Geometric optimization for structure-preserving model reduction of Hamiltonian systems

Bendokat, Thomas; Zimmermann, Ralf

Published in:
10th Vienna International Conference on Mathematical Modelling

DOI:
10.1016/j.ifacol.2022.09.137

Publication date:
2022

Document version:
Accepted manuscript

Citation for pulished version (APA):
Bendokat, T., & Zimmermann, R. (2022). Geometric optimization for structure-preserving model reduction of Hamiltonian systems. In *10th Vienna International Conference on Mathematical Modelling: MATHMOD 2022* (20 ed., Vol. 55, pp. 457-462). Elsevier. https://doi.org/10.1016/j.ifacol.2022.09.137

Geometric Optimization for Structure-Preserving Model Reduction of Hamiltonian Systems

Thomas Bendokat [*,**] **Ralf Zimmermann** [**]

[*] *Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany, (e-mail: bendokat@mpi-magdeburg.mpg.de)*
[**] *Department of Mathematics and Computer Science, University of Southern Denmark (SDU), Odense, Denmark (e-mail: zimmermann@imada.sdu.dk).*

Abstract:
Classical model reduction methods disregard the special symplectic structure associated with Hamiltonian systems. A key challenge in projection-based approaches is to construct a symplectic basis that captures the essential system information. This necessitates the computation of a so-called proper symplectic decomposition (PSD) of a given sample data set. The PSD problem allows for a canonical formulation as an optimization problem on the symplectic Stiefel manifold. However, as with their Euclidean counterparts, symplectic projectors only depend on the underlying symplectic subspaces and not on the selected symplectic bases. This motivates to tackle the PSD problem as a Riemannian optimization problem on the symplectic Grassmann manifold, i.e., the matrix manifold of symplectic projectors. Initial investigations on this manifold feature in a recent preprint of the authors. In this work, we investigate the feasibility and performance of this approach on two academic numerical examples. More precisely, we calculate an optimized PSD for snapshot matrices that stem from solving the one-dimensional linear wave equation and the one-dimensional nonlinear Schrödinger equation.

Keywords: Model Reduction, Hamiltonian Systems, Proper Symplectic Decomposition, Symplectic Grassmann Manifold, Symplectic Stiefel Manifold, Riemannian Optimization

1. INTRODUCTION

As a rule, simulating dynamical processes in real-life scenarios entails a tremendous amount of computational time and costs. *Model order reduction (MOR)* is concerned with developing techniques to emulate such large-scale systems in a fraction of the original computation time up to a satisfactory error. A key challenge in MOR is to construct a reduced basis that captures the essential system information. For snapshot-based MOR, this requires a data-sampling stage as an upfront investment. Once the reduced basis is at hand, the large-scale system is projected onto the associated subspace in projection-based MOR. The resulting *reduced-order model (ROM)* operates exclusively on the basis coefficients—usually some $p \in [10^2, 10^3]$—rather than on the original system dimension N, where $N \in [10^6, 10^8]$ is not rare in realistic applications. The basis construction paradigm consists of two generic steps.

- Data sampling: collect sample solutions, store data in a *snapshot matrix* $\mathbb{S}$.
- Basis construction: find basis of low rank $p \ll N$ that best represents the data in $\mathbb{S}$ with p coefficients.

The predictive power of the reduced basis is of paramount importance for the ROM's performance.

In this work, we consider aspects of MOR for Hamiltonian systems. Examples of Hamiltonian systems include applications in quantum mechanics (the Schrödinger equation), wave equations, the class of shallow water equations, celestial dynamics, etc. For the background theory, see Arnol'd (1997); Arnol'd and Givental' (2001). Yet, classical MOR methods provide bases that do not respect the invariants and the symplectic structure inherent to Hamiltonian systems. This results in ill-suited, often unstable ROMs. To obtain a basis that yields a structure-preserving ROM, Peng and Mohseni (2016) proposed to compute a *proper symplectic decomposition (PSD)* of the snapshot matrix $\mathbb{S}$ and to replace orthogonal projections with symplectic ones. The PSD problem is formalized via

$$\min_{U \in \mathbb{R}^{2N \times 2p}} \|\mathbb{S} - UU^+\mathbb{S}\|_F^2 \quad \text{s. t.} \quad U^T \mathbb{J}_{2N} U = \mathbb{J}_{2p}, \tag{1}$$

where $U^+ = \mathbb{J}_{2p}^T U^T \mathbb{J}_{2N}$ and

$$\mathbb{J}_{2m} = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix},$$

$m \in \{N, p\}$. Note that the problem only depends on the symplectic projector UU^+ and not the representative U. Furthermore, note the formal similarities with the archetype problem of dimension reduction

$$\min_{U \in \mathbb{R}^{N \times p}} \|\mathbb{S} - UU^T\mathbb{S}\|_F^2 \quad \text{s. t.} \quad U^T I_n U = I_p. \tag{2}$$

The solution of (2) is the singular value decomposition (SVD) of $\mathbb{S}$ truncated to the p dominant left singular vectors and is fundamental to the methods of proper orthogonal decomposition (POD), principal component analysis and Karhunen-Loève expansion. In contrast, there is no known closed-form solution to the PSD problem (1).

We briefly relate the state of the art. For more details, see Afkham and Hesthaven (2017); Buchfink et al. (2020); Musharbash et al. (2020); Peng and Mohseni (2016). So far, these are the probed methods:

- **(a)** *Cotangent lift*, (Peng and Mohseni (2016)): This method is SVD-based and yields the optimal symplectic projection basis under all candidates of block-diagonal form $\begin{pmatrix} \Phi & 0 \\ 0 & \Phi \end{pmatrix}$.
- **(b)** *Complex SVD*, (Peng and Mohseni (2016); Buchfink et al. (2019)): This method is SVD-based and yields the optimal symplectic projection basis under all candidates of block form $\begin{pmatrix} \Phi & -\Psi \\ \Psi & \Phi \end{pmatrix}$; however not for representing $\mathbb{S}$ but rather $(\mathbb{S}, -\mathbb{J}_{2N}\mathbb{S})$.
- **(c)** *Symplectic SVD-like decomposition*, (Buchfink et al. (2019)): Here, the SVD is replaced by a decomposition for which one of the orthogonal factors is instead symplectic, see Xu (2003) for details.
- **(d)** *Greedy basis generation*, (Afkham and Hesthaven (2017); Buchfink et al. (2020)). This approach is based on symplectic extension of an existing basis according to an error criterion, e.g., conservation of the Hamiltonian. The extension methods are based on an iterative, symplectic Gram-Schmidt process, or on complex SVD or SVD-like decomposition of a residual.
- **(e)** *Nonlinear programming*, (Peng and Mohseni (2016)): Here, (1) is tackled as a constrained optimization problem.

The methods (a), (b), (d) are, by design, in general not able to produce the global optimum for (1). The approach (e) has this potential but tends to be expensive. This has already been recognized in Peng and Mohseni (2016), who propose a restricted optimization that is bound to linear transformations of an initial suboptimal PSD. The approach (c) appears as a promising starting point but comes with the extra challenge that off-the-shelf methods for computing named symplectic SVD-like decomposition are lacking. For the experiments in Section 4, we utilize the MATLAB implementation '`svd_like_decomposition.m`' provided by Buchfink et al. (2019).

Original contribution. The PSD problem (1) can be considered as an unconstrained optimization problem on the *symplectic Stiefel manifold*,

$$\mathrm{SpSt}(2N, 2p) := \{U \in \mathbb{R}^{2N\times 2p} \mid U^T\mathbb{J}_{2N}U = \mathbb{J}_{2p}\}.$$

However, the objective function (1) actually does not depend on the symplectic Stiefel frame $U \in \mathrm{SpSt}(2N, 2p)$, but on the associated symplectic projector UU^+. The set

$$\begin{aligned}&\mathrm{SpGr}(2N, 2p) :=\\ &\{P \in \mathbb{R}^{2N\times 2N} \mid P^2 = P,\ \mathrm{rank}(P) = 2p,\ P^+ = P\}\end{aligned} \tag{3}$$

of symplectic projectors also constitutes a matrix manifold, which we refer to as the *real symplectic Grassmann manifold*. In the recent preprint Bendokat and Zimmermann (2021), we established the symplectic Grassmann manifold as a quotient space of $\mathrm{SpSt}(2N, 2p)$ and proposed Riemannian and Pseudo-Riemannian metrics. Moreover, we derived formulas for the associated geodesic lines on $\mathrm{SpGr}(2N, 2p)$ and introduced efficient retractions. In the work at hand, we present preliminary numerical experiments on tackling (1) as a Riemannian optimization problem on $\mathrm{SpGr}(2N, 2p)$ by applying Riemannian optimization to two examples of Hamiltonian systems and comparing the results to the first three methods mentioned above.

2. HAMILTONIAN SYSTEMS AND SYMPLECTIC STRUCTURES

2.1 Basics

The canonical symplectic vector space is $\mathbb{V} = (\mathbb{R}^{2N}, \omega_0)$, with symplectic form $\omega_0(v, w) = v^T\mathbb{J}_{2N}w$. The symplectic inverse of a matrix $A \in \mathbb{R}^{2N\times 2p}$ is

$$A^+ := \mathbb{J}_{2p}^T A^T \mathbb{J}_{2n}.$$

For symplectic Stiefel matrices $U \in \mathrm{SpSt}(2N, 2p)$, it holds $U^+U = I_{2p}$. Hamiltonian matrices are characterized by $A^+ = -A$. For $U \in \mathrm{SpSt}(2N, 2p)$, $P := UU^+$ represents the symplectic projection onto the $2p$-dimensional symplectic subspace $\mathrm{span}(U) \subset \mathbb{V}$.

An autonomous Hamiltonian system associated with a smooth Hamiltonian function $H : \mathbb{V} \to \mathbb{R}$ has the form

$$\begin{aligned}\dot z(t) = \mathbb{J}_{2N}\nabla_z H(z(t)), \quad z(t) = \begin{pmatrix} q(t) \\ p(t) \end{pmatrix} \in \mathbb{R}^{2N} \\ \Leftrightarrow \dot q(t) = \nabla_p H(q(t), p(t)),\ \ \dot p(t) = -\nabla_q H(q(t), p(t)),\end{aligned} \tag{4}$$

with initial state $(q(t_0)^T, p(t_0)^T)^T = z(t_0) = z_0$.

Hamiltonian systems preserve the Hamiltonian, identified with the total energy, see (Hairer et al., 2006, Section IV, Example 1.2). Hence, numerical schemes for advancing (4) in time that share this feature have been developed. Let $h = \Delta t$ be a time step size. We mention the symplectic Euler method

$$\begin{aligned}q_{n+1} &= q_n + h\nabla_p H(q_n, p_{n+1}),\\ p_{n+1} &= p_n - h\nabla_q H(q_n, p_{n+1})\end{aligned}$$

and the Störmer-Verlet scheme

$$\begin{aligned}q_{n+\frac12} &= q_n + \frac{h}{2}\nabla_p H(q_{n+\frac12}, p_n),\\ p_{n+1} &= p_n - \frac{h}{2}\left(\nabla_q H(q_{n+\frac12}, p_n) + \nabla_q H(q_{n+\frac12}, p_{n+1})\right),\\ q_{n+1} &= q_{n+\frac12} + \frac{h}{2}\nabla_p H(q_{n+\frac12}, p_{n+1}),\end{aligned} \tag{5}$$

and refer to (Hairer et al., 2006, Section VI.3) for the details and more options.

2.2 Model reduction for Hamiltonian systems

As with classical MOR, a formal reduction in dimension of a given Hamiltonian system can be obtained via projection. Let $P \in \mathrm{SpGr}(2N, 2p)$ be a symplectic projector as in (3). According to (Bendokat and Zimmermann, 2021, Prop. 4.1), there is $U \in \mathrm{SpSt}(2N, 2p)$ such that $P = UU^+$

and $P(\mathbb{R}^{2N}) = \mathrm{ran}(U)$. Let $z(t)$ be a solution of (4). Then, the projected trajectory $Pz(t) = UU^+z(t) = Uz_p(t)$ is uniquely defined by the coefficient function $z_p(t) = U^+z(t) \in \mathbb{R}^{2p}$. If $z(t)$ were perfectly contained in $\mathrm{ran}(U)$, then $z(t) = Pz(t) = Uz_p(t)$ and (4) yields

$$\begin{aligned}\dot{z}_p(t) &= U^+\mathbb{J}_{2N}\nabla_z H(Uz_p(t)) = -\mathbb{J}_{2p}^T U^T \nabla_z H(Uz_p(t)) \\ &= \mathbb{J}_{2p}\nabla_{z_p} H(Uz_p(t)). \end{aligned} \tag{6}$$

In this way, formally, the order-$2N$ system (4) has been replaced by an order-$2p$ system for the function $z_p(t)$ that defines the coefficients for the projected trajectory. For linear non-parametric Hamiltonian systems, the operators associated with the reduced system (6) can be precomputed. Nonlinear systems can be treated, e.g., with a symplectic discrete empirical interpolation approach, see Peng and Mohseni (2016) for details.

Due to space limitations, this contribution will focus entirely on the process of finding the projector $P = UU^+$. With such a projector, a reduced order model can be created using the techniques in Peng and Mohseni (2016); Afkham and Hesthaven (2017). In order to find the projector, we advance the full system (4) in time and collect snapshot solutions $\mathbb{S} = (z(t_{s_1}), \ldots, z(t_{s_m}))$ at selected time instants $t_{s_1}, \ldots, t_{s_m}$. Then, we conduct a Riemannian optimization process for the problem

$$\arg\min_P \|\mathbb{S} - P\mathbb{S}\|_F^2. \tag{7}$$

3. OPTIMIZATION ON THE SYMPLECTIC GRASSMANN MANIFOLD

Optimization on the symplectic Grassmann manifold $\mathrm{SpGr}(2N, 2p)$ can be efficiently executed by a lift to the symplectic Stiefel manifold $\mathrm{SpSt}(2N, 2p)$. Details on the connection between these two manifolds and on Riemannian optimization aspects can be found in Bendokat and Zimmermann (2021).

The tangent space at a symplectic projector $UU^+ \in \mathrm{SpGr}(2N, 2p)$ can be parameterized by choosing an associated representative $U \in \mathrm{SpSt}(2N, 2p)$ and setting

$$T_{UU^+}\mathrm{SpGr}(2N, 2p) = \{H \in \mathbb{R}^{2N\times 2p} \mid U^+H = 0\}.$$

For a chosen representative U, a Riemannian metric is defined point-wise by $g_{UU^+}\colon T_{UU^+}\mathrm{SpGr}(2N, 2p) \times T_{UU^+}\mathrm{SpGr}(2N, 2p) \to \mathbb{R}$,

$$\begin{aligned} g_{UU^+}\ &(H_1, H_2) = \\ &\mathrm{tr}(U^TU(H_2^TH_1)^+ - (U^TH_1)^+H_2^TU). \end{aligned} \tag{8}$$

Note that this Riemannian metric does not depend on the representative U, as the change of symplectic basis for U and H_i consist of the post-multiplication of the same symplectic matrix, which cancel each other out. A smooth function on $\mathrm{SpGr}(2N, 2p)$ can be defined as a smooth function $f\colon \mathrm{SpSt}(2N, 2p) \to \mathbb{R}$ fulfilling $f(U) = f(UN)$ for all symplectic basis changes N, i.e., all $N \in \mathbb{R}^{2p\times 2p}$ with $N^+N = I_{2p}$. Let ∇f_U be the Euclidean gradient of f at U. The Riemannian gradient $\mathrm{grad}_f(U) \in T_{UU^+}\mathrm{SpGr}(2N, 2p)$ of f with respect to (8) is given by

$$\mathrm{grad}_f(U) = (I_{2n} - UU^+)\mathbb{J}_{2N}^T\nabla f_U\mathbb{J}_{2p}.$$

Specifically, the Euclidean gradient of the objective function of (7), i.e., of $f\colon \mathrm{SpSt}(2N, 2p) \to \mathbb{R}$ with

$$f(U) = \|\mathbb{S} - UU^+\mathbb{S}\|_F^2, \tag{9}$$

is given by

$$\nabla f_U = -2\left(\mathcal{Q} + \mathcal{Q}^+\right)(U^+)^T,$$

where $\mathcal{Q} = (I_{2N} - UU^+)\mathbb{S}\mathbb{S}^T$. With knowledge of the Riemannian gradient, a gradient descent approach can be used to minimize (9). Stepping from a point in a given direction on the symplectic Grassmann manifold can be done via retractions that mediate between tangent space data, i.e., 'directions' and manifold data. The prime example of a retraction is the Riemannian exponential mapping, which corresponds to stepping along geodesic lines. An alternative is to use retractions based on the Cayley map. Explicit formulas for both of the aforementioned approaches are included in Bendokat and Zimmermann (2021). For the numerical experiments features in Section 4, we use the Cayley retraction exclusively, as it is cheaper in terms of the computational cost. Let $U \in \mathrm{SpSt}(2N, 2p)$ represent $UU^+ \in \mathrm{SpGr}(2N, 2p)$ and let $H \in T_{UU^+}\mathrm{SpGr}(2N, 2p)$. Furthermore, let

$$A = (H^TU)^+U^TU - (U^TU)^+H^TU \in \mathbb{R}^{2p\times 2p}$$

and

$$T = (UH^+ - HU^+)^TU - UA \in \mathbb{R}^{2N\times 2p}.$$

Then the Cayley retraction from U in direction H is given by

$$\mathrm{Cay}_U(tH) = -U + (tT + 2U)(\frac{t^2}{4}T^+T - \frac{t}{2}A + I_{2p})^{-1}.$$

For the experiments in the next Section 4, we use (Gao et al., 2021, Algorithm 1) as the line-search gradient descent method, adapted to the symplectic Grassmann manifold, i.e., using the aforementioned Riemannian metric, Riemannian gradient and Cayley retraction.

4. NUMERICAL EXPERIMENTS

We test the approach of computing a PSD via optimizing (7) on $\mathrm{SpGr}(2N, 2p)$ by means of numerical experiments. To this end, we consider the linear wave equation and the nonlinear Schrödinger equation, where we follow the settings of (Peng and Mohseni, 2016, Section 6.2) and (Afkham and Hesthaven, 2017, Section 5.2), respectively.

4.1 Linear wave equation

The one-dimensional linear wave equation with constant wave speed c reads

$$u_{tt}(t, x) = c^2u_{xx}(t, x).$$

We consider the spatial domain $[0, l]$ and impose periodic boundary conditions. Discretizing in space on steps $x_i := i\Delta x, i = 0, \ldots, N-1$, $\Delta x = \frac{l}{N}$ and introducing $q_i(t) = u(t, x_i)$ and $p_i(t) = u_t(t, x_i)$, the Hamiltonian formulation

$$\begin{pmatrix}\dot{q}\\ \dot{p}\end{pmatrix} = \mathbb{J}_{2n}\begin{pmatrix}-c^2D_{xx} & 0\\ 0 & I\end{pmatrix}\begin{pmatrix}q\\ p\end{pmatrix} = \mathbb{J}_{2n}\begin{pmatrix}\nabla_q H(q,p)\\ \nabla_p H(q,p)\end{pmatrix} \tag{10}$$

is obtained, where D_{xx} is the central difference operator corresponding to the second-order spatial derivative. Essentially, this is the tridiagonal matrix $\frac{1}{\Delta x^2}\mathrm{diags}(\mathbf{1}, -\mathbf{2}, \mathbf{1})$, but with the extra entries $D_{xx}[N, 1] = \frac{1}{\Delta x^2} = D_{xx}[1, N]$ that encode the periodic boundary conditions. As an initial condition $q(0) := \left(c(10|x_1 - \frac{1}{2}|), \ldots, c(10|x_N - \frac{1}{2}|)\right)^T$, $p(0) := 0 \in \mathbb{R}^N$ is used, where $c(s)$ is the spline function defined piece wise by $c|_{[0,1]}(s) = 1 - \frac{3}{2}s^2 + \frac{3}{4}s^3$, $c|_{(1,2]}(s) = \frac{1}{4}(2 - s)^3$, $c|_{(2,\infty)}(s) = 0$.

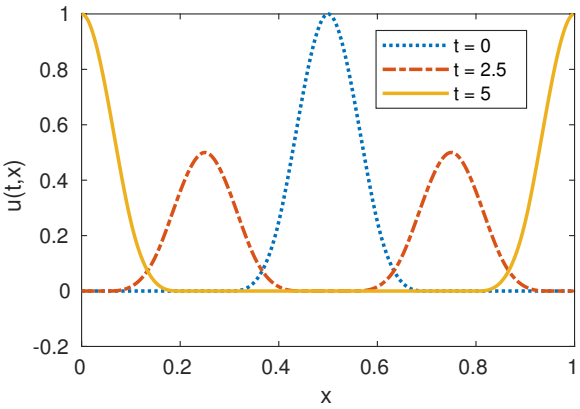

Fig. 1. Corresponding to Subsection 4.1: The function $u(t,x)$ based on the numerical solution of (10) at time $t = 0, 2.5, 5$.

The full-order data is now obtained by applying the symplectic Störmer-Verlet time-stepping scheme (5) with the parameters $l = 1$, $N = 500$, $\Delta x = \frac{l}{N}$, $t_0 = 0$, $t_{\text{end}} = 5$, $h = \Delta t = 0.01$, $c^2 = 0.01$. Fig. 1 displays snapshot solution functions $x \mapsto u(t,x) = q(t,x)$ at time instants $t = 0, 2.5, 5$.

We take snapshots at every 10th time step. This produces a snapshot matrix

$$\mathbb{S} = \left(\begin{pmatrix} q(t_1) \\ p(t_1) \end{pmatrix}, \ldots, \begin{pmatrix} q(t_m) \\ p(t_m) \end{pmatrix} \right) \in \mathbb{R}^{1000 \times 51}.$$

We use a fixed reduced dimension of $p = 5$. As symplectic starting points for optimizing (7), we consider $P_0 = U_0 U_0^+ \in \mathrm{SpGr}(2N, 2p)$, with the following options for constructing $U_0 \in \mathrm{SpSt}(2N, 2p)$:

(0) $U_0 = E := \begin{pmatrix} I_p & 0 & 0 & 0 \\ 0 & 0 & I_p & 0 \end{pmatrix}^T$;

(a) U_0 of the same dimensions produced by 'cotangent lift';

(b) U_0 of the same dimensions produced by 'complex SVD';

(c) U_0 of the same dimensions produced by 'SVD-like decomposition'.

Methods (a) and (b) are as discussed in (Peng and Mohseni, 2016, Section 4). Method (c) is from Buchfink et al. (2019), Algorithm 1. As termination conditions for the descent algorithm, we use a tolerance of 10^{-3} for the Frobenius norm of the Riemannian gradient and a tolerance of 10^{-6} for both $\frac{\|U_{k-1} - U_k\|_F}{\sqrt{2n}}$ and $\frac{|f(U_{k-1}) - f(U_k)|}{|f(U_k)|+1}$.

The results are stated in Table 1. The largest error in symplecticity $\|U^+U - I_{2p}\|_F$ after optimization from all starting points is generated by method (b) and amounts to 1.0968×10^{-13}, so that symplecticity can be considered as numerically preserved.

As an example, the convergence history corresponding to optimizing (7) when starting from 'complex SVD' is displayed in Fig. 2. It is worth mentioning that the optimization process arrives at different matrix representatives U_a^*, U_b^*, U_c^* depending on the chosen starting point, but that the associated optimal projector $P^* = U_a^*(U_a^*)^+ =$

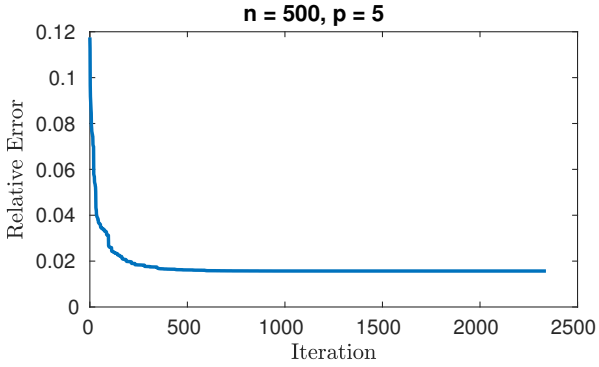

Fig. 2. Corresponding to Subsection 4.1: Relative error $\|\mathbb{S} - U_i U_i^+ \mathbb{S}\|_F / \|\mathbb{S}\|_F$ of the convergence history for minimizing (9) with starting point $P_0 = U_0 U_0^+$ obtained from the method of 'complex SVD'.

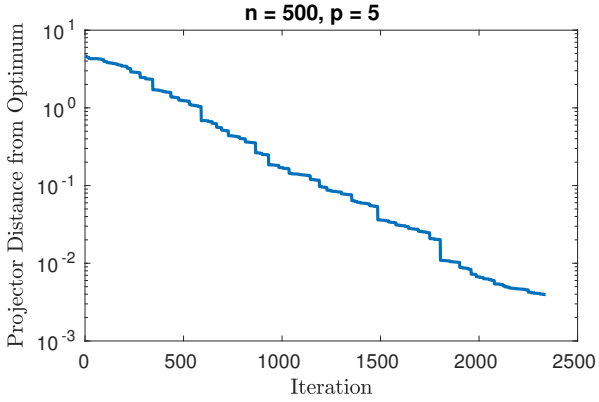

Fig. 3. Corresponding to Subsection 4.1: Projector distance (11) between the iterates with starting point $P_0 = U_0 U_0^+$ obtained from the method of 'complex SVD' and the optimum P^*.

$U_b^*(U_b^*)^+ = U_c^*(U_c^*)^+$ is the same up to the convergence accuracy. In Fig. 3 we denote the projector distance

$$\mathrm{dist}(P_1, P_2) := \|P_1 - P_2\|_F \tag{11}$$

between the iterates from case (b) and the assumed optimum P^*, where the latter is obtained by letting the descent algorithm run for 5000 iterations from the starting point obtained by the SVD-like decomposition. To further strengthen the conjecture that P^* is a *global* optimum, we started the algorithm from a random starting point P_0 with initial projector distance of $\mathrm{dist}(P_0, P^*) = 109.771$ from P^*, and observed that the algorithm converged against the same optimum after 24492 iterations. Moreover, the experiments show that it is possible to represent the snapshot data $\mathbb{S}$ with a rank-10 symplectic projector up to a relative error of less than 1.6%, which is a significant improvement on the initial guess, when started

Table 1. Optimizing the symplectic basis for the test case of Section 4.1. The relative error $\|\mathbb{S} - UU^+\mathbb{S}\|_F / \|\mathbb{S}\|_F$ is reported.

Start	init. error	error after opt.	iters.
(0) $U_0 = E$	0.9906	0.0157	3537
(a) cotangent lift	0.2407	0.0157	2200
(b) complex SVD	0.1277	0.0157	2338
(c) SVD-like decomp.	0.0160	0.0157	1444

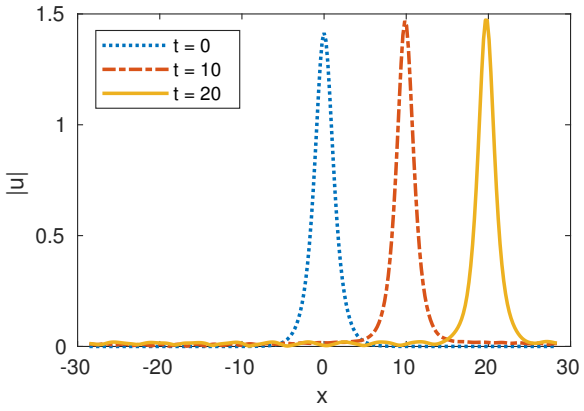

Fig. 4. Corresponding to Subsection 4.2: The function $|u(t,x)|$ based on the numerical solution of (12) at time $t = 0, 10, 20$.

from 'cotangent lift' or 'complex SVD'. The method 'SVD-like decomposition' produces a symplectic projector that is already close to numerical optimum so that the improvement is minor. For comparison, the relative error of POD projection of the same dimension is 0.0140, which is the optimum for any $2p$ dimensional subspace. The symplectic structure however is lost in that case, which may lead to high errors in the reduced model (Peng and Mohseni, 2016, Section 6).

4.2 Nonlinear Schrödinger equation

The one-dimensional parametric Schrödinger equation for a complex wave function $u(t,x)$ reads

$$\iota u_t(t,x) = -u_{xx}(t,x) - \epsilon|u(t,x)|^2 u(t,x),$$

with initial condition $u(t_0,x) = u_0(x)$. Here, $\iota = \sqrt{-1}$ denotes the complex unit. The spatial domain is $[-\frac{l}{2}, \frac{l}{2}]$. Again, we impose periodic boundary conditions and discretizing in space via D_{xx} as introduced above and steps $-\frac{l}{2}+i\Delta x, i = 0,\dots,N-1$, $\Delta x = \frac{l}{N}$. By writing $u = p+\iota q$, we arrive at the Hamiltonian system

$$\begin{pmatrix}\dot q\\ \dot p\end{pmatrix} = \mathbb{J}_{2n}\left(\begin{pmatrix}D_{xx} & 0\\ 0 & D_{xx}\end{pmatrix}\begin{pmatrix}q\\ p\end{pmatrix} + \epsilon\begin{pmatrix}a(q,p)\\ b(q,p)\end{pmatrix}\right) \qquad (12)$$

with nonlinear terms $a = (a_1,\dots,a_N)^T$ and $b = (b_1,\dots,b_N)^T$ defined by

$$a_i(q,p) = (q_i^2+p_i^2)q_i, \quad b_i(q,p) = (q_i^2+p_i^2)p_i.$$

The initial condition is set to be

$$u(0,x) = \frac{\sqrt{2}}{\cosh(x)}\exp\left(\frac{\iota}{2}cx\right), \quad x\in[-\frac{l}{2},\frac{l}{2}],$$

so that $p(0,x) = \Re(u(0,x))$ is the real part and $q(0,x) = \Im(u(0,x))$ is the imaginary part of $u(0,x)$.

Full-order simulations are conducted with the Störmer-Verlet time-stepping scheme (5) for the parameters $l = \frac{2\pi}{0.11}$, $N = 256$, $\Delta x = \frac{l}{N}$, $t_0 = 0$, $t_{\text{end}} = 20$, $h = \Delta t = 0.01$, $\epsilon = 1.0932$ and $c = 1$. Fig. 4 displays the function $|u(t,x)| = \sqrt{q^2(t,x)+p^2(t,x)}$, which is to be interpreted as a probability density function in quantum mechanics, for time instants $t = 0, 10, 20$.

We take snapshots at every 10th time step. This produces a snapshot matrix

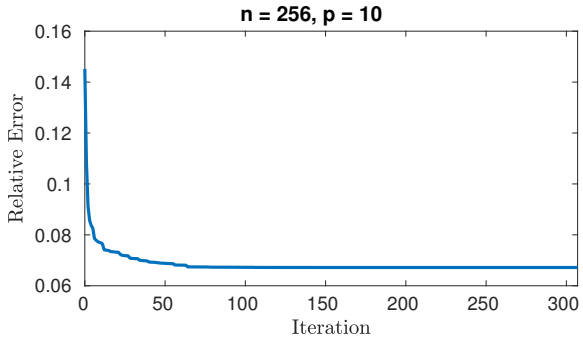

Fig. 5. Corresponding to Subsection 4.2: Relative error $\|\mathbb{S} - U_iU_i^+\mathbb{S}\|_F/\|\mathbb{S}\|_F$ of the convergence history for minimizing (9) with starting point $P_0 = U_0U_0^+$ obtained from the method of 'cotangent lift'.

$$\mathbb{S} = \left(\begin{pmatrix}q(t_1)\\ p(t_1)\end{pmatrix},\dots,\begin{pmatrix}q(t_m)\\ p(t_m)\end{pmatrix}\right) \in \mathbb{R}^{512\times 201}.$$

We use a fixed reduced dimension of $p = 10$. As symplectic starting points for optimizing (7), we consider the same three options as in Subsection 4.1: (0) $U_0 = E \in \mathrm{SpSt}(2N,2p)$; (a) U_0 via 'cotangent lift', (b) U_0 via 'complex SVD', (c) U_0 via 'SVD-like decomposition', and the same tolerances for convergence. The results are stated in Table 2. For comparison, the relative error obtained from POD is 0.0581. The largest error in symplecticity

Table 2. Optimizing the symplectic basis for the test case of Section 4.2. The relative error $\|\mathbb{S} - UU^+\mathbb{S}\|_F/\|\mathbb{S}\|_F$ is reported.

Start	init. error	error after opt.	iters.
(0) $U_0 = E$	1.0	0.0672	658
(a) cotangent lift	0.2608	0.0672	307
(b) complex SVD	0.1750	0.0672	409
(c) SVD-like decomp.	0.0857	0.0672	258

among all methods after optimization amounts to 8.4269×10^{-14} for case (a), so that symplecticity can be considered as numerically preserved. The convergence history corresponding to optimizing (7) when starting from (b) 'cotangent lift' is displayed in Fig. 5.

As in the previous example, upon convergence, the optimized projector $P^* = U_a^*(U_a^*)^+ = U_b^*(U_b^*)^+ = U_c^*(U_c^*)^+$ is the same (up to the convergence accuracy), regardless of which initial guess (0), (a), (b), (c) is chosen, even though the matrix representatives U_a^*, U_b^*, U_c^* differ. In Fig. 6 we show the projector distance between the iterates from case (a) and the assumed optimum P^*, obtained by letting the descent algorithm run for 5000 iterations from the starting point obtained by the SVD-like decomposition. Also in this case, the optimized projector improves significantly on the initial guesses produced by the first three methods. The best starting projector is provided by the method of 'SVD-like decomposition'; however, the experiment shows that this projector is not the global optimum.

5. CONCLUSION

From the experiments conducted in Section 4, we infer the following main conclusions:

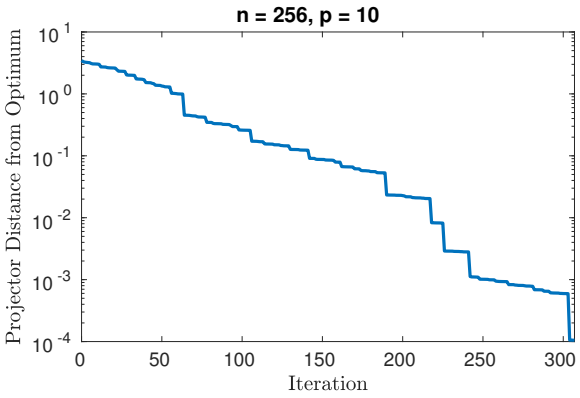

Fig. 6. Corresponding to Subsection 4.2: Projector distance (11) between the iterates with starting point $P_0 = U_0 U_0^+$ obtained from the method of 'cotangent lift' and the optimum P^*.

- It is appropriate to tackle the PSD problem (1) as an optimization problem (7) on the symplectic Grassmann manifold. The experiments suggest that from the Grassmann perspective, the problem is unimodal, i.e., it appears that there is an optimal symplectic projector which is the unique global minimum.
- The proposed Riemannian optimization approach to tackle (7) has shown the potential to significantly improve on a given initial guess. On the other hand, the experiments confirm, as was to be expected, that the SVD-based approaches 'cotangent lift' and 'complex SVD' produce suboptimal projectors.
- The question posed in Buchfink et al. (2019) whether the SVD-like decomposition produces the global optimum for the PSD problem can be answered negatively. However, it produces a projector that seems to be reasonably close to the optimum compared to other methods. Future investigations whether there is a matrix decomposition that produces the global optimum directly, analogously to the SVD in POD, seem worthwhile.
- A generic gradient descent method is able to reach the minimum. However, the method proceeds with small steps and the associated iteration count is rather high. This observation triggers the need to tailor-made optimization procedures, preferably of second order. Moreover, Figs. 2 & 5 show that, for practical purposes, the optimization procedure can be stopped earlier without compromising the prediction quality of the resulting projector. Hence, problem-adapted starting and stopping conditions are also of interest.
- Given that the conjectured unimodality of the optimization problem (7) holds true, it is fair to assume that the optimal symplectic projector associated with a given snapshot matrix $\mathbb{S}$ depends smoothly on any additional system parameters, provided that $\mathbb{S}$ is smooth in these parameters. This would then justify parametric interpolation of optimized symplectic projectors in the sense of Zimmermann (2021). All the necessary algorithmic tools for such a course of action are provided by Bendokat and Zimmermann (2021).

REFERENCES

Afkham, B.M. and Hesthaven, J.S. (2017). Structure preserving model reduction of parametric Hamiltonian systems. *SIAM Journal on Scientific Computing*, 39(6), A2616–A2644. doi:10.1137/17M1111991.

Arnol'd, V.I. and Givental', A.B. (2001). *Symplectic Geometry*, 1–138. Springer, Berlin, Heidelberg.

Arnol'd, V. (1997). *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics. Springer, New York.

Bendokat, T. and Zimmermann, R. (2021). The real symplectic Stiefel and Grassmann manifolds: metrics, geodesics and applications. URL `arxiv.org/abs/2108.12447`.

Buchfink, P., Haasdonk, B., and Rave, S. (2020). PSD-greedy basis generation for structure-preserving model order reduction of Hamiltonian systems. *Proceedings of the Conference Algoritmy*, 151–160.

Buchfink, P., Bhatt, A., and Haasdonk, B. (2019). Symplectic model order reduction with non-orthonormal bases. *Mathematical and Computational Applications*, 24(2), 43. doi:10.3390/mca24020043.

Gao, B., Son, N.T., Absil, P.A., and Stykel, T. (2021). Riemannian optimization on the symplectic Stiefel manifold. *SIAM Journal on Optimization*, 31(2), 1546–1575. doi:10.1137/20M1348522.

Hairer, E., Lubich, C., and Wanner, G. (2006). *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition. doi:10.1007/3-540-30666-8.

Musharbash, E., Nobile, F., and Vidličková, E. (2020). Symplectic dynamical low rank approximation of wave equations with random parameters. *BIT*, 60(4), 1153–1201. doi:10.1007/s10543-020-00811-6.

Peng, L. and Mohseni, K. (2016). Symplectic model reduction of Hamiltonian systems. *SIAM Journal on Scientific Computing*, 38(1), A1–A27. doi:10.1137/140978922.

Xu, H. (2003). An SVD-like matrix decomposition and its applications. *Linear Algebra and its Applications*, 368, 1–24. doi:10.1016/S0024-3795(03)00370-7.

Zimmermann, R. (2021). Manifold interpolation. In P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L.M. Silveira (eds.), *System- and Data-Driven Methods and Algorithms*, volume 1 of *Model Order Reduction*. De Gruyter, Boston.