Dynamical Reduced Basis Methods for Hamiltonian Systems

Cecilia Pagliantini*

October 21, 2019. Updated on November 2019.

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.

MSC 2010. 37N30, 65P10, 15A24, 78M34.

Keywords. Hamiltonian dynamics, symplectic manifolds, dynamical low-rank approximation, reduced basis methods (RBM), Lie group integrators.

1 Introduction

We consider parameterized finite-dimensional canonical Hamiltonian systems describing nondissipative flows or ensuing from the numerical discretization of partial differential equations that can be derived from action principles. Let $\mathcal{T} := (t_0, T]$ be a temporal interval and let $\mathcal{V}_{2N}$ be a $2N$-dimensional symplectic 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 $\mathcal{V}_{2N}$.

In the context of long-time and many-query simulations model order reduction aims at pruning the computational cost of solving systems like (1.1) by replacing the original high-dimensional problem with a simplified model, such as a surrogate or low-fidelity model, without compromising the overall

*Chair of Computational Mathematics and Simulation Science (MCSS), École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. Email: `cecilia.pagliantini@epfl.ch`

accuracy. 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.

Reduced basis methods (RBM) relying on projection techniques 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 [15] and [5] using a variational Lagrangian formulation of the problem, in [23, 2, 3] for canonically symplectic dynamical systems, and in [12] to deal with Hamiltonian problems whose phase space is endowed with a nonlinear 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 [22]. Hence, local reduced spaces seems 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

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

where the reduced basis $\{U_i\}_i \subset \mathbb{R}^{2N}$, and the expansion coefficients $\{Z_i\}_i \subset \mathbb{R}$ 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 $\{U_i\}_i$ 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. [17]. In the finite dimensional setting, a similar approach, known as dynamical low-rank approximation [14], 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 [24] 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 [21]. 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 [10]. To the best of our knowledge the only work to address structure-preserving dynamical low-rank approximations is the work by Musharbash and Nobile [20], where the authors develop a DO discretization of stochastic PDEs possessing a symplectic Hamiltonian structure. The method proposed in [20] 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 [10] and yields an evolution equation for the reduced solution analogous to [20], 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 $\{Z_i\}_i$ 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. [11, 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.* [11, Section 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. In the first algorithm we extend the initial value problem for the reduced basis $\{U_i\}_i$ on the orthosymplectic matrix manifold to a quadratic Lie group and, similarly to [16], use Lie group methods and conventional multi-stage explicit RK integrators to solve the equivalent system on the corresponding Lie algebra. By exploiting the structure of the 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 [7, 8] on intrinsic temporal integrators for orthogonal flows.

The remainder of the paper is organized as follows. In Section 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 Section 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 Section 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 Section 5 efficient and structure-preserving temporal integrators for the reduced basis evolution problem are derived. We present some concluding remarks and open questions in Section 6.

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 structure). Let $\mathcal{V}_{2N}$ be a finite $2N$-dimensional smooth manifold. A *symplectic structure* on $\mathcal{V}_{2N}$ is a nondegenerate closed 2-form ω. A manifold $\mathcal{V}_{2N}$ endowed with a symplectic structure ω is called a symplectic manifold and denoted with $(\mathcal{V}_{2N}, \omega)$. If $\mathcal{V}_{2N}$ is a vector space, then $(\mathcal{V}_{2N}, \omega)$ is called symplectic vector space.

The algebraic structure of a symplectic manifold $(\mathcal{V}_{2N}, \omega)$ can be characterized through the definition of a bracket: Let $\mathsf{d}\mathcal{F}$ be the 1-form given by the exterior derivative of a given smooth function $\mathcal{F} \in C^\infty(\mathcal{V}_{2N})$, then

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

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 application $\mathcal{J}_{2N}: T^*\mathcal{V}_{2N} \to T\mathcal{V}_{2N}$ 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].

A Hamiltonian system is a triple $(\mathcal{V}_{2N}, \omega, \mathcal{H})$ where $(\mathcal{V}_{2N}, \omega)$ is a symplectic manifold and $\mathcal{H} \in C^\infty(\mathcal{V}_{2N})$ is a smooth function, called *Hamiltonian*, such that

$$d\mathcal{H} = \mathsf{i}_{\mathcal{X}_\mathcal{H}}\omega, \tag{2.2}$$

where i denotes the contraction operator and $\mathcal{X}_\mathcal{H} \in T\mathcal{V}_{2N}$. Since ω is nondegenerate, the vector field $\mathcal{X}_\mathcal{H} \in T\mathcal{V}_{2N}$, called *Hamiltonian vector field*, is unique. A 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 map $\Phi^t_{\mathcal{X}_\mathcal{H}}$ of a vector field $\mathcal{X}_\mathcal{H} \in T\mathcal{V}_{2N}$ is Hamiltonian if and only if $\Phi^t_{\mathcal{X}_\mathcal{H}}$ is a symplectic diffeomorphism (symplectomorphism) on its domain, i.e., for each $t \in \mathcal{T}$, $(\Phi^t_{\mathcal{X}_\mathcal{H}})^*\omega = \omega$.

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 $\psi(u) = (q^1(u),\dots,q^N(u),p_1(u),\dots,p_N(u))$, 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 $i,j = 1,\dots,N$.

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 define a global basis that is symplectic and orthonormal.

Definition 2.5 (Orthosymplectic basis)**.** Let $(\mathcal{V}_{2N}, \omega)$ be a $2N$-dimensional symplectic vector space and let ω be the canonical symplectic form. Then, the set of vectors $\{e_i\}_{i=1}^{2N}$ is said to be *orthosymplectic* in $\mathcal{V}_{2N}$ if

$$\omega(e_i,e_j) = (J_{2N})_{i,j}\,, \quad \text{and} \quad (e_i,e_j) = \delta_{i,j}\,, \qquad \forall i,j = 1\dots,2N,$$

where $(\cdot,\cdot)$ is the Euclidean inner product and J_{2N} is the canonical symplectic tensor on $\mathcal{V}_{2N}$.

Let $(\mathcal{V}_{2N}, \omega)$ be a symplectic vector space, and let $\mathcal{U}$ be a subspace of $\mathcal{V}_{2N}$. The symplectic complement of $\mathcal{U}$ in $\mathcal{V}_{2N}$ is the set of $u \in \mathcal{V}_{2N}$ such that $\omega(u, v) = 0$ for all $v \in T\mathcal{V}_{2N} \cong \mathcal{V}_{2N}$. A subspace of a symplectic vector space $(\mathcal{V}_{2N}, \omega)$ is called *Lagrangian* if it coincides with its symplectic complement in $\mathcal{V}_{2N}$. As a consequence of the fact that any basis of a Lagrangian subspace of a symplectic vector space $(\mathcal{V}_{2N}, \omega)$ can be extended to a symplectic basis in $(\mathcal{V}_{2N}, \omega)$, every symplectic vector space admits an orthosymplectic basis, *cf.* for example [4, Section 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

$$\begin{cases} \partial_t u(t, \eta) = J_{2N} \nabla_u \mathcal{H}(u(t, \eta); \eta), & \text{for } t \in \mathcal{T}, \\ u(t_0, \eta) = u_0(\eta), \end{cases} \tag{2.3}$$

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 (2.3) 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 symplectic 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 mapping $M_+ : (\mathcal{V}_{2N}, \omega) \to (\mathcal{V}_{2n}, \omega)$ is symplectic, in the sense of Definition 2.2, if and only if $M_+ J_{2N} M_+^\top = J_{2n}$.*

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 equivalently $M^\top M = I_{2n}$. Moreover, 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 $M \in \mathrm{Sp}(2n, \mathbb{R}^{2N}) := \{L \in \mathbb{R}^{2N \times 2n} :\ L^\top J_{2N} L = J_{2n}\}$.

In this work we consider symplectic spaces spanned by basis which are also orthonormal.

Definition 3.2. A matrix $M \in \mathbb{R}^{2N \times 2n}$ is called *orthosymplectic* if $M \in \mathcal{U}(2n, \mathbb{R}^{2N}) := \mathrm{St}(2n, \mathbb{R}^{2N}) \cap \mathrm{Sp}(2n, \mathbb{R}^{2N})$, 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 $M = [A \,|\, B]$ 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, $M = [A \,|\, J_{2N}^\top A]$ with $A^\top A = I_n$ and $A^\top J_{2N} A = 0_n$. The definition of symplectic inverse yields $M_+^\top = J_{2N}^\top M J_{2n} = J_{2N}^\top [A \,|\, J_{2N}^\top A] J_{2n} = [J_{2N}^\top A \,|\, -A] J_{2n} = [A \,|\, J_{2N}^\top A] = M$.

Conversely, the symplecticity of M_+ implies $M_+ M_+^\top = M_+ M = M_+ J_{2N} M_+^\top J_{2n}^\top = I_{2n}$. ∎

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

$$\mathfrak{sp}(2n) := \{M \in \mathbb{R}^{2n\times 2n} :\ MJ_{2n} + J_{2n}M^\top = 0_{2n}\}.$$

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 the matrix commutator as bracket.

4 Orthosymplectic dynamical reduced basis method

Assume we want to solve the parameterized Hamiltonian problem (2.3) at $p\in\mathbb{N}$ samples of the parameter $\{\eta_j\}_{j=1}^p =: \Gamma_h \subset \mathbb{R}^{pd}$. To ease 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 (2.3) 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 $\mathcal{R}_0(\eta_h) := \big[u_0(\eta_1)|\ldots|u_0(\eta_p)\big] \in \mathbb{R}^{2N\times p}$, 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

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

where $U = \big[\mathbf{U}_1|\ldots|\mathbf{U}_{2n}\big]\in\mathbb{R}^{2N\times 2n}$, 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

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

where

$$\begin{aligned} \mathcal{M} &:= \mathcal{U}(2n,\mathbb{R}^{2N}) = \{U\in\mathbb{R}^{2N\times 2n} :\ U^\top U = I_{2n},\ U^\top J_{2N}U = J_{2n}\},\\ V^{p\times 2n} &:= \{Z\in\mathbb{R}^{p\times 2n} :\ \operatorname{rank}(Z^\top Z + J_{2n}^\top Z^\top Z J_{2n}) = 2n\}. \end{aligned} \tag{4.4}$$

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 [20, 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),

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

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 [14, Section 5.3].

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, $(\mathcal{M} \times V^{p\times 2n}, \mathcal{M}_{2n}^{\text{spl}}, \phi)$ is a fiber bundle with fibers given by the group of unitary matrices $\mathcal{U}(2n)$, so that $\mathcal{M}_{2n}^{\text{spl}}$ 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 [24] a unique characterization of the reduced solution is obtained by fixing a gauge constraint in the tangent space of the reduced solution manifold. For the manifold $\mathcal{M}_{2n}^{\text{spl}}$ the tangent space at $R \in \mathcal{M}_{2n}^{\text{spl}}$ 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 $U(t)Z^\top(t) \in \mathcal{M}_{2n}^{\text{spl}}$ 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 $U \in \text{Sp}(2n, \mathbb{R}^{2N})$. Therefore, the tangent space of $\mathcal{M}_{2n}^{\text{spl}}$ at $UZ^\top$ is defined as

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

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

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

This means that the map

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

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 $X_U Z^\top + U X_Z^\top \in T_{UZ^\top}\mathcal{M}_{2n}^{\text{spl}}$.

To show that the map Ψ is injective, we take $X = 0 \in T_{UZ^\top}\mathcal{M}_{2n}^{\text{spl}}$. 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

$$0 = X_U Z^\top Z + J_{2N} X_U Z^\top Z J_{2n}^\top = X_U Z^\top Z + J_{2N}^\top X_U J_{2n} J_{2n} Z^\top Z J_{2n}^\top = X_U(Z^\top Z + J_{2n} Z^\top Z J_{2n}^\top),$$

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

For the surjectivity of Ψ we show that

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

Any $X \in T_{UZ^\top}\mathcal{M}_{2n}^{\text{spl}}$ 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

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

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 horizontal space $H_{(U,Z)}$. From the orthogonality of U it easily follows that

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

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

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

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

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

□

Proposition 4.1 entails that 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.

4.1 Dynamical low-rank symplectic variational principle

For $\eta \in \Gamma$ fixed in (2.3), the vector field $\mathcal{X}_{\mathcal{H}}$ at time t belongs to $T_{u(t)}\mathcal{V}_{2N}$. Taking the cue from dynamical low-rank approximations [14], the dynamics on the reduced space $\mathcal{M}_{2n}^{\text{spl}}$ is obtained 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 $\mathcal{X}_{\mathcal{H}}$ at every point on the manifold $\mathcal{V}_{2N}$ in the Frobenius norm. 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 4.2. *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 $R := UZ^\top \in \mathcal{M}_{2n}^{\text{spl}}$. Then the map*

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

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 $S := Z^\top Z + J_{2n}Z^\top ZJ_{2n} \in \mathbb{R}^{2n\times 2n}$ and $F_w := wZ + J_{2N}wZJ_{2n}^\top \in \mathbb{R}^{2N\times 2n}$, then

$$\begin{aligned} X_U(\Pi w) &= (I_{2N} - UU^\top)\big((I_{2N} - UU^\top)F_wS^{-1}Z^\top Z + J_{2N}(I_{2N} - UU^\top)F_wS^{-1}Z^\top ZJ_{2n}^\top\big)S^{-1} \\ &= X_U(w)Z^\top ZS^{-1} + J_{2N}X_U(w)Z^\top ZJ_{2n}^\top S^{-1}. \end{aligned}$$

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 $y = Y_UZ^\top + UY_Z^\top$ where $Y_Z \in \mathbb{R}^{1\times 2n}$ and $Y_U \in H_U$. Therefore,

$$\omega(v - \Pi v, y) = \omega(v - \Pi v, Y_UZ^\top) + \omega(v, UY_Z^\top) - \omega(X_UZ^\top + UX_Z^\top, UY_Z^\top),$$

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

$$\omega(UX_Z^\top, UY_Z^\top) = \omega(UU^\top v, UY_Z^\top) = \omega(v, J_{2N}^\top UJ_{2n}U^\top UY_Z^\top) = \omega(v, J_{2N}^\top UJ_{2n}Y_Z^\top) = \omega(v, UY_Z^\top).$$

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

$$\omega(v - \Pi v, 2y) = \omega(v, Y_UZ^\top + J_{2N}^\top Y_UJ_{2n}Z^\top) - \omega(X_UZ^\top + UX_Z^\top, Y_UZ^\top + J_{2N}^\top Y_UJ_{2n}Z^\top) =: T_1 - T_2.$$

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

$$\begin{aligned} T_1 &= \omega\Big(v, \sum_{i=1}^{2n}(Y_U)_iZ_i\Big) + \omega\Big(J_{2N}v, \sum_{i=1}^{2n}(Y_U)_i(J_{2n}Z^\top)_i\Big) \\ &= \sum_{i=1}^{2n}\omega\big(v, (Y_U)_i\big)Z_i + \sum_{i=1}^{2n}\omega\big(J_{2N}v, (Y_U)_i\big)(J_{2n}Z^\top)_i = \sum_{i=1}^{2n}\omega\big(vZ_i + J_{2N}v(ZJ_{2n}^\top)_i, (Y_U)_i\big). \end{aligned}$$

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

$$\begin{aligned} T_2 &= \omega\big(X_UZ^\top, Y_UZ^\top\big) + \omega\big(X_UJ_{2n}Z^\top, Y_UJ_{2n}Z^\top\big) \\ &= \sum_{i,j=1}^{2n}\omega\big((X_U)_jZ_j, (Y_U)_iZ_i\big) + \omega\big((X_U)_j(J_{2n}Z^\top)_j, (Y_U)_i(J_{2n}Z^\top)_i\big) \\ &= \sum_{i,j=1}^{2n}\omega\big((X_U)_j, (Y_U)_i\big)\big(Z_jZ_i + (J_{2n}Z^\top)_j(ZJ_{2n}^\top)_i\big). \end{aligned}$$

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

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 $\text{St}(n, \mathbb{C}^N)$, *cf.* Lemma 4.6. 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 $R \in C^1(\mathcal{T}, \mathcal{M}_{2n}^{\text{spl}})$ such that

$$\begin{cases} \dot{R}(t) = \Pi_{T_R\mathcal{M}_{2n}^{\text{spl}}}\mathcal{X}_{\mathcal{H}}(R(t), \eta_h), & \text{for } t \in \mathcal{T}, \\ R(t_0) = U_0Z_0^\top. \end{cases} \tag{4.9}$$

For any $1 \leq j \leq p$ 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 $Y(t) := [Y_1|\ldots|Y_p] \in \mathbb{R}^{2N\times p}$ where $Y_j := \nabla_{UZ_j^\top}\mathcal{H}(UZ_j^\top, \eta_j) \in \mathbb{R}^{2N\times 1}$, and $\nabla_{UZ_j^\top}$ denotes the gradient with respect to $UZ_j^\top$. Using the decomposition $R = UZ^\top$ 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

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

The reduced problem (4.10) is analogous to the system derived in [20, 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 $1 \leq j \leq p$, 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 4.3. *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, let $S := Z^\top Z + J_{2n}Z^\top Z J_{2n}^\top \in \mathbb{R}^{2n\times 2n}$, since $SJ_{2n} = J_{2n}S$, the constraint $\dot{U} = J_{2N}^\top \dot{U} J_{2n}$ is satisfied if $\dot{U}SJ_{2n}^\top = J_{2N}^\top \dot{U}S$. One can easily show that $A := J_{2N}YZ - YZJ_{2n}^\top = J_{2N}AJ_{2n}^\top$. Therefore, $\dot{U}SJ_{2n}^\top = (I_{2N} - UU^\top)AJ_{2n}^\top = J_{2N}^\top(I_{2N} - UU^\top)J_{2N}AJ_{2n}^\top = J_{2N}^\top \dot{U}S$. ☐

Remark 4.4. 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 [12, Section 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 Section 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 4.2, 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,

$$\sum_{j=1}^p \frac{d}{dt}\mathcal{H}_j(R(t)) = \sum_{j=1}^p (\nabla_{R_j}\mathcal{H}_j(R), \dot{R}_j) = \sum_{j=1}^p \omega(J_{2N}\nabla_{R_j}\mathcal{H}_j(R), \dot{R}_j) = \sum_{j=1}^p \omega_j(\dot{R}, \dot{R}) = 0.$$

Therefore, if $\mathcal{R}_0(\eta_h) \in \text{span}\{U_0\}$ 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 $\mathbb{R}^{2N} \to \text{span}\{U(t)\}$ 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 [12, Lemma 3.9] holds.

Lemma 4.5. *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 $\mathcal{H} \in \text{Im}(\pi^*_{+,t})$ for any $t \in \mathcal{T}$. Then, $\mathcal{I}(t) \in C^\infty(\mathbb{R}^{2n})$ is an invariant of $\Phi^t_{X_{\pi^*_{+,t}\mathcal{H}}}$ if and only if $(\pi^*_{+,t}\mathcal{I})(t) \in C^\infty(\mathbb{R}^{2N})$ is an invariant of $\Phi^t_{X_\mathcal{H}}$ in $\text{Im}(\pi^*_{+,t})$.*

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 [10, Section 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 $\mathcal{M}^{\text{spl}}_{2n}$ defined in (4.3) and the manifold $\mathcal{M}_n$ of rank-n complex matrices, already established in [20, 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 [10] 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

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

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 $\mathcal{M}_n := \{C \in \mathbb{C}^{N\times p} : \ \text{rank}(C) = n\}$. Any $C \in \mathcal{M}_n$ can be decomposed, up to unitary $n \times n$ transformations, as $C = WY^\top$ where $W \in \text{St}(n, \mathbb{C}^N) = \{M \in \mathbb{C}^{N\times n} : \ M^\mathsf{H} M = I_n\}$, and $Y \in \mathcal{V}^{p\times n} := \{M \in \mathbb{C}^{p\times n} : \ \text{rank}(M) = n\}$. Analogously to [20, Lemma 6.1] one can establish the following result.

Lemma 4.6. *The manifolds $\mathcal{M}_n$ and $\mathcal{M}^{\text{spl}}_{2n}$ are isomorphic via the map*

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

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

For $C(t_0) \in \mathcal{M}_n$ associated with $R(t_0) \in \mathcal{M}^{\text{spl}}_{2n}$ 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

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

The so-called dynamically orthogonal condition $X_W^{\mathsf{H}} W = 0$, 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 4.7. *At every $C = WY^\top \in \mathcal{M}_n$, the map*

$$\begin{aligned} \Pi_{T_C\mathcal{M}_n} : \mathbb{C}^{N\times p} &\longrightarrow T_C\mathcal{M}_n \\ w &\longmapsto (I_N - WW^{\mathsf{H}})w\, Y^*(Y^\top Y^*)^{-1} Y^\top + WW^{\mathsf{H}} w, \end{aligned} \tag{4.15}$$

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 [10, 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 $X_W^{\mathsf{H}} W = 0$. □

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

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

Proposition 4.8. *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). □

In view of Proposition 4.8, we can revert to the error estimate established in [10].

Theorem 4.9 ([10, 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

$$\|C(t) - \Pi_{\mathcal{M}_n}\mathcal{C}(t)\| \le \int_{\mathcal{T}} \left(L_{\mathcal{X}} + \frac{\|\mathcal{X}_{\mathcal{H}}(\mathcal{C}(s), \eta_h)\|}{\sigma_n(\mathcal{C}(s)) - \sigma_{n+1}(\mathcal{C}(s))} \right) \|\mathcal{C}(s) - \Pi_{\mathcal{M}_n}\mathcal{C}(s)\| e^{\mu(t-s)}\, ds,$$

where $L_{\mathcal{X}} \in \mathbb{R}$ denotes the Lipschitz continuity constant of $\mathcal{X}_{\mathcal{H}}$ and $\mu \in \mathbb{R}$ is defined as

$$\mu := L_{\mathcal{X}} + 2 \sup_{t\in\mathcal{T}} \frac{\|\mathcal{X}_{\mathcal{H}}(\mathcal{C}(t), \eta_h)\|}{\sigma_n(\mathcal{C}(t))}.$$

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. [11, Section 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 then be performed using symplectic integrators, *cf.* [11, Section VI].

Observe that using 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 associated with the projection step, which 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 a future work the investigation and the numerical study of splitting methods preserving 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 ease 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 $t \in \mathcal{T}$, the function $\mathcal{F} : \mathcal{M} \to T\mathcal{M}$ in (5.1) is defined as

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

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 4.3, 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 [11, Section IV.5]. The analyticity and the favorable computational properties of the Cayley transform, *cf.* Proposition 5.2 and [13], 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. [11, Section 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 $\mathrm{cay} : \mathfrak{g}_{2N} \to \mathbb{R}^{2N\times 2N}$ be the Cayley transform defined as*

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

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* $M \in \mathfrak{g}_{2N}$ *is the map* $\mathrm{dcay}_M : T_M\mathfrak{g}_{2N} \cong \mathfrak{g}_{2N} \to T_{\mathrm{cay}(M)}\mathcal{G}_{2N}$,

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

and its inverse is

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

(iii) [9, Theorem 3] Let $\sigma(A)$ denote the spectrum of $A \in \mathbb{R}^{2N\times 2N}$. If $M \in C^1(\mathbb{R}, \mathfrak{g}_{2N})$ then $A := \mathrm{cay}(M) \in C^1(\mathbb{R}, \mathcal{G}_{2N})$. 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 $M \in C^1(\mathbb{R}, \mathfrak{g}_{2N})$ such that $M = \mathrm{cay}^{-1}(A) = 2(A - I_{2N})(A + I_{2N})^{-1}$.

Proof. Let $M \in \mathfrak{g}_{2N}$ and let $\overline{M} := M/2$. Since M is skew-symmetric then $I - \overline{M}$ is invertible.
(i) The Cayley transform defined in (5.3) can be recast as

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

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

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

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

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

(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.* [11, Section 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 Section 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 $M \in \mathfrak{g}_{2N}$ and let $k \in \mathbb{N}$, $k < 2N$. If M admits the low-rank splitting*

$$M = \alpha\beta^\top, \qquad \alpha,\, \beta \in \mathbb{R}^{2N\times k}, \tag{5.6}$$

then, for any $Y \in \mathbb{R}^{2N\times k}$, $\mathrm{cay}(M)Y \in \mathbb{R}^{2N\times k}$ can be evaluated with computational complexity of order $O(Nk^2) + O(k^3)$.

Proof. To evaluate the Cayley transform of $M = \alpha\beta^\top$ 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 is analytic with a pole at $z = 0$ and its Taylor expansion reads

$$f(z) = z^{-1}(\mathrm{cay}(z) - 1) = \sum_{m=0}^{\infty} 2^{-m} z^m.$$

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

$$\mathrm{cay}(M) = I_{2N} + \sum_{m=1}^{\infty} 2^{1-m} M^m = I_{2N} + \sum_{m=1}^{\infty} 2^{1-m}\alpha(\beta^\top\alpha)^{m-1}\beta^\top = I_{2N} + \alpha f(\beta^\top\alpha)\beta^\top.$$

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

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

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 $O(Nk^2)$ 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 [6, Proposition 4], one could alternatively envision the use of Woodbury matrix identity [25] to compute the matrix inverse appearing in the definition (5.3) of the Cayley transform. This yields the formula

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

which can be evaluated in $O(Nk^2) + O(k^3)$ operations. □

5.2 Numerical integrators based on Lie groups

In this Section we propose a numerical scheme for the solution of (5.1) based on Lie group methods. The idea is to extend the rectangular system (5.1) on $\mathcal{M}$ to the Lie group $\mathcal{U}(2N)$ of square orthosymplectic matrices. A local coordinate map, as the Cayley transform, is employed to derive a differential equation on the Lie algebra $\mathfrak{g}_{2N}$, that is then solved using RK methods. Indeed, since RK methods preserve linear invariants, they allow to derive discrete trajectories that remain in the Lie algebra.

The method we develop extends to the orthosymplectic manifold the scheme proposed in [16] for the numerical solution of matrix differential systems on quadratic groups. The main computational cost of this approach rests on the evaluation of the coordinate map at each RK stage in every temporal interval. However, we derive an extension of problem (5.1) that possesses a low-rank structure amenable to efficient algorithms for its numerical approximation.

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*

$$\begin{cases} \dot{U}(t) = \mathcal{L}(U(t))\,U(t), & \text{for } t\in\mathcal{T},\\ U(t_0) = Q, \end{cases} \tag{5.7}$$

where $\mathcal{L}:\mathcal{M}\to\mathbb{R}^{2N\times 2N}$ is a skew-symmetric and Hamiltonian operator defined as

$$\mathcal{L}(U) = \frac{1}{2}\big(\mathcal{S}(U) + J_{2N}^\top\mathcal{S}(U)J_{2N}\big), \qquad \mathcal{S}(U) := (I_{2N} - UU^\top)\mathcal{F}(U)U^\top - U\mathcal{F}(U)^\top. \tag{5.8}$$

If $\mathcal{F}:\mathcal{M}\to T\mathcal{M}$ is defined as in (5.2) then,

$$\mathcal{L}(U) = \mathcal{F}(U)U^\top - U\mathcal{F}(U)^\top. \tag{5.9}$$

Proof. First, observe that by introducing the function $\mathcal{A}:\mathcal{M}\to\mathbb{R}^{2N\times 2N}$ defined as $\mathcal{A}(U) := \mathcal{F}(U)U^\top$ for any $U\in\mathcal{M}$, the evolution equation (5.1) can be trivially recast as $\dot{U}(t) = \mathcal{A}(U(t))U(t)$ for all $t\in\mathcal{T}$. By deriving (in time) the orthogonality and symplecticity constraints on U, it can be easily shown that the operator $\mathcal{A}$ is weakly skew-symmetric and weakly Hamiltonian, that is

$$U^\top\big(\mathcal{A}(U) + \mathcal{A}(U)^\top\big)U = 0, \qquad U^\top\big(\mathcal{A}(U)^\top J_{2N} + J_{2N}\mathcal{A}(U)\big)U = 0, \qquad \forall\, U\in\mathcal{M}.$$

Let us now 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 $Y(t) = [U(t)\,|\,W(t)]\in\mathcal{U}(2N)$. Since $Y(t)$ is orthosymplectic by construction, it holds

$$\begin{aligned} 0 &= \frac{d}{dt}(Y^\top Y) = \dot{Y}^\top Y + Y^\top\dot{Y}, &&\Longrightarrow&& \dot{Y} = -Y\dot{Y}^\top Y,\\ 0 &= \frac{d}{dt}(Y^\top J_{2N}Y) = \dot{Y}^\top J_{2N}Y + Y^\top J_{2N}\dot{Y}, &&\Longrightarrow&& \dot{Y} = -J_{2N}^\top Y\dot{Y}^\top J_{2N}Y. \end{aligned} \tag{5.10}$$

From (5.10) it follows that $\dot{Y} = \mathcal{B}(Y,\dot{Y})Y$ for all $t \in \mathcal{T}$ with

$$\mathcal{B}(Y,\dot{Y}) = -\frac{1}{2}\big(Y\dot{Y}^\top + J_{2N}^\top Y\dot{Y}^\top J_{2N}\big).$$

The orthogonality of Y implies that $\mathcal{B}(Y,\dot{Y}) \in \mathbb{R}^{2N\times 2N}$ is skew-symmetric and Hamiltonian. Moreover, by writing $\mathcal{B}(Y,\dot{Y})$ explicitly in terms of U and W, and using the evolution equation satisfied by U, yields

$$\mathcal{B}(Y,\dot{Y}) = -\frac{1}{2}\big(UU^\top\mathcal{A}(U)^\top + J_{2N}^\top UU^\top\mathcal{A}(U)^\top J_{2N} + W\dot{W}^\top + J_{2N}^\top W\dot{W}^\top J_{2N}\big). \tag{5.11}$$

The skew-symmetric condition $\mathcal{B}(Y,\dot{Y}) + \mathcal{B}(Y,\dot{Y})^\top = 0$ can be written as

$$\dot{W}W^\top + W\dot{W}^\top = -\mathcal{A}(U)UU^\top - UU^\top\mathcal{A}(U)^\top. \tag{5.12}$$

If $W \in \mathbb{R}^{2N\times 2(N-n)}$ is such that $\dot{W}W^\top = -UU^\top\mathcal{A}(U)^\top(I_{2N} - UU^\top)$, then it satisfies (5.12) owing to the weak skew-symmetry of $\mathcal{A}$. Substituting this expression in (5.11) yields

$$\mathcal{L}(U) = \mathcal{B}(Y,\dot{Y}) = \frac{1}{2}\big(\mathcal{S}(U) + J_{2N}^\top\mathcal{S}(U)J_{2N}\big),$$

with $\mathcal{S}(U) := -UU^\top\mathcal{A}(U)^\top + \mathcal{A}(U)UU^\top - UU^\top\mathcal{A}(U)UU^\top$ and the expression (5.8) is recovered.

In particular, if $\mathcal{F}$ is defined as in (5.2), then $U^\top\mathcal{F}(U) = 0$ for any $U \in \mathcal{M}$. Substituting in (5.8) and using the fact that $J_{2N}^\top\mathcal{F}(U)U^\top J_{2N} = \mathcal{F}(U)U^\top$ yields (5.9). □

Analogously to [9, Theorem 5] it can be shown that, if $U(t) \in \mathcal{M}$ is solution of (5.7) with $-1 \notin \bigcup_{t\in\mathcal{T}}\sigma(U(t))$, then $U(t) = \mathrm{cay}(M(t))U(t_0)$ where $M(t) \in \mathfrak{g}_{2N}$ satisfies

$$\begin{cases} \dot{M}(t) = \mathrm{dcay}_{M(t)}^{-1}\big(\mathcal{L}\big(U(t)\big)\big), & \text{for } t \in \mathcal{T},\\ M(t_0) = \mathrm{cay}^{-1}\big([\,U(t_0)\,|\,W_0]\big). \end{cases} \tag{5.13}$$

Here $W_0 \in \mathbb{R}^{2N\times 2(N-n)}$ provides an orthosymplectic extension of $U(t_0)$ as described in the proof of Proposition 5.3. Note that, since $\mathrm{dcay}_{M(t)}^{-1} : T_{\mathrm{cay}(M)}\mathcal{G}_{2N} \to \mathfrak{g}_{2N}$, the solution of (5.13) belongs to $\mathfrak{g}_{2N}$ at any time $t \in \mathcal{T}$. In order to numerically solve (5.13) on the Lie algebra $\mathfrak{g}_{2N}$, one can then apply Runge–Kutta methods analogous to [16, Section 2], which we briefly report here for the sake of completeness. The skew-symmetric and Hamiltonian conditions on M are satisfied at any $t \in \mathcal{T}$ since RK methods preserve linear invariants (besides explicit methods only involve linear combinations of elements belonging to the Lie algebra). 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 an s-stage explicit Runge–Kutta method. Then, the numerical approximation of (5.13) in the interval $(t^m, t^{m+1}]$ reads: Given $U_m \in \mathcal{M}$, compute

$$\begin{aligned} M_{m+1} &= \Delta t\sum_{i=1}^{s} b_i\,\mathrm{dcay}_{M_m^i}^{-1}\big(\mathcal{L}(U_m^i)\big),\\ &\quad M_m^1 = 0,\\ &\quad M_m^i = \Delta t\sum_{j=1}^{i-1} a_{i,j}\,\mathrm{dcay}_{M_m^j}^{-1}\big(\mathcal{L}(U_m^j)\big), && i = 2,\ldots,s,\\ &\quad U_m^i = \mathrm{cay}(M_m^i)U_m, && i = 1,\ldots,s,\\ U_{m+1} &= \mathrm{cay}(M_{m+1})U_m. \end{aligned} \tag{5.14}$$

Note that choosing a sufficiently small time step for the temporal integrator can prevent the numerical solution of of (5.13) from having an eigenvalue close to -1 for some $t \in \mathcal{T}$. Alternatively, restarting procedures of the algorithm (5.14) can be implemented similarly to [9, pp. 323-324].

The computational cost of algorithm (5.14) is assessed in the following result.

Proposition 5.4. *In any fixed temporal interval $(t^m, t^{m+1}] \subset \mathcal{T}$, the Cayley Runge–Kutta algorithm (5.14) for the numerical solution of the evolution problem (5.13) has computational complexity of order $O(Nn^2s^5) + O(n^3s^7) + C_{\mathcal{F}}$, where s is the number of stages of the Runge–Kutta method, and $C_{\mathcal{F}}$ is the complexity of the algorithm to compute $\mathcal{F}(U)$ in (5.2) at any given $U \in \mathcal{M}$.*

Proof. Assume that the temporal interval $(t^m, t^{m+1}]$ is fixed. In view of Proposition 5.2, we want to show that the terms $\{M_m^i\}_{i=1}^s$ and M_{m+1} admit a low-rank splitting of the form (5.6) and find the dimension k of this splitting.

Each term $\{\mathcal{L}(U_m^i)\}_{i=1}^s$, with $\mathcal{L}$ defined in (5.9), can be written as $\mathcal{L}(U_m^i) = c_i d_i^\top$ where

$$c_i := \big[\mathcal{F}(U_m^i) \,|\, -U_m^i\big] \in \mathbb{R}^{2N\times 4n}, \qquad d_i := \big[U_m^i \,|\, \mathcal{F}(U_m^i)\big] \in \mathbb{R}^{2N\times 4n}. \tag{5.15}$$

For any $i = 1, \dots, s$ let us introduce the $2N \times 2N$ matrix

$$A_i := \mathrm{dcay}^{-1}_{M_m^i}\big(\mathcal{L}(U_m^i)\big) = \mathrm{dcay}^{-1}_{M_m^i}(c_i d_i^\top).$$

We prove that $A_i = e_i f_i^\top$ with $e_i, f_i \in \mathbb{R}^{2N\times 4ni}$. For $i = 1$, $M_m^1 = 0$ and hence $A_1 = \mathcal{L}(U_m^1) = c_1 d_1^\top =: e_1 f_1^\top$ owing to (5.15). For $i = 2$, it holds

$$M_m^2 = \Delta t\, a_{2,1} A_1 = \Delta t\, a_{2,1} e_1 f_1^\top.$$

Hence, $A_2 = \mathrm{dcay}^{-1}_{M_m^2}(c_2 d_2^\top) = e_2 f_2^\top$ with $e_2, f_2 \in \mathbb{R}^{2N\times 2(4n)}$ defined as

$$\begin{aligned} e_2 &:= \left[c_2 - \frac{\Delta t}{2} a_{2,1} e_1 (f_1^\top c_2) \,\middle|\, \frac{\Delta t}{2} a_{2,1} c_2 (d_2^\top e_1) - \frac{\Delta t^2}{4} a_{2,1}^2 e_1 (f_1^\top c_2 (d_2^\top e_1))\right], \\ f_2 &:= \big[d_2 \,|\, f_1\big]. \end{aligned}$$

Note that the computation of each term in e_2 can be performed in $O(Nn^2)$ operations. Proceeding in this way, we obtain that for any $i \leq s$ it holds

$$A_i = \left(I_{2N} - \frac{M_m^i}{2}\right) c_i d_i^\top \left(I_{2N} + \frac{\Delta t}{2} \sum_{\ell=1}^{i-1} a_{i,\ell} e_\ell f_\ell^\top\right) =: e_i f_i^\top,$$

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

$$\begin{aligned} e_i &:= \left[\left(I_{2N} - \frac{M_m^i}{2}\right) c_i \,\middle|\, \frac{\Delta t}{2} a_{i,1} \left(I_{2N} - \frac{M_m^i}{2}\right) c_i (d_i^\top e_1) \,\middle|\, \dots \,\middle|\, \frac{\Delta t}{2} a_{i,i-1} \left(I_{2N} - \frac{M_m^i}{2}\right) c_i (d_i^\top e_{i-1})\right], \\ f_i &:= \big[d_i \,|\, f_1 \,|\, \dots \,|\, f_{i-1}\big]. \end{aligned}$$

This implies that

$$M_{m+1} = \Delta t \sum_{i=1}^s b_i A_i = \Delta t \sum_{i=1}^s b_i e_i f_i^\top = \Delta t \big[b_1 e_1 \,|\, \dots \,|\, b_s e_s\big] \big[f_1 \,|\, \dots \,|\, f_s\big]^\top =: \alpha \beta^\top,$$

where the matrices α and β have size $2N \times \sum_{i=1}^s 4ni = 2N \times 2ns(s+1)$. Analogously, for every $i = 2, \dots, s$, it holds

$$M_m^i = \Delta t \sum_{j=1}^{i-1} a_{i,j} A_j = \Delta t \sum_{j=1}^{i-1} a_{i,j} e_j f_j^\top = \Delta t \big[a_{i,1} e_1 \,|\, \dots \,|\, a_{i,i-1} e_{i-1}\big] \big[f_1 \,|\, \dots \,|\, f_{i-1}\big]^\top =: \gamma_i \delta_i^\top.$$

The matrices γ_i and δ_i have size $2N \times \sum_{j=1}^{i-1} 4nj = 2N \times 2ni(i-1)$, for any fixed $i = 2, \dots, s$.

In conclusion, we can apply the result of Proposition 5.2 for the computation of U_{m+1} with $k = 2ns(s+1)$ which implies a computational cost of order $O(Nn^2s^4) + O(n^3s^6)$. For the terms $\{U_m^i\}_{i=1}^s$, we proceed with an analogous reasoning and then sum over i the number of operations required to compute each U_m^i. Simple algebraic calculations yield an overall computational cost of order $O(Nn^2s^5) + O(n^3s^7)$. ☐

Despite the scheme is linear in the full dimension, algorithm (5.13) might still incur a high computational cost associated with the dependence on the number s of stages of the RK algorithm. Note that, in the practical implementation of algorithm (5.14), the computational complexity derived in Proposition 5.4 might prove to be pessimistic in terms of the polynomial dependence on s. In principle one can envision to solve (5.1) by recasting problem (5.7) as an evolution equation on the Lie algebra $\mathfrak{g}_{2N}$, similarly to (5.13) but using the matrix exponential exp as coordinate map instead of the Cayley transform, in the spirit of [19]. In this case, one would need to evaluate the exponential map at matrices $M \in \mathbb{R}^{2N\times 2N}$ that admit a low-rank splitting of the form $M = \sum_{i=1}^{r} \alpha_i\beta_i^\top$ for some mild constant $r \in \mathbb{N}$. One can then approximate $\exp(tM) \approx E(tM)$ with the function $E(tM) := \Pi_{i=1}^{r} \exp(t\,\alpha_i\beta_i^\top)$. Similarly to Proposition 5.2, it is possible to show that the product $E(tM)Y$ for $Y \in \mathbb{R}^{2N\times 2n}$ can be performed in $O(Nn^2)$ operations and results in a milder dependence on the number s of the RK scheme. However, in order to guarantee that $\exp(tM)$ belongs to $\mathcal{G}_{2N}$ we need that each low-rank term $\{\alpha_i\beta_i^\top\}_i$ belongs to the Lie algebra $\mathfrak{g}_{2N}$. Moreover, since these terms do not necessarily commute, a truncation of the Baker–Campbell–Hausdorff series would yield only an approximation of the exponential coordinate map. Standard approximations of the exponential map yield methods of order one, and stricter conditions on the factors of the low-rank splitting are required for higher order approximations. These are however not satisfied in general by the low-rank factors $\{\alpha_i\beta_i^\top\}_i$ from Proposition 5.4 obtained in each RK stage. The rather crude estimate in terms of stages of the RK scheme might also be mitigated with techniques that exploit the structure of the operators involved in algorithm (5.14).

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 $O(Nn^2s)$, 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 [7, Section 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

$$U(t) = \mathcal{R}_Q(V(t)), \qquad V(t) \in T_Q\mathcal{M}. \tag{5.16}$$

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 $\mathcal{R}'_Q : TT_Q\mathcal{M} \cong T_Q\mathcal{M} \times T_Q\mathcal{M} \longrightarrow T\mathcal{M}$ be the tangent of the map $\mathcal{R}_Q$. Let us fix the first argument of $\mathcal{R}'_Q$ so that, for any $U, V \in \mathcal{M}$, the tangent map $\mathcal{R}'_Q\big|_U : T_Q\mathcal{M} \to T_{\mathcal{R}_Q(U)}\mathcal{M}$ is defined as $\mathcal{R}'_Q\big|_U(V) = \mathcal{R}'_Q(U, V)$. Assume that the local rigidity condition $\mathcal{R}'_Q\big|_0 = \mathrm{Id}_{T_Q\mathcal{M}}$ 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 (5.16) in the tangent space of $\mathcal{M}$ at Q. Indeed, using (5.1) we can derive an evolution equation for $V(t)$ as

$$\dot{U}(t) = \mathcal{R}'_Q\big|_{V(t)}(\dot{V}(t)) = \mathcal{F}\big(\mathcal{R}_Q(V(t))\big).$$

By the continuity of V and the local rigidity condition, the map $\mathcal{R}'_Q\big|_{V(t)}$ is invertible for sufficiently small t (i.e., $V(t)$ sufficiently close to $0 \in T_Q\mathcal{M}$) and hence

$$\dot{V}(t) = \left(\mathcal{R}'_Q\big|_{V(t)}\right)^{-1} \mathcal{F}\big(\mathcal{R}_Q(V(t))\big). \tag{5.17}$$

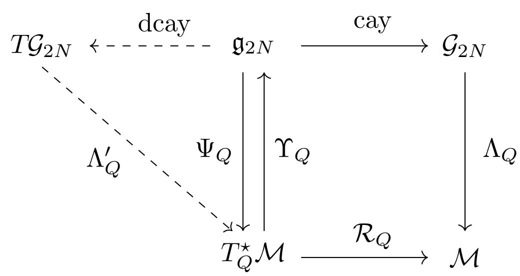
Since the initial condition is $U(t_0) = Q$ it holds $V(t_0) = 0 \in T_Q\mathcal{M}$.

This strategy allows to solve the ODE (5.17) 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 (5.16). 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 (5.17) 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 [8] for the solution of differential equations on the Stiefel manifold. Observe that the velocity field (5.2) describing the flow of the reduced basis on the manifold $\mathcal{M}$ belongs to the linear subspace of $T\mathcal{M}$ defined as $T^\star\mathcal{M} := T\mathcal{M} \cap \{M \in \mathbb{R}^{2N\times 2n} : MJ_{2n} = J_{2N}M\}$. We therefore construct a tangent method restricted to this subspace. In more details, for any $Q \in \mathcal{M}$, we construct a retraction $\mathcal{R}_Q$ as composition of three applications: a linear map Υ_Q from the space $T^\star_Q\mathcal{M}$ 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 $\Lambda : \mathcal{G}_{2N} \times \mathcal{M} \to \mathcal{M}$,

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

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 $\Upsilon_Q : T^\star_Q\mathcal{M} \subset T_Q\mathcal{M} \to \mathfrak{g}_{2N}$ such that $\Psi_Q \circ \Upsilon_Q = \mathrm{Id}_{T^\star_Q\mathcal{M}}$ where $\Psi_Q = \Lambda'_Q\big|_e \circ \mathrm{dcay}_0$.

To build the retraction $\mathcal{R}_Q$, let us characterize the tangent space to the manifold $\mathcal{M}$ at $Q \in \mathcal{M}$. Analogously to (4.6), temporal integration of the constraints on $\mathcal{M}$ yields

$$T_Q\mathcal{M} = \{V \in \mathbb{R}^{2N\times 2n} :\ Q^\top V \in \mathfrak{g}_{2n}\}, \quad \text{and} \quad T^\star_Q\mathcal{M} = \{V \in \mathbb{R}^{2N\times 2n} :\ Q^\top V \in \mathfrak{so}(2n),\ VJ_{2n} = J_{2N}V\}. \tag{5.18}$$

The tangent space of $\mathcal{M}$ can be equivalently characterized as follows.

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

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

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

$$\begin{aligned} Q^\top V &= Q^\top(\Theta Q^\top - Q\Theta^\top)Q = Q^\top\Theta - \Theta^\top Q = -Q^\top(Q\Theta^\top - \Theta Q^\top)Q = -V^\top Q. \\ Q^\top VJ_{2n} &= (Q^\top\Theta - \Theta^\top Q)J_{2n} = -J_{2n}(\Theta^\top Q - Q^\top\Theta) = -J_{2n}V^\top Q. \end{aligned}$$

($\Longrightarrow$) 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

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

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

$$\Theta - Q\Theta^\top Q = V + QS - Q\frac{Q^\top V}{2} - QV^\top Q - Q\left(S^\top - \frac{V^\top Q}{2}\right) = V - Q\frac{Q^\top V}{2} - Q\frac{V^\top Q}{2} = V. \tag{5.19}$$

$\square$

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

$$T_Q\mathcal{M} = \{V \in \mathbb{R}^{2N\times 2n} :\ V = (\Theta_Q^S(V)Q^\top - Q\Theta_Q^S(V)^\top)Q,$$
$$\text{with } \Theta_Q^S(V) := V + Q\left(S - \frac{Q^\top V}{2}\right),\ \text{for } S \in \mathrm{Sym}(2n)\cap \mathfrak{sp}(2n)\}.$$

This suggests that the linear map Υ_Q can be defined as

$$\begin{array}{rccl} \Upsilon_Q : & T_Q^\star\mathcal{M} & \longrightarrow & \mathfrak{g}_{2N}, \\ & V & \longmapsto & \Theta_Q^S(V)Q^\top - Q\Theta_Q^S(V)^\top. \end{array} \tag{5.20}$$

Note that $\Upsilon_Q(V) = VQ^\top - QV^\top + QV^\top QQ^\top$ and hence $\Upsilon_Q(V) \in \mathfrak{g}_{2N}$ in view of the characterization of $T_Q^\star\mathcal{M}$ in (5.18). Indeed, since $\Lambda'_Q\big|_e(G) = GQ$ and $\mathrm{dcay}_0 = I$, it holds $(\Psi_Q \circ \Upsilon_Q)(V) = \Upsilon_Q(V)Q = V$ for any $V \in T_Q^\star\mathcal{M}$. This stems from the definition of Υ_Q in (5.20) since

$$\begin{aligned} \Psi_Q(\Upsilon_Q(V)) &= \big(\Lambda_Q\big|_e \circ \mathrm{dcay}_0 \circ \Upsilon_Q\big)(V) = \Lambda_Q\big|_e(\Upsilon_Q(V)) \\ &= \Upsilon_Q(V)Q = \big(\Theta_Q^S(V)Q^\top - Q\Theta_Q^S(V)^\top\big)\,Q = V, \end{aligned}$$

where the last equality follows by (5.19). Note that $\Psi_Q = \Lambda'_Q\big|_e \circ \mathrm{dcay}_0$ is not injective as $\Upsilon_Q(T_Q^\star\mathcal{M})$ is a proper subspace of $\mathfrak{g}_{2N}$.

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*

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

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

$$\mathcal{R}_Q(V) = \mathrm{cay}(\Theta_Q^S(V)Q^\top - Q\Theta_Q^S(V)^\top)Q, \tag{5.21}$$

is a retraction.

Proof. We follow [8, Proposition 2.2]. Let $V = 0 \in T_Q\mathcal{M}$, then $\Theta_Q^S(0) = QS$ and then, using the fact that $S \in \mathrm{Sym}(2n)$ and $\mathrm{cay}(0) = I_{2N}$, it holds $\mathcal{R}_Q(0) = \mathrm{cay}\big(Q(S - S^\top)Q^\top\big)Q = \mathrm{cay}(0)Q = Q$.

Let Υ_Q be defined as in (5.20). Since, by construction Υ_Q admits left inverse it is injective and then $\Upsilon_Q(V) = 0$ if and only if $V = 0 \in T_Q^\star\mathcal{M}$. Then, $\mathcal{R}_Q(V) = Q$ if and only if $\mathrm{cay}(\Upsilon_Q(V)) = I_{2N}$, which implies $V = 0 \in T_Q^\star\mathcal{M}$. Moreover, since $\mathcal{R}_Q = \Lambda_Q \circ \mathrm{cay} \circ \Upsilon_Q$, the definition of group action and the linearity of Υ result in $\mathcal{R}'_Q\big|_0 = \Psi_Q \circ \Upsilon_Q = \mathrm{Id}_{T_Q^\star\mathcal{M}}$. It can be easily verified that $\mathcal{R}_Q(V) \in \mathcal{M}$ for any $V \in T_Q^\star\mathcal{M}$. ∎

Note that the matrix $S \in \mathrm{Sym}(2n)\cap\mathfrak{sp}(2n)$ in the definition of the retraction (5.21) is of the form

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

Its choice affects the numerical performances of the algorithm for the computation of the retraction and its inverse tangent map, as pointed out in [8, Section 3].

In the following Subsections we propose a temporal discretization of (5.17) 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 (5.17), reads

$$\dot{V}(t) = \left(\mathcal{R}'_{U_m}\Big|_{V(t)}\right)^{-1} \mathcal{F}\big(\mathcal{R}_{U_m}(V(t))\big) =: f_m(V(t)).$$

Let $(b_i, a_{i,j})$ for $i = 1, \dots, s$ and $j = 1, \dots, i-1$ be the coefficients of the Butcher tableau describing the s-stage explicit Runge–Kutta method. Then the numerical approximation of (5.17)-(5.16) reads: Given $U_0 := Q \in \mathcal{M}$ and $V_0 = 0 \in T_Q\mathcal{M}$, for $m = 0, 1, \dots$

$$\begin{aligned} V_{m+1} = \quad & \Delta t \sum_{i=1}^{s} b_i A_m^i, \\ & A_m^1 = \mathcal{F}(U_m), \\ & A_m^i = f_m\Big(\Delta t \sum_{j=1}^{i-1} a_{i,j} A_m^j\Big), \qquad\qquad i = 2, \dots, s, \\ U_{m+1} = \quad & \mathcal{R}_{U_m}(V_{m+1}). \end{aligned} \tag{5.22}$$

Other than the evaluation of the velocity field $\mathcal{F}$ at $\mathcal{R}_{U_m}(V)$, the crucial points of algorithm (5.22) 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 (5.22) has an overall arithmetic complexity of order $O(Nn^2s) + C_{\mathcal{F}}s$, 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 $\mathcal{R}_Q$ (5.21) at the matrix $V \in \mathbb{R}^{2N\times 2n}$ requires $O(N^2n)$ for the multiplication between $\mathrm{cay}(\Upsilon_Q(V))$ and Q, plus the computational cost to evaluate the Cayley transform at $\Upsilon_Q(V) \in \mathbb{R}^{2N\times 2N}$. However, for any $V \in T_Q^\star\mathcal{M}$, 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

$$\alpha := \big[\,\Theta_Q^S(V)\,|\,{-Q}\,\big] \in \mathbb{R}^{2N\times 4n}, \qquad \beta := \big[\,Q\,|\,\Theta_Q^S(V)\,\big] \in \mathbb{R}^{2N\times 4n}. \tag{5.23}$$

We can revert to the results of Proposition 5.2 (with $k = 4n$) so that the retraction (5.21) can be computed as

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

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

Computation of the inverse tangent map of the retraction. Let $Q \in \mathcal{M}$ and $V \in T_Q^\star\mathcal{M}$. Using the definition of retraction (5.21) we have

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

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

$$\mathcal{R}'_Q = \Lambda'_Q \circ \mathrm{cay}' \circ \Upsilon'_Q : TT_Q^\star\mathcal{M} \longrightarrow T\mathfrak{g}_{2N} \cong \mathfrak{g}_{2N} \longrightarrow T\mathcal{G}_{2N} \longrightarrow T_Q\mathcal{M}.$$

Fixing the fiber on $TT_Q^\star\mathcal{M}$ corresponding to $V \in T_Q^\star\mathcal{M}$ results in

$$\begin{aligned} \mathcal{R}'_Q\Big|_V(\widetilde{V}) = \mathcal{R}'_Q(V,\widetilde{V}) &= \Lambda'_Q\Big|_{\mathrm{cay}(\Upsilon_Q(V))} \circ \mathrm{dcay}_{\Upsilon_Q(V)}(\Upsilon_Q(\widetilde{V})) \\ &= \mathrm{dcay}_{\Upsilon_Q(V)}(\Upsilon_Q(\widetilde{V}))\,\mathrm{cay}(\Upsilon_Q(V))Q = \mathrm{dcay}_{\Upsilon_Q(V)}(\Upsilon_Q(\widetilde{V}))\,\mathcal{R}_Q(V), \end{aligned}$$

where we have used the linearity of the map Υ_Q.

Assume we know $W \in T^{\star}_{\mathcal{R}_Q(V)}\mathcal{M}$. We want to compute $\widetilde{V} \in T^{\star}_{Q}\mathcal{M}$ such that

$$\mathcal{R}'_Q\big|_V(\widetilde{V}) = \mathrm{dcay}_{\Upsilon_Q(V)}(\Upsilon_Q(\widetilde{V}))\,\mathcal{R}_Q(V) = W. \tag{5.24}$$

It is possible to solve problem (5.24) with arithmetic complexity $O(Nn^2)$ by proceeding as in [8, Section 3.2.1]. Since, for our algorithm, the result of [8] can be extended to the case of arbitrary matrix $S \in \mathrm{Sym}(2n) \cap \mathfrak{sp}(2n)$ in (5.21), we report the more general derivation in Appendix A.

Note that for $S = 0$ and explicit Euler scheme the two numerical integrators (5.14) and (5.22) are equivalent.

5.3.2 Convergence estimates for the tangent method

Since the retraction and its inverse tangent map in (5.22) 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 (5.17) in the tangent space.

Proposition 5.7. *The retraction map $\mathcal{R} : T\mathcal{M} \to \mathcal{M}$ defined in (5.21) is locally Lipschitz continuous in the Frobenius $\|\cdot\|$-norm, namely for any $Q \in \mathcal{M}$, $\mathcal{R}_Q : T^{\star}_Q\mathcal{M} \to \mathcal{M}$ satisfies*

$$\|\mathcal{R}_Q(V) - \mathcal{R}_Q(W)\| \leq 3\|V - W\|, \qquad \forall\, V,\, W \in T^{\star}_Q\mathcal{M}.$$

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 $\overline{\Upsilon}_Q(\cdot) := \Upsilon_Q(\cdot)/2$,

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

Since Υ_Q is skew-symmetric $\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}$ is normal. Then $\|\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}\|_2 = \rho\big[\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}\big]$ and $\rho\big[\big(I_{2N} - \overline{\Upsilon}_Q(V)\big)^{-1}\big] \leq 1$. Hence, since Q and Y are (semi-)orthogonal matrices, it holds

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

Using the definition of Υ_Q from (5.20) results in

$$\|\Upsilon_Q(V) - \Upsilon_Q(W)\| = \|(V - W)Q^\top - Q(V - W) + Q(V^\top - W^\top)QQ^\top\| \leq 3\|V - W\|.$$

□

It follows that the solution of (5.22) 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 (5.21). Let $U(t^m) = \mathcal{R}_Q(V(t^m))$, where $V(t^m)$ is the exact solution of (5.17) at a given time t^m and let $U_m = \mathcal{R}_Q(V_m)$, where V_m is the numerical solution of (5.17) at time t^m obtained with algorithm (5.22). 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 Concluding remarks and future work

Nonlinear dynamical reduced basis methods for parameterized finite-dimensional Hamiltonian systems have been developed. These 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, restarting procedures of the Cayley RK algorithm (5.14), and the investigation of the robustness of the temporal integrator under over-approximation of the full model solution [18].

Our dynamical reduced basis method engenders a smooth approximation of the solution of the full Hamiltonian model. Analogous to dynamical low-rank approximations of matrices, we rely on the assumption that the projection error of the full model solution onto the evolving reduced space (of fixed rank) remains small at all times. However, dynamical modes that have been neglected in the reduction might become relevant over time (the problem of crossing of singular values [14]). Restarting algorithms and dynamical approximations with adaptive rank might be envisioned in this case.

Finally, 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.

Acknowledgment. The author gratefully acknowledges many fruitful discussions with Jan S. Hesthaven, and would like to thank Elena Celledoni for pointing out references [7, 8]. This work was partially supported by AFOSR under grant FA9550-17-1-9241.

References

[1] R. Abraham and J. E. Marsden. *Foundations of mechanics. Second edition.* Addison-Wesley Publishing Company, Inc., Redwood City, CA., 1987.

[2] B. M. Afkham and J. S. Hesthaven. "Structure preserving model reduction of parametric Hamiltonian systems." *SIAM J. Sci. Comput.* 39.6 (2017), A2616–A2644. URL: https://doi.org/10.1137/17M1111991.

[3] P. Buchfink, A. Bhatt, and B. Haasdonk. "Symplectic model order reduction with non-orthonormal bases." *Math. Comput. Appl.* 24.2 (2019), Paper No. 43, 26.

[4] A. Cannas da Silva. *Lectures on symplectic geometry.* Vol. 1764. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001. URL: https://doi.org/10.1007/978-3-540-45330-7.

[5] K. Carlberg, R. Tuminaro, and P. Boggs. "Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics." *SIAM J. Sci. Comput.* 37.2 (2015), B153–B184. URL: https://doi.org/10.1137/140959602.

[6] E. Celledoni and A. Iserles. "Approximating the exponential from a Lie algebra to a Lie group." *Math. Comp.* 69.232 (2000), pp. 1457–1480. URL: https://doi.org/10.1090/S0025-5718-00-01223-0.

[7] E. Celledoni and B. Owren. "A class of intrinsic schemes for orthogonal integration." *SIAM J. Numer. Anal.* 40.6 (2002), pp. 2069–2084. URL: https://doi.org/10.1137/S0036142901385143.

[8] E. Celledoni and B. Owren. "On the implementation of Lie group methods on the Stiefel manifold." *Numer. Algorithms* 32.2-4 (2003), pp. 163–183. URL: https://doi.org/10.1023/A:1024079724094.

[9] F. Diele, L. Lopez, and R. Peluso. "The Cayley transform in the numerical solution of unitary differential systems." *Adv. Comput. Math.* 8.4 (1998), pp. 317–334. URL: https://doi.org/10.1023/A:1018908700358.

[10] F. Feppon and P. F. J. Lermusiaux. "A geometric approach to dynamical model order reduction." *SIAM J. Matrix Anal. Appl.* 39.1 (2018), pp. 510–538. URL: https://doi.org/10.1137/16M1095202.

[11] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration.* Second. Vol. 31. Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2006.

[12] J. S. Hesthaven and C. Pagliantini. *Structure-Preserving Reduced Basis Methods for Hamiltonian Systems with a Nonlinear Poisson Structure.* Tech. rep. EPFL, Switzerland, 2018. URL: http://infoscience.epfl.ch/record/256097.

[13] A. Iserles. "On Cayley-transform methods for the discretization of Lie-group equations." *Found. Comput. Math.* 1.2 (2001), pp. 129–160. URL: https://doi.org/10.1007/s102080010003.

[14] O. Koch and C. Lubich. "Dynamical low-rank approximation." *SIAM J. Matrix Anal. Appl.* 29.2 (2007), pp. 434–454. URL: https://doi.org/10.1137/050639703.

[15] S. Lall, P. Krysl, and J. E. Marsden. "Structure-preserving model reduction for mechanical systems." *Phys. D* 184.1-4 (2003), pp. 304–318. URL: https://doi.org/10.1016/S0167-2789(03)00227-6.

[16] L. Lopez and T. Politi. "Applications of the Cayley approach in the numerical solution of matrix differential systems on quadratic groups." *Appl. Numer. Math.* 36.1 (2001), pp. 35–55. URL: https://doi.org/10.1016/S0168-9274(99)00049-5.

[17] C. Lubich. *From quantum to classical molecular dynamics: reduced models and numerical analysis.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008. URL: https://doi.org/10.4171/067.

[18] C. Lubich and I. V. Oseledets. "A projector-splitting integrator for dynamical low-rank approximation." *BIT* 54.1 (2014), pp. 171–188. URL: https://doi.org/10.1007/s10543-013-0454-0.

[19] H. Munthe-Kaas. "High order Runge-Kutta methods on manifolds." *Proceedings of the NSF/CBMS Regional Conference on Numerical Analysis of Hamiltonian Differential Equations (Golden, CO, 1997)*. Vol. 29. 1. 1999, pp. 115–127. URL: https://doi.org/10.1016/S0168-9274(98)00030-0.

[20] E. Musharbash and F. Nobile. *Symplectic Dynamical Low Rank approximation of wave equations with random parameters*. Tech. rep. 18.2017. EPFL, Switzerland, 2017.

[21] E. Musharbash, F. Nobile, and T. Zhou. "Error analysis of the dynamically orthogonal approximation of time dependent random PDEs." *SIAM J. Sci. Comput.* 37.2 (2015), A776–A810. URL: https://doi.org/10.1137/140967787.

[22] M. Ohlberger and S. Rave. "Reduced Basis Methods: Success, Limitations and Future Challenges." *Proceedings of the Conference Algoritmy* (2016), pp. 1–12. URL: http://www.iam.fmph.uniba.sk/amuc/ojs/index.php/algoritmy/article/view/389.

[23] L. Peng and K. Mohseni. "Symplectic model reduction of Hamiltonian systems." *SIAM J. Sci. Comput.* 38.1 (2016), A1–A27. URL: https://doi.org/10.1137/140978922.

[24] T. P. Sapsis and P. F. J. Lermusiaux. "Dynamically orthogonal field equations for continuous stochastic dynamical systems." *Phys. D* 238.23-24 (2009), pp. 2347–2360. URL: https://doi.org/10.1016/j.physd.2009.09.017.

[25] M. A. Woodbury. *Inverting modified matrices*. Statistical Research Group, Memo. Rep. no. 42. Princeton University, Princeton, N. J., 1950.

A Efficient computation of the inverse tangent map

We propose an algorithm to solve (5.24) with a computational cost of order $O(Nn^2)$. We proceed exactly as in [8, Section 3.2.1] with the only difference that we consider any arbitrary $S \in \mathrm{Sym}(2n) \cap \mathfrak{sp}(2n)$.

Using the definition of the derivative of the Cayley transform (5.4) we can recast (5.24) as

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

Moreover, using the definition of $\mathcal{R}_Q(V)$ in (5.21) results in

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

Therefore, substituting in (A.1) and using the definition of Υ_Q from (5.20) gives

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

We proceed by solving problem (A.2) for $\widetilde{\Theta} := \Theta^S_Q(\widetilde{V}) \in T_Q\,\mathrm{Sp}(2n, \mathbb{R}^{2N})$ and then, in view of (5.19), we recover $\widetilde{V} \in T^\star_Q\mathcal{M}$ 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 $\Theta^S_Q(\widetilde{V}) = QT_1(\widetilde{V}) + T_2$, where

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

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 (5.23), 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 $T_1(\widetilde{V}) = Q^\top\Theta^S_Q(\widetilde{V}) + QT_2$. Using the fact that $Q^\top\Theta^S_Q(\widetilde{V}) + \Theta^S_Q(\widetilde{V})^\top Q = 2S$, the symmetric part of T_1 reads $T_1 + T_1^\top = 2S - Q^\top T_2 + T_2^\top Q$. Moreover,

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

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,

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

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