H_Q \subset T_Q\mathcal{M} \to \mathfrak{g}_{2N}$ such that $\Psi_Q \circ \Upsilon_Q = \mathrm{Id}_{H_Q}$, where $\Psi_Q = \mathrm{d}\Lambda_Q\big|_e \circ \mathrm{dcay}_0$, and $T_Q\mathcal{M} = \{V \in \mathbb{R}^{2N \times 2n} : Q^\top V \in \mathfrak{g}_{2n}\}$. The space $T_Q\mathcal{M}$ can be characterized as follows.

Proposition 5.5 *Let $Q \in \mathcal{M}$ be arbitrary. Then, $V \in T_Q\mathcal{M}$ if and only if*

$$\exists\, \Theta \in \mathbb{R}^{2N \times 2n} \ \text{ with } \ Q^\top \Theta \in \mathfrak{sp}(2n) \ \text{ such that } \ V = (\Theta Q^\top - Q\Theta^\top)Q.$$

Proof ($\Longleftarrow$) Assume that $V \in \mathbb{R}^{2N \times 2n}$ is of the form $V = (\Theta Q^\top - Q\Theta^\top)Q$ for some $\Theta \in \mathbb{R}^{2N \times 2n}$ with $Q^\top \Theta \in \mathfrak{sp}(2n)$. To prove that $V \in T_Q\mathcal{M}$, we verify that $Q^\top V \in \mathfrak{g}_{2n}$. Using the orthogonality of Q, and the assumption $Q^\top \Theta \in \mathfrak{sp}(2n)$ results in

$$Q^\top V = Q^\top \left(\Theta Q^\top - Q\Theta^\top\right) Q = -Q^\top \left(Q\Theta^\top - \Theta Q^\top\right) Q = -V^\top Q.$$

$$Q^\top V J_{2n} = \left(Q^\top \Theta - \Theta^\top Q\right) J_{2n} = -J_{2n}\left(\Theta^\top Q - Q^\top \Theta\right) = -J_{2n} V^\top Q.$$

($\Longrightarrow$) Let $V \in T_Q\mathcal{M}$, i.e. $Q^\top V \in \mathfrak{g}_{2n}$. Let $\Theta := V + Q\big(S - \frac{Q^\top V}{2}\big)$ with $S \in \mathrm{Sym}(2n) \cap \mathfrak{sp}(2n)$ arbitrary. We first verify that $Q^\top \Theta \in \mathfrak{sp}(2n)$. Using the orthogonality of Q, the fact that $V \in T_Q\mathcal{M}$ and $S \in \mathfrak{sp}(2n)$ results in

$$Q^\top \Theta J_{2n} + J_{2n}\Theta^\top Q = \frac{Q^\top V}{2} J_{2n} + J_{2n}\frac{V^\top Q}{2} + S J_{2n} + J_{2n} S^\top$$
$$= S J_{2n} + J_{2n} S^\top = 0.$$

We then verify that, with the above definition of Θ, the matrix $(\Theta Q^\top - Q\Theta^\top)Q = \Theta - Q\Theta^\top Q$ coincides with V. Using the fact that $S \in \mathrm{Sym}(2n)$ and $V \in T_Q\mathcal{M}$ yields

$$\Theta - Q\Theta^\top Q = V + QS - Q\frac{Q^\top V}{2} - QV^\top Q - Q\left(S^\top - \frac{V^\top Q}{2}\right)$$
$$= V - Q\frac{Q^\top V}{2} - Q\frac{V^\top Q}{2} = V. \tag{5.16}$$

$\square$

We can therefore characterize the tangent space of the orthosymplectic matrix manifold as

$$T_Q\mathcal{M} = \{V \in \mathbb{R}^{2N \times 2n} : V = \left(\Theta^S_Q(V)Q^\top - Q\Theta^S_Q(V)^\top\right)Q,$$

$$\text{with } \Theta^S_Q(V) := V + Q\left(S - \frac{Q^\top V}{2}\right), \text{ for } S \in \mathrm{Sym}(2n) \cap \mathfrak{sp}(2n)\}.$$

This suggests that the linear map Υ_Q can be defined as

$$\Upsilon_Q : H_Q \longrightarrow \mathfrak{g}_{2N},$$
$$V \longmapsto \Theta^S_Q(V)Q^\top - Q\Theta^S_Q(V)^\top. \tag{5.17}$$

Indeed, since $\mathrm{d}\Lambda_Q\big|_e(G) = GQ$ and $\mathrm{dcay}_0 = I$, it holds $(\Psi_Q \circ \Upsilon_Q)(V) = \Upsilon_Q(V)Q = V$ for any $V \in H_Q$. This stems from the definition of Υ_Q in (5.17) since

$$\Psi_Q(\Upsilon_Q(V)) = \left(\Lambda_Q\big|_e \circ \mathrm{dcay}_0 \circ \Upsilon_Q\right)(V) = \Lambda_Q\big|_e(\Upsilon_Q(V))$$
$$= \Upsilon_Q(V)Q = \left(\Theta^S_Q(V)Q^\top - Q\Theta^S_Q(V)^\top\right)Q = V,$$

where the last equality follows by (5.16). Note that $\Psi_Q = \mathrm{d}\Lambda_Q\big|_e \circ \mathrm{dcay}_0$ is not injective as $\Upsilon_Q(H_Q)$ is a proper subspace of $\mathfrak{g}_{2N}$. Observe that, for any $V \in T_Q\mathcal{M}$, it holds $\Upsilon_Q(V) = VQ^\top - QV^\top + QV^\top QQ^\top$ and, hence, $\Upsilon_Q(V) \in \mathfrak{g}_{2N}$.

Proposition 5.6 *Let* $\mathrm{cay} : \mathfrak{g}_{2N} \to \mathcal{G}_{2N}$ *be the Cayley transform defined in* (5.3). *For any* $Q \in \mathcal{M}$ *and* $S \in \mathrm{Sym}(2n) \cap \mathfrak{sp}(2n)$, *we define*

$$\Theta^S_Q : T_Q\mathcal{M} \longrightarrow T_Q\,\mathrm{Sp}(2n, \mathbb{R}^{2N}) = \{M \in \mathbb{R}^{2N \times 2n} : Q^\top M \in \mathfrak{sp}(2n)\}$$
$$V \longmapsto V + Q\left(S - \frac{1}{2}Q^\top V\right).$$

Then the map $\mathcal{R}_Q : H_Q \to \mathcal{M}$ *defined for any* $V \in H_Q$ *as*

$$\mathcal{R}_Q(V) = \mathrm{cay}(\Theta^S_Q(V)Q^\top - Q\Theta^S_Q(V)^\top)Q, \tag{5.18}$$

is a retraction.

Proof We follow [10, Proposition 2.2]. Let $V = 0 \in T_Q\mathcal{M}$, then $\Theta^S_Q(0) = QS$ and then, using the fact that $S \in \mathrm{Sym}(2n)$ and $\mathrm{cay}(0) = I_{2N}$, it holds $\mathcal{R}_Q(0) = \mathrm{cay}\left(Q(S - S^\top)Q^\top\right)Q = \mathrm{cay}(0)Q = Q$.

Let Υ_Q be defined as in (5.17). Since, by construction Υ_Q admits left inverse it is injective and then $\Upsilon_Q(V) = 0$ if and only if $V = 0 \in H_Q$. Then, $\mathcal{R}_Q(V) = Q$ if and only if $\text{cay}(\Upsilon_Q(V)) = I_{2N}$, which implies $V = 0 \in H_Q$. Moreover, since $\mathcal{R}_Q = \Lambda_Q \circ \text{cay} \circ \Upsilon_Q$, the definition of group action and the linearity of Υ result in $d\mathcal{R}_Q\big|_0 = \Psi_Q \circ \Upsilon_Q = \text{Id}_{H_Q}$. It can be easily verified that $\mathcal{R}_Q(V) \in \mathcal{M}$ for any $V \in H_Q$. □

Note that the matrix $S \in \text{Sym}(2n) \cap \mathfrak{sp}(2n)$ in the definition of the retraction (5.18) is of the form

$$S = \begin{pmatrix} A & B \\ B & -A \end{pmatrix}, \qquad \text{with} \quad A, B \in \text{Sym}(n).$$

Its choice affects the numerical performances of the algorithm for the computation of the retraction and its inverse tangent map, as pointed out in [10, Sect. 3].

In the following Subsections we propose a temporal discretization of (5.15) with an s-stage explicit Runge–Kutta method and show that the resulting algorithm has arithmetic complexity of order $C_{\mathcal{F}} + O(Nn^2)$ at every stage of the temporal solver.

5.3.1 Efficient computation of retraction and inverse tangent map

In the interval $(t^m, t^{m+1}]$ the local evolution on the tangent space, corresponding to (5.15), reads

$$\dot{V}(t) = \left(d\mathcal{R}_{U_m}\big|_{V(t)}\right)^{-1} \mathcal{F}\left(\mathcal{R}_{U_m}(V(t))\right) =: f_m(V(t)).$$

Let $(b_i, a_{i,j})$ for $i = 1, \ldots, s$ and $j = 1, \ldots, i-1$ be the coefficients of the Butcher tableau describing the s-stage explicit Runge–Kutta method. Then the numerical approximation of (5.15)–(5.14) with $U_0 := Q \in \mathcal{M}$ and $V_0 = 0 \in T_Q\mathcal{M}$ is given in Algorithm 2.

Algorithm 2 Explicit RK-tangent scheme in $(t^m, t^{m+1}]$

Input: $U_m \in \mathcal{M}, \{b_i\}_{i=1}^s, \{a_{i,j}\}_{i,j=1}^s$

1: $A_m^1 = \mathcal{F}(U_m)$
2: **for** $i = 2, \ldots, s$ **do**
3: $\quad A_m^i = f_m\left(\Delta t \sum_{j=1}^{i-1} a_{i,j} A_m^j\right),$
4: **end for**
5: $V_{m+1} = \Delta t \sum_{i=1}^s b_i A_m^i,$
6: **return** $U_{m+1} = \mathcal{R}_{U_m}(V_{m+1}) \in \mathcal{M}$

Other than the evaluation of the velocity field $\mathcal{F}$ at $\mathcal{R}_{U_m}(V)$, the crucial points of Algorithm 2 in terms of computational cost, are the evaluation of the retraction and

the computation of its inverse tangent map. If we assume that both operations can be performed with a computational cost of order $O(Nn^2)$, then Algorithm 2 has an overall arithmetic complexity of order $O(Nn^2 s) + C_{\mathcal{F}} s$, where $C_{\mathcal{F}}$ is the cost to compute $\mathcal{F}(U)$ in (5.2) at any given $U \in \mathcal{M}$.

Computation of the retraction. A standard algorithm to compute the retraction $\mathcal{R}_Q$ (5.18) at the matrix $V \in \mathbb{R}^{2N \times 2n}$ requires $O(N^2 n)$ for the multiplication between $\text{cay}(\Upsilon_Q(V))$ and Q, plus the computational cost to evaluate the Cayley transform at $\Upsilon_Q(V) \in \mathbb{R}^{2N \times 2N}$. However, for any $V \in H_Q$, the matrix $\Upsilon_Q(V) \in \mathfrak{g}_{2N}$ admits the low-rank splitting

$$\Upsilon_Q(V) = \Theta_Q^S(V) Q^\top - Q \Theta_Q^S(V)^\top = \alpha \beta^\top,$$

where

$$\alpha := \left[\Theta_Q^S(V) \,|\, -Q \right] \in \mathbb{R}^{2N \times 4n}, \qquad \beta := \left[Q \,|\, \Theta_Q^S(V) \right] \in \mathbb{R}^{2N \times 4n}. \quad (5.19)$$

We can revert to the results of Proposition 5.2 (with $k = 4n$) so that the retraction (5.18) can be computed as

$$\mathcal{R}_Q(V) = \text{cay}(\Upsilon_Q(V)) Q = Q + \alpha (\beta^\top \alpha)^{-1} \big(\text{cay}(\beta^\top \alpha) - I_{4n} \big) \beta^\top Q,$$

with computational cost of order $O(Nn^2)$.

Computation of the inverse tangent map of the retraction. Let $Q \in \mathcal{M}$ and $V \in H_Q$. Using the definition of retraction (5.18) we have

$$\mathcal{R}_Q(V) = \text{cay}(\Upsilon_Q(V)) Q = (\Lambda_Q \circ \text{cay} \circ \Upsilon_Q)(V).$$

Then, the tangent map $\mathrm{d}\mathcal{R}_Q$ reads

$$\mathrm{d}\mathcal{R}_Q = \mathrm{d}\Lambda_Q \circ \mathrm{dcay} \circ \mathrm{d}\Upsilon_Q : T H_Q \longrightarrow T\mathfrak{g}_{2N} \cong \mathfrak{g}_{2N} \longrightarrow T\mathcal{G}_{2N} \longrightarrow T_Q \mathcal{M}.$$

Fixing the fiber on $T H_Q$ corresponding to $V \in H_Q$ results in

$$\begin{aligned}
\mathrm{d}\mathcal{R}_Q\big|_V (\widetilde{V}) = \mathrm{d}\mathcal{R}_Q\big(V, \widetilde{V}\big) &= \mathrm{d}\Lambda_Q\big|_{\text{cay}\big(\Upsilon_Q(V)\big)} \circ \mathrm{dcay}_{\Upsilon_Q(V)}\big(\Upsilon_Q(\widetilde{V})\big) \\
&= \mathrm{dcay}_{\Upsilon_Q(V)}\big(\Upsilon_Q(\widetilde{V})\big) \text{cay}\big(\Upsilon_Q(V)\big) Q = \mathrm{dcay}_{\Upsilon_Q(V)}\big(\Upsilon_Q(\widetilde{V})\big) \mathcal{R}_Q(V),
\end{aligned}$$

where we have used the linearity of the map Υ_Q.

Assume we know $W \in H_{\mathcal{R}_Q(V)}$. We want to compute $\widetilde{V} \in H_Q$ such that

$$\mathrm{d}\mathcal{R}_Q\big|_V (\widetilde{V}) = \mathrm{dcay}_{\Upsilon_Q(V)}\big(\Upsilon_Q(\widetilde{V})\big) \mathcal{R}_Q(V) = W. \quad (5.20)$$

It is possible to solve problem (5.20) with arithmetic complexity $O(Nn^2)$ by proceeding as in [10, Sect. 3.2.1]. Since, for our algorithm, the result of [10] can be

extended to the case of arbitrary matrix $S \in \mathrm{Sym}(2n) \cap \mathfrak{sp}(2n)$ in (5.18), we report the more general derivation in Appendix A. Note that, for $S = 0$ and explicit Euler scheme, the two Algorithms 1 and 2 are equivalent.

5.3.2 Convergence estimates for the tangent method

Since the retraction and its inverse tangent map in Algorithm 2 can be computed exactly, the smoothness properties of $\mathcal{R}$ allow to derive error estimates for the approximate reduced basis in terms of the numerical solution of the evolution problem (5.15) in the tangent space.

Proposition 5.7 *The retraction map $\mathcal{R} : T\mathcal{M} \to \mathcal{M}$ defined in (5.18) is locally Lipschitz continuous in the Frobenius $\|\cdot\|$-norm, namely for any $Q \in \mathcal{M}$, $\mathcal{R}_Q : H_Q \to \mathcal{M}$ satisfies*

$$\|\mathcal{R}_Q(V) - \mathcal{R}_Q(W)\| \le 3\|V - W\|, \quad \forall\, V,\, W \in H_Q.$$

Proof Let $U := \mathcal{R}_Q(V) = \mathrm{cay}(\Upsilon_Q(V))Q$ and $Y := \mathcal{R}_Q(W) = \mathrm{cay}(\Upsilon_Q(W))Q$. Using the definition of Cayley transform (5.3) we have, for $\overline{\Upsilon}_Q(\cdot) := \Upsilon_Q(\cdot)/2$,

$$\begin{aligned}
0 &= \big(I_{2N} - \overline{\Upsilon}_Q(V)\big)U - \big(I_{2N} - \overline{\Upsilon}_Q(W)\big)Y \\
&\quad - \big(I_{2N} + \overline{\Upsilon}_Q(V)\big)Q - \big(I_{2N} + \overline{\Upsilon}_Q(W)\big)Q \\
&= \big(I_{2N} - \overline{\Upsilon}_Q(V)\big)(U - Y) - \big(\overline{\Upsilon}_Q(V) - \overline{\Upsilon}_Q(W)\big)(Q + Y).
\end{aligned}$$

Since Υ_Q is skew-symmetric $\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}$ is normal. Then,

$$\|\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}\|_2 = \rho\big[\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}\big] \le 1.$$

Hence, since Q and Y are (semi-)orthogonal matrices, it holds

$$\|U - Y\| \le \|\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}\|_2 \|\Upsilon_Q(V) - \Upsilon_Q(W)\| \le \|\Upsilon_Q(V) - \Upsilon_Q(W)\|.$$

Using the definition of Υ_Q from (5.17) results in

$$\begin{aligned}
\|\Upsilon_Q(V) - \Upsilon_Q(W)\| &= \|(V - W)Q^\top - Q(V - W) + Q(V^\top - W^\top)QQ^\top\| \\
&\le 3\|V - W\|.
\end{aligned}$$

$\square$

It follows that the solution of Algorithm 2 can be computed with the same order of accuracy of the RK temporal scheme.

Corollary 5.8 *For $Q \in \mathcal{M}$ given, let $\mathcal{R}_Q$ be the retraction map defined in (5.18). Let $U(t^m) = \mathcal{R}_Q(V(t^m))$, where $V(t^m)$ is the exact solution of (5.15) at a given time t^m and let $U_m = \mathcal{R}_Q(V_m)$, where V_m is the numerical solution of (5.15) at time t^m*

obtained with Algorithm 2. Assume that the numerical approximation of the evolution equation for the unknown V on the tangent space of $\mathcal{M}$ is of order $O(\Delta t^k)$. Then, it holds

$$\|U(t^m) - U_m\| = O(\Delta t^k).$$

6 Numerical experiment

To gauge the performances of the proposed method, we consider the numerical simulation of the finite-dimensional parametrized Hamiltonian system arising from the spatial approximation of the one-dimensional shallow water equations (SWE). The shallow water equations are used in oceanography to describe the kinematic behavior of thin inviscid single fluid layers flowing over a changing topography. Under the assumptions of irrotational flow and flat bottom topography, the fluid is described by the scalar potential ϕ and the height h of the free-surface, normalized by its mean value, via the nonlinear system of PDEs

$$\begin{cases} \partial_t h + \partial_x (h\,\partial_x \phi) = 0, & \text{in } (-L, L) \times (0, T], \\ \partial_t \phi + \dfrac{1}{2}|\partial_x \phi|^2 + h = 0, & \text{in } (-L, L) \times (0, T], \end{cases} \qquad (6.1)$$

where $L = 10$, $T = 7$, $h, \phi : [-L, L] \times (0, T] \times \Gamma \to \mathbb{R}$ are the state variables, and $\Gamma \subset \mathbb{R}^2$ is a compact set of parameters. Here we consider $\Gamma := [0.1, 0.15] \times [0.2, 1.5]$. The system is provided with periodic boundary conditions for both state variables, and with parametric initial conditions $(h^0(x; \eta), \phi^0(x; \eta)) = (1 + \alpha e^{-\beta x^2}, 0)$, where α controls the amplitude of the initial hump in the depth, β describes its width, and $\eta = (\alpha, \beta)$.

For the numerical discretization in space, we consider a Cartesian mesh on $[-L, L]$ with $N-1$ equispaced intervals and we denote with Δx the mesh width. The degrees of freedom of the problem are the nodal values of the height and potential, i.e. $(h_h(t; \eta), \phi_h(t; \eta)) = (h_1, \ldots, h_N, \phi_1, \ldots, \phi_N)$. The discrete set of parameters Γ_h is obtained by uniformly sampling Γ with 10 samples per dimension, for a total of $p = 100$ different configurations. This implies that the full model variable $\mathcal{R}$ in (4.1) is the $N \times p$ matrix given by $\mathcal{R}_{i,k}(t) = h_i(t; \eta_h^k)$ if $1 \leq i \leq N$ and $\mathcal{R}_{i,k}(t) = \phi_{i-N}(t; \eta_h^k)$ if $N+1 \leq i \leq 2N$, for any $k \in \{1, \ldots, p\}$, where η_h^k denotes the k-th entry of the vector $\eta_h \in \mathbb{R}^p$ containing the samples of the parameters. We consider second order accurate centered finite difference schemes to discretize the first order spatial derivative in (6.1). The evolution problem (6.1) admits a canonical symplectic Hamiltonian. Spatial discretization with centered finite differences yields a Hamiltonian dynamical system where the Hamiltonian associated with the k-th parameter is given by

Fig. 1 SWE-1D: Error between the full model solution and the reduced solution at final time vs. the algorithm runtime
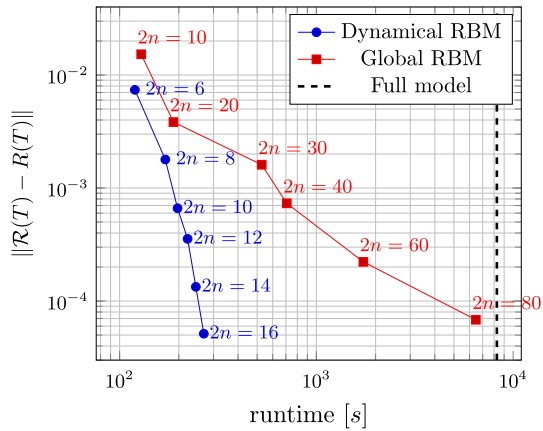
$$\mathcal{H}_k(\mathcal{R}(t)) = \frac{1}{2} \sum_{i=1}^{N} \left(h_i(t; \eta_h^k) \left(\frac{\phi_{i+1}(t; \eta_h^k) - \phi_{i-1}(t; \eta_h^k)}{2\Delta x} \right)^2 + h_i^2(t; \eta_h^k) \right).$$

Wave-type phenomena often exhibit a low-rank behavior only locally in time, and, hence, global (in time) model order reduction proves ineffective in these situations. We show this behavior by comparing the performances of our dynamical reduced basis method with the global symplectic reduced basis approach of [30] based on complex SVD. For the latter, a symplectic reduced space is obtained from the full model obtained by discretizing (6.1) with centred finite differences in space, with $N = 1000$, and the implicit midpoint rule in time, with $\Delta t = 10^{-3}$. We consider snapshots every 10 time steps and 4 uniformly distributed samples of Γ per dimension. Concerning the dynamical reduced model, we evaluate the initial condition $(h_h(0; \eta_h), \phi_h(0; \eta_h))$ at all values η_h and compute the matrix $\mathcal{R}_0 \in \mathbb{R}^{2N \times p}$ having as columns each of the evaluations. As initial condition for the reduced system (4.10), we use $U(0) = U_0 \in \mathbb{R}^{2N \times 2n}$ obtained via complex SVD of the matrix $\mathcal{R}_0$ truncated at n, while $Z(0) = U_0^T \mathcal{R}_0$. Then, we solve system (4.10) with a 2-stage partitioned Runge-Kutta method obtained as follows: the evolution equation for the coefficients Z is discretized with the implicit midpoint rule; while the evolution equation (5.1) for the reduced basis is solved using the tangent method described in Algorithm 2 with the explicit midpoint scheme, i.e. $s = 2$, $b_1 = 0$, $b_2 = 1$, and $a_{1,1} = a_{1,2} = a_{2,2} = 0$, $a_{2,1} = 1/2$. Note that the resulting partitioned RK method has order of accuracy 2 and the numerical integrator for Z is symplectic [15, Sect. III.2]. Finally, the nonlinear quadratic operator in (6.1), is reduced by using tensorial techniques [32].

In Fig. 1 we report the error in the Frobenius norm, at final time, between the full model solution and the reduced solution obtained with the two different approaches and various dimensions of the reduced space. Note that the runtime includes also the offline phase for the global approach.

The results of Fig. 1 show that the dynamical reduced basis method outperforms the global approach by reaching comparable accuracy at a reduced computational cost. Moreover, as the dimension of the reduced space increases, the runtime of the global

Fig. 2 Evolution of the error in the conservation of the Hamiltonian
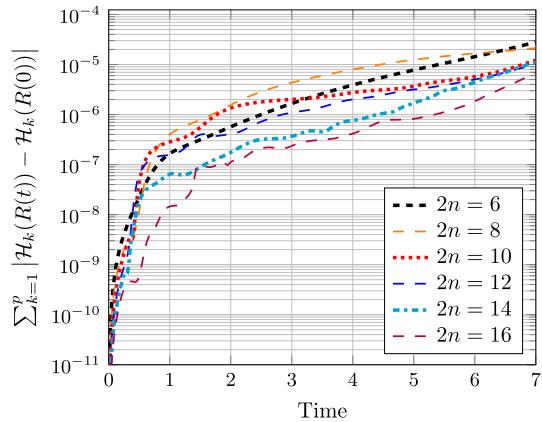
method becomes comparable to the one required to solve the high-fidelity problem, meaning that there is no gain in performing global model order reduction.

Figure 2 shows the evolution of the error in the conservation of the discrete Hamiltonian, averaged over all p values of the parameter. Since the Hamiltonian is a cubic quantity, we do not expect exact conservation associated with the proposed partitioned RK scheme. In addition, as pointed out at the end of Sect. 4, we cannot guarantee exact preservation of the invariants at the interface between temporal intervals, since the reduced solution is projected into the space spanned by the updated basis. However, the preservation of the symplectic structure both in the reduction and in the discretization yields a good control on the Hamiltonian error, as it can be observed in Fig. 2.

7 Concluding remarks and future work

Nonlinear dynamical reduced basis methods for parameterized finite-dimensional Hamiltonian systems have been developed to mitigate the computational burden of large-scale, multi-query and long-time simulations. The proposed techniques provide an attractive computational approach to deal with the local low-rank nature of Hamiltonian dynamics while preserving the geometric structure of the phase space even at the discrete level.

Possible extensions of this work involve the numerical study of the proposed algorithm including high order splitting temporal integrators, numerical approximations ensuring the exact conservation of Hamiltonian, and restarting procedures of the Cayley RK algorithm. Moreover, the extension of dynamical reduced basis methods to Hamiltonian systems with a nonlinear Poisson structure would allow nonlinear structure-preserving model order reduction of a large class of problems, including Euler and Vlasov–Maxwell equations. Some of these topics will be investigated in forthcoming works.

A Exponential map

Let us consider Algorithm 1 with the exponential as coordinate map, namely $\psi = \exp : \mathfrak{g}_{2N} \to \mathfrak{g}_{2N}$. For any $\Omega_m^i \in \mathfrak{g}_{2N}$ and $U_m^i \in \mathcal{M}$, the matrix $\mathrm{dexp}_{\Omega_m^i}^{-1}(\mathcal{L}(U_m^i))$ in Line 3 can be approximated by truncating the Baker–Campbell–Hausdorff (BCH) formula as

$$\mathrm{dexp}_{\Omega_m^i}^{-1}(\mathcal{L}(U_m^i)) \approx \widehat{\Lambda}_i := \sum_{k=0}^{q} \frac{\mathcal{B}_k}{k!} \, \mathrm{ad}_{\Omega_m^i}^{k}(\mathcal{L}(U_m^i)), \quad \text{for some } q \in \mathbb{N}, \quad (A.1)$$

where $\mathcal{B}_k$ denotes the k-th Bernoulli number and $\mathrm{ad}_{\Omega_m^i}^{0}(\mathcal{L}(U_m^i)) = \mathcal{L}(U_m^i)$. Observe that $\widehat{\Lambda}_i \in \mathbb{R}^{2N \times 2N}$ belongs to the Lie algebra $\mathfrak{g}_{2N}$, and, hence, each Ω_m^i is in $\mathfrak{g}_{2N}$ and the solution of the RK-MK method remains on $\mathcal{M}$, see e.g. [15, Theorem 8.4].

To assess the computational complexity of Algorithm 1, we need to consider two operations: the evaluation of $\widehat{\Lambda}_i$ and the computation of $U_m^i = \exp(\Omega_m^i)U_m$, for any $i = 2, \ldots, s$. To this aim, we first rewrite each commutator in (A.1) as a matrix polynomial.

Lemma A.1 *Let* $A, B \in \mathbb{R}^{2N \times 2N}$. *For any fixed* $k \geq 0$*, there exist coefficients* $\{c_h^{(k)}\}_{h=0}^{k} \subset \mathbb{R}$ *such that*

$$\mathrm{ad}_A^k(B) = \sum_{h=0}^{k} c_h^{(k)} A^h B A^{k-h}. \quad (A.2)$$

Proof We proceed by induction on k. For $k = 0$, $\mathrm{ad}_A^0(B) = B$ and $c_0^{(0)} = 1$. For $k = 1$, $\mathrm{ad}_A(B) = [A, B] = AB - BA = c_0^{(1)} BA + c_1^{(1)} AB$, with $c_0^{(1)} = -1$ and $c_0^{(1)} = 1$. Assume that $\mathrm{ad}_A^{k-1}(B)$, with $k \geq 2$, can be expressed in polynomial form. Then,

$$\mathrm{ad}_A^k(B) = [A, \mathrm{ad}_A^{k-1}(B)] = \sum_{h=0}^{k-1} c_h^{(k-1)} A^{h+1} B A^{k-1-h} - \sum_{h=0}^{k-1} c_h^{(k-1)} A^h B A^{k-h},$$

which is of the form (A.2) with $c_0^{(k)} = -c_0^{(k-1)}, c_k^{(k)} = c_{k-1}^{(k-1)}$ and $c_h^{(k)} = c_{h-1}^{(k-1)} - c_h^{(k-1)}$ for any $1 \leq h \leq k - 1$. □

The matrix polynomial form (A.2) allows to estimate the rank of the $\{\widehat{\Lambda}_i\}_{i=1}^s$.

Lemma A.2 *Let $\widehat{\Lambda}_i \in \mathfrak{g}_{2N}$ be defined as in (A.1). Then,*

$$\text{rank}(\widehat{\Lambda}_i) \leq \min\{2^{i-1}, q+1\} \, \text{rank}(\mathcal{L}(U_m^i)). \tag{A.3}$$

Proof Using Lemma A.1, we have that, for any $A, B \in \mathfrak{g}_{2N}$,

$$C := \sum_{k=0}^{q} \text{ad}_A^k(B) = \sum_{k=0}^{q} \sum_{h=0}^{k} c_h^{(k)} A^h B A^{k-h} = \sum_{h=0}^{q} A^h B \sum_{k=h}^{q} c_h^{(k)} A^{k-h}.$$

Since the rank of a matrix product is bounded by the minimum among the ranks of the factors, this implies that $\text{rank}(C) \leq (q+1)\text{rank}(B)$ and, hence, $\text{rank}(\widehat{\Lambda}_i) \leq (q+1)\,\text{rank}(\mathcal{L}(U_m^i))$.

We now prove that, for any $0 \leq k \leq q$, there exists matrices $E_k, D_k \in \mathbb{R}^{2N \times 2N}$ such that $\text{ad}_A^k(B) = AE_k + BD_k$. We proceed by induction on k. For $k = 0$, $\text{ad}_A^0(B) = B$ so that $E_0 = 0$ and $D_0 = I_{2N}$. For $k = 1$, $\text{ad}_A(B) = [A, B] = AB - BA$ so that $E_1 = B$ and $D_1 = -A$. Assume that the statement holds for $k - 1$, with $k \geq 3$, then $\text{ad}_A^k(B) = A \, \text{ad}_A^{k-1}(B) - \text{ad}_A^{k-1}(B)A = A(AE_{k-1} + BD_{k-1}) - (AE_{k-1} + BD_{k-1})A = AE_k + BD_k$ with $E_k = AE_{k-1} + BD_{k-1} - E_{k-1}A$ and $D_k = -D_{k-1}A$. Therefore,

$$C = \sum_{k=0}^{q}(AE_k + BD_k) = A\left(\sum_{k=1}^{q} E_k\right) + B\left(\sum_{k=1}^{q} D_k\right),$$

and, hence, $\text{rank}(C) \leq \text{rank}(A) + \text{rank}(B)$. This is equivalent to $\text{rank}(\widehat{\Lambda}_i) \leq \text{rank}(\Omega_m^i) + \text{rank}(\mathcal{L}(U_m^i))$. Using the definition of Ω_m^i from Line 3 of Algorithm 1, the rank of $\widehat{\Lambda}_i$, for any $i \geq 2$, can be bounded as

$$\text{rank}(\widehat{\Lambda}_i) \leq \text{rank}(\mathcal{L}(U_m^i)) + \sum_{j=1}^{i-1} \text{rank}(\widehat{\Lambda}_j).$$

Since $\text{rank}(\widehat{\Lambda}_1) = \text{rank}(\mathcal{L}(U_m))$, it easily follows by induction that $\text{rank}(\widehat{\Lambda}_i) \leq 2^{i-1}\,\text{rank}(\mathcal{L}(U_m^i))$. □

Observe that the factorization (5.13) implies that each term $\{\mathcal{L}(U_m^i)\}_{i=1}^s$, with $\mathcal{L}$ defined in (5.8), has rank at most $4n$. Therefore, from Lemma A.2, it follows that

$$r_i := \text{rank}(\Omega_m^i) \leq \sum_{j=1}^{i-1} \text{rank}(\widehat{\Lambda}_j) \leq 4n \sum_{j=1}^{i-1} \min\{2^{i-1}, q+1\}. \tag{A.4}$$

It can be inferred from (A.4) that the bound $q + 1$ is the one dominating in the computation of Ω_m^i whenever the number s of RK stages is sufficiently large. An

optimal number q_{opt} of commutators to achieve the accuracy of the corresponding RK method can be derived as in [3,7]. We consider a few examples from [7, Table 3.1]: RKF45 has $s = 6$, $q_{opt} = 5$, and hence $q_{opt} + 1 \leq 2^{i-1}$ for $i \geq 4$; DVERK has $s = 8$, $q_{opt} = 10$, and hence $q_{opt} + 1 \leq 2^{i-1}$ for $i \geq 5$; Butcher7 has $s = 9$, $q_{opt} = 21$, and hence $q_{opt} + 1 \leq 2^{i-1}$ for $i \geq 6$. In light of these considerations, we consider the bound $r_i \leq 4n(i-1)(q+1)$ although for small i this might not be sharp. With the estimate (A.4) on the rank of Ω^i_m, we can assess the cost of computing U^i_m at each stage of the RK-MK Algorithm 1. The computation of the exponential of a matrix in $\mathfrak{g}_{2N}$ requires $O(N^3)$ operations, but this cost can be mitigated whenever the argument of the exponential is of low-rank. Similarly to Proposition 5.2, it can be shown that the cost to compute $\exp(uv^\top)Y$ with $u, v \in \mathbb{R}^{2N \times k}$ and $Y \in \mathbb{R}^{2N \times 2n}$ is $O(k^3 + k^2n + Nnk)$ [8, Proposition 3]. In Algorithm 1 we need to evaluate the exponential of $\{\Omega^i_m\}^s_{i=1}$, with $\text{rank}(\Omega^i_m) \leq 4n(i-1)(q+1)$. Therefore, the computation of all U^i_m in Line 4, for $2 \leq i \leq s$, requires $O(Nn^2s^2q + n^3s^4q^3)$.

The other contribution to the computational cost of Algorithm 1 comes from the evaluation of each $\widehat{\Lambda}_i$ in (A.1). To estimate this cost, we resort to the polynomial expression (A.2) and the low-rank splitting of $\Omega^i_m = \alpha_i \beta_i^\top$, with $\alpha_i, \beta_i \in \mathbb{R}^{2N \times r_i}$, and of $\mathcal{L}(U^i_m) = \gamma_i \delta_i^\top$ with $\gamma_i, \delta_i \in \mathbb{R}^{2N \times 4n}$. Let $\hat{c}^{(k)}_j := c^{(k)}_j \mathcal{B}_k/k!$, for any $0 \leq k, j \leq q$, then

$$\widehat{\Lambda}_i = \hat{c}^{(0)}_0 \gamma_i \delta_i^\top + \sum_{j=1}^q \left(\hat{c}^{(j)}_0 \gamma_i \delta_i^\top (\alpha_i \beta_i^\top)^j + \hat{c}^{(j)}_j (\alpha_i \beta_i^\top)^j \gamma_i \delta_i^\top \right)$$

$$+ \sum_{j=1}^q \alpha_i (\beta_i^\top \alpha_i)^{j-1} \beta_i^\top \gamma_i \delta_i^\top \alpha_i \left(\sum_{k=j+1}^q \hat{c}^{(k)}_j (\beta_i^\top \alpha_i)^{k-j-1} \right) \beta_i^\top$$

$$= \hat{c}^{(0)}_0 \gamma_i \delta_i^\top + \sum_{j=1}^q \left(\hat{c}^{(j)}_0 \gamma_i \delta_i^\top \alpha_i (\beta_i^\top \alpha_i)^{j-1} \beta_i^\top + \hat{c}^{(j)}_j \alpha_i (\beta_i^\top \alpha_i)^{j-1} \beta_i^\top \gamma_i \delta_i^\top \right)$$

$$+ \sum_{j=1}^q \alpha_i (\beta_i^\top \alpha_i)^{j-1} \beta_i^\top \gamma_i \delta_i^\top \alpha_i P_j \beta_i^\top, \quad \text{with } P_j := \sum_{k=0}^{q-j-1} \hat{c}^{(k+j+1)}_j (\beta_i^\top \alpha_i)^k.$$

The terms:

- $\{P_j\}^q_{j=1}$ can be computed in $O(Nr_i^2 + q^2r_i^3)$ operations;
- $\{(\beta_i^\top \alpha_i)^j\}^q_{j=1}$ can be computed in $O(qr_i^3)$ operations;
- $\{\beta_i^\top \gamma_i \delta_i^\top \alpha_i P_j\}^q_{j=1}$ can be computed in $O(nqr_i^2 + Nnr_i)$ operations.

Therefore, the overall computational cost to evaluate each $\widehat{\Lambda}_i$ is $O(Nr_i^2 + q^2r_i^3 + nqr_i^2 + Nnr_i)$. Using $r_i = \text{rank}(\Omega^i_m) \leq 4n(i-1)(q+1)$ and summing over the stages of the RK scheme, all terms involved in Algorithm 1 can be evaluated with arithmetic complexity $O(Nn^2q^2s^3 + n^3q^5s^4) + C_{\mathcal{F}}$, where $C_{\mathcal{F}}$ is the complexity of the algorithm to compute $\mathcal{F}(U)$ in (5.2) at any given $U \in \mathcal{M}$. The latter is, thus, the computational complexity of Algorithm 1 with $\psi = \exp$. Since each Ω^i_m can be written as the sum of elements in the Lie algebra $\mathfrak{g}_{2N}$, namely $\Omega^i_m = \sum_{j=1}^{i-1} \alpha_j \beta_j^\top$ with $\alpha_j, \beta_j \in \mathbb{R}^{2N \times r_i}$,

one might suggest to approximate $\exp(\Omega_m^i)$ with $E(\Omega_m^i) := \Pi_{j=1}^{i-1}\exp(\alpha_j\beta_j^\top)$ in the spirit of [8]. However, such an approximation does not bring significant computational savings nor it is guaranteed to provide a good approximation of the exponential map.

A Efficient computation of the inverse tangent map

We propose an algorithm to solve (5.20) with a computational cost of order $O(Nn^2)$. We proceed exactly as in [10, Sect. 3.2.1] with the only difference that we consider any arbitrary $S \in \text{Sym}(2n) \cap \mathfrak{sp}(2n)$.

Using the definition of the derivative of the Cayley transform (5.4) we can recast (5.20) as

$$
\begin{aligned}
&\Upsilon_Q(\widetilde{V})(I_{2N} + \overline{\Upsilon}_Q(V))^{-1}\mathcal{R}_Q(V) - (I_{2N} - \overline{\Upsilon}_Q(V))W = 0, \\
&\overline{\Upsilon}_Q(V) := \frac{\Upsilon_Q(V)}{2}.
\end{aligned} \tag{A.1}
$$

Moreover, using the definition of $\mathcal{R}_Q(V)$ in (5.18) results in

$$
\begin{aligned}
2\mathcal{R}_Q(V) &= (I_{2N} + \overline{\Upsilon}_Q(V))\mathcal{R}_Q(V) + (I_{2N} - \overline{\Upsilon}_Q(V))\mathcal{R}_Q(V) \\
&= (I_{2N} + \overline{\Upsilon}_Q(V))\mathcal{R}_Q(V) + (I_{2N} - \overline{\Upsilon}_Q(V))(I_{2N} - \overline{\Upsilon}_Q(V))^{-1}(I_{2N} + \overline{\Upsilon}_Q(V))Q \\
&= (I_{2N} + \overline{\Upsilon}_Q(V))(\mathcal{R}_Q(V) + Q).
\end{aligned}
$$

Therefore, substituting in (A.1) and using the definition of Υ_Q from (5.17) gives

$$
\Theta_Q^S(\widetilde{V})Q^\top(\mathcal{R}_Q(V)+Q) - Q\Theta_Q^S(\widetilde{V})^\top(\mathcal{R}_Q(V)+Q) - (2I_{2N} - \Upsilon_Q(V))W = 0. \tag{A.2}
$$

We proceed by solving problem (A.2) for $\widetilde{\Theta} := \Theta_Q^S(\widetilde{V}) \in T_Q \text{Sp}(2n, \mathbb{R}^{2N})$ and then, in view of (5.16), we recover $\widetilde{V} \in H_Q$ as $\widetilde{V} = \widetilde{\Theta} - Q\widetilde{\Theta}^\top Q$, at a computational cost of order $O(Nn^2)$.

It is possible to recast problem (A.2) as $\Theta_Q^S(\widetilde{V}) = QT_1(\widetilde{V}) + T_2$, where

$$
\begin{aligned}
T_1(\widetilde{V}) &:= \Theta_Q^S(\widetilde{V})^\top(\mathcal{R}_Q(V) + Q)(Q^\top\mathcal{R}_Q(V) + I_{2n})^{-1}, \\
T_2 &:= (2I_{2N} - \Upsilon_Q(V))W(Q^\top\mathcal{R}_Q(V) + I_{2n})^{-1}.
\end{aligned}
$$

The term T_2, independent of $\widetilde{V}$, can be computed in $O(Nn^2 + n^3)$ operations. Indeed, since $\Upsilon_Q(V) = \alpha\beta^\top$ as defined in (5.19), the term $\Upsilon_Q(V)W$ can be computed as $\alpha(\beta^\top W)$ in $O(Nn^2)$ flops. The term $T_1(\widetilde{V})$ can be expressed as $T_1(\widetilde{V}) = Q^\top\Theta_Q^S(\widetilde{V}) + QT_2$. Using the fact that $Q^\top\Theta_Q^S(\widetilde{V}) + \Theta_Q^S(\widetilde{V})^\top Q = 2S$, the symmetric part of T_1

reads $T_1 + T_1^\top = 2S - Q^\top T_2 + T_2^\top Q$. Moreover,

$$
\begin{aligned}
(\mathcal{R}_Q(V) + Q)^\top \Theta_Q^S(\widetilde{V}) &= (\mathcal{R}_Q(V) + Q)^\top T_2 + (\mathcal{R}_Q(V)^\top Q + I_{2n}) T_1(\widetilde{V}), \\
(\mathcal{R}_Q(V)^\top Q + I_{2n}) T_1(\widetilde{V})^\top & \\
&= (Q^\top \mathcal{R}_Q(V) + I_{2n})^\top (Q^\top \mathcal{R}_Q(V) + I_{2n})^{-\top} (\mathcal{R}_Q(V) + Q)^\top \Theta_Q^S(\widetilde{V}).
\end{aligned}
$$

The skew-symmetric part of T_1 is then $T_1 - T_1^\top = -(\mathcal{R}_Q(V)^\top Q + I_{2n})^{-1}(\mathcal{R}_Q(V) + Q)^\top T_2$. Therefore,

$$
\begin{aligned}
2T_1(\widetilde{V}) &= \left((T_1(\widetilde{V}) + T_1(\widetilde{V})^\top) + (T_1(\widetilde{V}) - T_1(\widetilde{V})^\top) \right) \\
&= 2S - (Q^\top T_2 - T_2^\top Q) - (\mathcal{R}_Q(V)^\top Q + I_{2n})^{-1}(\mathcal{R}_Q(V) + Q)^\top T_2.
\end{aligned}
$$

It is straightforward to show that all operations involved in the computation of T_1 can be done with complexity of order $O(Nn^2)$.

References

1. Abraham, R., Marsden, J.E.: Foundations of Mechanics, 2nd edn. Addison-Wesley Publishing Company Inc, Redwood City (1987)
2. Afkham, B.M., Hesthaven, J.S.: Structure preserving model reduction of parametric Hamiltonian systems. SIAM J. Sci. Comput. **39**(6), A2616–A2644 (2017). https://doi.org/10.1137/17M1111991
3. Blanes, S., Casas, F., Ros, J.: High order optimized geometric integrators for linear differential equations. BIT **42**(2), 262–284 (2002). https://doi.org/10.1023/A:1021942823832
4. Buchfink, P., Bhatt, A., Haasdonk, B.: Symplectic model order reduction with non-orthonormal bases. Math. Comput. Appl. (2019). https://doi.org/10.3390/mca24020043
5. Cannas da Silva, A.: Lectures on Symplectic Geometry. Lecture Notes in Mathematics, vol. 1764. Springer, Berlin (2001). https://doi.org/10.1007/978-3-540-45330-7
6. Carlberg, K., Tuminaro, R., Boggs, P.: Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics. SIAM J. Sci. Comput. **37**(2), B153–B184 (2015). https://doi.org/10.1137/140959602
7. Casas, F., Owren, B.: Cost efficient Lie group integrators in the RKMK class. BIT **43**(4), 723–742 (2003). https://doi.org/10.1023/B:BITN.0000009959.29287.d4
8. Celledoni, E., Iserles, A.: Approximating the exponential from a Lie algebra to a Lie group. Math. Comput. **69**(232), 1457–1480 (2000). https://doi.org/10.1090/S0025-5718-00-01223-0
9. Celledoni, E., Owren, B.: A class of intrinsic schemes for orthogonal integration. SIAM J. Numer. Anal. **40**(6), 2069–2084 (2002). https://doi.org/10.1137/S0036142901385143
10. Celledoni, E., Owren, B.: On the implementation of Lie group methods on the Stiefel manifold. Numer. Algorithms **32**(2–4), 163–183 (2003). https://doi.org/10.1023/A:1024079724094
11. DeVore, R.A.: The theoretical foundation of reduced basis methods. In: Model Reduction and Approximation: Theory and Algorithms, Chap. 3, pp. 137–168 . Society for Industrial and Applied Mathematics (2017)
12. Diele, F., Lopez, L., Peluso, R.: The Cayley transform in the numerical solution of unitary differential systems. Adv. Comput. Math. **8**(4), 317–334 (1998). https://doi.org/10.1023/A:1018908700358
13. Edelman, A., Arias, T.A., Smith, S.T.: The geometry of algorithms with orthogonality constraints. SIAM J. Matrix Anal. Appl. **20**(2), 303–353 (1999). https://doi.org/10.1137/S0895479895290954
14. Feppon, F., Lermusiaux, P.F.J.: A geometric approach to dynamical model order reduction. SIAM J. Matrix Anal. Appl. **39**(1), 510–538 (2018). https://doi.org/10.1137/16M1095202
15. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration. Second. Springer Series in Computational Mathematics, vol. 31. Springer, Berlin (2006). https://doi.org/10.1007/3-540-30666-8
16. Hesthaven, J.S., Pagliantini, C.: Structure-preserving reduced basis methods for Poisson systems. Math. Comput. (2020). https://doi.org/10.1090/mcom/3618

17. Iserles, A.: On Cayley-transform methods for the discretization of Lie-group equations. Found. Comput. Math. **1**(2), 129–160 (2001). https://doi.org/10.1007/s102080010003
18. Iserles, A., Munthe-Kaas, H.Z., Nørsett, S.P., Zanna, A.: Lie-group methods. Acta Numer. **9**, 215–365 (2000). https://doi.org/10.1017/S0962492900002154
19. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry. Wiley Classics Library, vol. 1. Wiley (1963)
20. Koch, O., Lubich, C.: Dynamical low-rank approximation. SIAM J. Matrix Anal. Appl. **29**(2), 434–454 (2007). https://doi.org/10.1137/050639703
21. Lall, S., Krysl, P., Marsden, J.E.: Structure-preserving model reduction for mechanical systems. Phys. D **184**(1–4), 304–318 (2003). https://doi.org/10.1016/S0167-2789(03)00227-6
22. Lopez, L., Politi, T.: Applications of the Cayley approach in the numerical solution of matrix differential systems on quadratic groups. Appl. Numer. Math. **36**(1), 35–55 (2001). https://doi.org/10.1016/S0168-9274(99)00049-5
23. Lubich, C.: From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis. European Mathematical Society (EMS), Zürich (2008). https://doi.org/10.4171/067
24. Munthe-Kaas, H.: Lie–Butcher theory for Runge–Kutta methods. BIT **35**(4), 572–587 (1995). https://doi.org/10.1007/BF01739828
25. Munthe-Kaas, H.: Runge–Kutta methods on Lie groups. BIT **38**(1), 92–111 (1998). https://doi.org/10.1007/BF02510919
26. Munthe-Kaas, H.: High order Runge–Kutta methods on manifolds. Appl. Numer. Math. **29**(1), 115–127 (1999). https://doi.org/10.1016/S0168-9274(98)00030-0
27. Munthe-Kaas, H., Zanna, A.: Numerical integration of differential equations on homogeneous manifolds. Found. Comput. Math. (1997). https://doi.org/10.1007/978-3-642-60539-0_24
28. Musharbash, E., Nobile, F.: Symplectic Dynamical Low Rank approximation of wave equations with random parameters. Technical report 18.2017. EPFL, Switzerland (2017)
29. Musharbash, E., Nobile, F., Zhou, T.: Error analysis of the dynamically orthogonal approximation of time dependent random PDEs. SIAM J. Sci. Comput. **37**(2), A776–A810 (2015). https://doi.org/10.1137/140967787
30. Peng, L., Mohseni, K.: Symplectic model reduction of Hamiltonian systems. SIAM J. Sci. Comput. **38**(1), A1–A27 (2016). https://doi.org/10.1137/140978922
31. Sapsis, T.P., Lermusiaux, P.F.J.: Dynamically orthogonal field equations for continuous stochastic dynamical systems. Phys. D **238**(23–24), 2347–2360 (2009). https://doi.org/10.1016/j.physd.2009.09.017
32. Stefănescu, R., Sandu, A., Navon, I.M.: Comparison of POD reduced order strategies for the nonlinear 2D shallow water equations. Int. J. Numer. Methods Fluids **76**(8), 497–521 (2014). https://doi.org/10.1002/fld.3946
33. Woodbury, M.A.: Inverting Modified Matrices. Princeton University, Princeton (1950)