



P. Buchfink^a · A. Bhatt^b · B. Haasdonk^a

Symplectic Model Order Reduction with Non-Orthonormal Bases

Stuttgart, February 2019

^a Institute of Applied Analysis and Numerical Simulation,
University of Stuttgart,
Pfaffenwaldring 57, D-70569 Stuttgart
{patrick.buchfink,haasdonk}@mathematik.uni-stuttgart.de
www.ians.uni-stuttgart.de

^b Indian Institute of Technology (ISM),
Dhanbad, Jharkhand 826004, India
ashishbhatt@iitism.ac.in
www.iitism.ac.in

Abstract Parametric high-fidelity simulations are of interest for a wide range of applications. But the restriction of computational resources renders such models to be inapplicable in a real-time context or in multi-query scenarios. Model order reduction (MOR) is used to tackle this issue. Recently, MOR is extended to preserve specific structures of the model throughout the reduction, e.g. structure-preserving MOR for Hamiltonian systems. This is referred to as symplectic MOR. It is based on the classical projection-based MOR and uses a symplectic reduced order basis (ROB). Such a ROB can be derived in a data-driven manner with the Proper Symplectic Decomposition (PSD) in the form of a minimization problem. Due to the strong nonlinearity of the minimization problem, it is unclear how to efficiently find a global optimum. In our paper, we show that current solution procedures almost exclusively yield suboptimal solutions by restricting to orthonormal ROBs. As new methodological contribution, we propose a new method which eliminates this restriction by generating non-orthonormal ROBs. In the numerical experiments, we examine the different techniques for a classical linear elasticity problem and observe that the non-orthonormal technique proposed in this paper shows superior results with respect to the error introduced by the reduction.

Keywords Symplectic model order reduction · proper symplectic decomposition (PSD) · structure preservation of symplecticity · Hamiltonian system

Stuttgart Research Centre for Simulation Technology (SRC SimTech)

SimTech – Cluster of Excellence
Pfaffenwaldring 5a
70569 Stuttgart
publications@simtech.uni-stuttgart.de
www.simtech.uni-stuttgart.de

Symplectic Model Order Reduction with Non-Orthonormal Bases

Patrick Buchfink^{1,*}, Ashish Bhatt^{2,†} and Bernard Haasdonk^{1,*}

¹ Institute of Applied Analysis and Numerical Simulation,
University of Stuttgart, 70569 Stuttgart, Germany

² Indian Institute of Technology (ISM), Dhanbad, Jharkhand 826004, India

February 26, 2019

Abstract

Parametric high-fidelity simulations are of interest for a wide range of applications. But the restriction of computational resources renders such models to be inapplicable in a real-time context or in multi-query scenarios. Model order reduction (MOR) is used to tackle this issue. Recently, MOR is extended to preserve specific structures of the model throughout the reduction, e.g. structure-preserving MOR for Hamiltonian systems. This is referred to as symplectic MOR. It is based on the classical projection-based MOR and uses a symplectic reduced order basis (ROB). Such a ROB can be derived in a data-driven manner with the Proper Symplectic Decomposition (PSD) in the form of a minimization problem. Due to the strong nonlinearity of the minimization problem, it is unclear how to efficiently find a global optimum. In our paper, we show that current solution procedures almost exclusively yield suboptimal solutions by restricting to orthonormal ROB. As new methodological contribution, we propose a new method which eliminates this restriction by generating non-orthonormal ROB. In the numerical experiments, we examine the different techniques for a classical linear elasticity problem and observe that the non-orthonormal technique proposed in this paper shows superior results with respect to the error introduced by the reduction.

Keywords— Symplectic model order reduction, proper symplectic decomposition (PSD), structure preservation of symplecticity, Hamiltonian system

1 Introduction

Simulations enable researchers of all fields to run virtual experiments that are too expensive or impossible to be carried out in the real world. In many contexts, high-fidelity models are indispensable to represent the simulated process accurately. These high-fidelity simulations typically come with the burden of large computational cost such that an application in real-time or an evaluation for many different parameters is impossible respecting the given restrictions of computational resources at hand. Model order reduction (MOR) techniques can be used to reduce the computational cost of evaluations of the high-fidelity model by approximating these with a surrogate reduced-order model (ROM) [4].

One class of high-fidelity models are systems of ordinary differential equations (ODEs) with a high order, i.e. a high dimension in the unknown variable. Such models typically arise from fine discretizations of time-dependent partial differential equations (PDEs). Since each point in the discretization requires one or multiple unknowns, fine discretizations with many discretization points yield a system of ODEs with a high order. In some cases, the ODE system takes the form of a finite-dimensional Hamiltonian system. Examples are linear elastic models [6] or gyro systems [20].

Symplectic MOR [19] allows to derive a ROM for high-dimensional Hamiltonian systems by lowering the order of the system while maintaining the Hamiltonian structure. Thus, it is also referred to as structure-preserving MOR for Hamiltonian systems [15]. Technically speaking, a Petrov–Galerkin projection is used in combination with a symplectic reduced-order basis (ROB).

*{patrick.buchfink,haasdonk}@ians.uni-stuttgart.de

†ashishbhatt@iitism.ac.in

For a data-driven generation of the ROB, the conventional methods e.g. the Proper Orthogonal Decomposition (POD) [4] are not suited since they do not necessarily compute a symplectic ROB. To this end, the referenced works introduce the Proper Symplectic Decomposition (PSD) which is a data-driven basis generation technique for symplectic ROBs. Due to the high nonlinearity of the optimization problem, an efficient solution strategy is yet unknown for the PSD. The existing PSD methods (Cotangent Lift, Complex SVD, a nonlinear programming approach [19] and a greedy procedure introduced in [15]) each restrict to a specific subset of symplectic ROBs from which they select optimal solutions which might be globally suboptimal.

The present paper classifies the existing symplectic basis generation techniques in two classes of methods which either generate orthonormal or non-orthonormal bases. To this end, we show that the existing basis generation techniques for symplectic bases almost exclusively restrict to orthonormal bases. Furthermore, we prove that Complex SVD is the optimal solution of the PSD on the set of orthonormal, symplectic bases. During the proof, an alternative formulation of the Complex SVD for symplectic matrices is introduced. To leave the class of orthonormal, symplectic bases, we propose a new basis generation technique, namely the PSD SVD-like decomposition. It is based on an SVD-like decomposition of arbitrary matrices $\mathbf{B} \in \mathbb{R}^{n \times 2m}$ introduced in [21].

This paper is organized in the following way: Section 2 is devoted to the structure-preserving MOR for autonomous and non-autonomous, parametric Hamiltonian systems and thus, introduces symplectic geometry, Hamiltonian systems and symplectic MOR successively. The data-driven generation of a symplectic ROB with PSD is discussed in Section 3. The numerical results are presented and elaborated in Section 4 exemplified by a Lamé-Navier type elasticity model which we introduce at the beginning of that section together with a short comment on the software that is used for the experiments. The paper is summarized and concluded in Section 5.

2 Symplectic model reduction

Symplectic MOR for autonomous Hamiltonian systems is introduced in [19]. We repeat the essentials for the sake of completeness and to provide a deeper understanding of the methods used. In the following $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$ describe $p \in \mathbb{N}$ parameters of the system from the parameter set \mathcal{P} . We might skip the explicit dependence on the parameter vector $\boldsymbol{\mu}$ if it is not relevant in this specific context.

2.1 Symplectic geometry in finite dimensions

Definition 1 (Symplectic form over \mathbb{R}). *Let \mathbb{V} be a finite-dimensional vector space over \mathbb{R} . We consider a skew-symmetric and non-degenerate bilinear form $\omega : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$, i.e. for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$, it holds*

$$\omega(\mathbf{v}_1, \mathbf{v}_2) = -\omega(\mathbf{v}_2, \mathbf{v}_1) \quad \text{and} \quad \omega(\mathbf{v}_2, \mathbf{v}_3) = 0 \quad \forall \mathbf{v}_3 \in \mathbb{V} \implies \mathbf{v}_3 = \mathbf{0}.$$

The bilinear form ω is called symplectic form on \mathbb{V} and the pair (\mathbb{V}, ω) is called symplectic vector space.

It can be shown that \mathbb{V} is necessarily of even dimension [8]. Thus, \mathbb{V} is isomorphic to \mathbb{R}^{2n} which is why we restrict to $\mathbb{V} = \mathbb{R}^{2n}$ and write ω_{2n} instead of ω in the following. In context of the theory of Hamiltonians, \mathbb{R}^{2n} refers to the phase space which consists, in the context of classical mechanics, of position states $\mathbf{q} = [q_1, \dots, q_n]^T \in \mathbb{R}^n$ of the configuration space and momentum states $\mathbf{p} = [p_1, \dots, p_n]^T \in \mathbb{R}^n$ which form together the state $\mathbf{x} = [q_1, \dots, q_n, p_1, \dots, p_n]^T \in \mathbb{R}^{2n}$.

It is guaranteed [8] that there exists a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n, \mathbf{f}_1, \dots, \mathbf{f}_n\} \subset \mathbb{R}^{2n}$ such that the symplectic form takes the canonical structure

$$\omega_{2n}(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1^T \mathbb{J}_{2n} \mathbf{v}_2 \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{2n}, \quad \mathbb{J}_{2n} := \begin{bmatrix} \mathbf{0}_n & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0}_n \end{bmatrix}, \quad (1)$$

where $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ is the identity matrix, $\mathbf{0}_n \in \mathbb{R}^{n \times n}$ is the matrix of all zeros and \mathbb{J}_{2n} is called Poisson matrix. Thus, we restrict to symplectic forms of the canonical structure in the following. For the Poisson matrix, it holds for any $\mathbf{v} \in \mathbb{R}^{2n}$

$$\mathbb{J}_{2n} \mathbb{J}_{2n}^T = \mathbf{I}_{2n}, \quad \mathbb{J}_{2n} \mathbb{J}_{2n} = \mathbb{J}_{2n}^T \mathbb{J}_{2n}^T = -\mathbf{I}_{2n}, \quad \mathbf{v}^T \mathbb{J}_{2n} \mathbf{v} = 0. \quad (2)$$

These properties are intuitively understandable as the Poisson matrix is a $2n$ -dimensional, 90° rotation matrix and the matrix $-\mathbf{I}_{2n}$ can be interpreted as a rotation by 180° in this context.

Definition 2 (Symplectic map). *Let $A : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2n}$, $\mathbf{y} \mapsto \mathbf{A}\mathbf{y}$, $\mathbf{A} \in \mathbb{R}^{2n \times 2m}$ be a linear mapping for $n, m \in \mathbb{N}$ and $m \leq n$. We call A a linear symplectic map and \mathbf{A} a symplectic matrix with respect to ω_{2n} and ω_{2m} if*

$$\mathbf{A}^T \mathbb{J}_{2n} \mathbf{A} = \mathbb{J}_{2m}. \quad (3)$$

where ω_{2m} is the canonical symplectic form on \mathbb{R}^{2m} (and is equal to ω_{2n} if $n = m$).

Let $U \subset \mathbb{R}^{2m}$ be an open set and $\mathbf{g} : U \rightarrow \mathbb{R}^{2n}$ a differentiable map on U . We call \mathbf{g} a symplectic map if the Jacobian matrix $\frac{d}{d\mathbf{y}} \mathbf{g}(\mathbf{y}) \in \mathbb{R}^{2n \times 2m}$ is a symplectic matrix for every $\mathbf{y} \in U$.

For a linear map, it is easy to check that the condition (3) is equivalent to the preservation of the symplectic form, i.e. for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{2m}$

$$\omega_{2n}(\mathbf{A}\mathbf{v}_1, \mathbf{A}\mathbf{v}_2) = \mathbf{v}_1^\top \mathbf{A}^\top \mathbb{J}_{2n} \mathbf{A} \mathbf{v}_2 = \mathbf{v}_1^\top \mathbb{J}_{2m} \mathbf{v}_2 = \omega_{2m}(\mathbf{v}_1, \mathbf{v}_2).$$

Now we give the definition of the so-called symplectic inverse which will be used in Section 2.3.

Definition 3 (Symplectic inverse). *For each symplectic matrix $\mathbf{A} \in \mathbb{R}^{2n \times 2m}$, we define the symplectic inverse*

$$\mathbf{A}^+ = \mathbb{J}_{2m}^\top \mathbf{A}^\top \mathbb{J}_{2n} \in \mathbb{R}^{2m \times 2n}. \quad (4)$$

The symplectic inverse \mathbf{A}^+ exists for every symplectic matrix and it holds the following inverse relation

$$\mathbf{A}^+ \mathbf{A} = \mathbb{J}_{2m}^\top \mathbf{A}^\top \mathbb{J}_{2n} \mathbf{A} = \mathbb{J}_{2m}^\top \mathbb{J}_{2m} = \mathbf{I}_{2m}.$$

2.2 Finite-dimensional, autonomous Hamiltonian systems

To begin with, we introduce the Hamiltonian system in a finite-dimensional, autonomous setting.

Definition 4 (Finite-dimensional, autonomous Hamiltonian system). *Let $\mathcal{H} : \mathbb{R}^{2n} \times \mathcal{P} \rightarrow \mathbb{R}$ be a scalar-valued function that we require to be continuously differentiable in the first argument and which we call Hamiltonian (function). Hamilton's equation is an initial value problem with the prescribed initial data $t_0 \in \mathbb{R}$, $\mathbf{x}_0(\boldsymbol{\mu}) \in \mathbb{R}^{2n}$ which describes the evolution of the solution $\mathbf{x}(t, \boldsymbol{\mu}) \in \mathbb{R}^{2n}$ for all $t \in [t_0, t_{\text{end}}]$, $\boldsymbol{\mu} \in \mathcal{P}$ with*

$$\frac{d}{dt} \mathbf{x}(t, \boldsymbol{\mu}) = \mathbb{J}_{2n} \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x}(t, \boldsymbol{\mu}), \boldsymbol{\mu}) =: \mathbf{X}_{\mathcal{H}}(\mathbf{x}(t, \boldsymbol{\mu}), \boldsymbol{\mu}), \quad \mathbf{x}(t_0, \boldsymbol{\mu}) = \mathbf{x}_0(\boldsymbol{\mu}) \quad (5)$$

where $\mathbf{X}_{\mathcal{H}}(\bullet, \boldsymbol{\mu})$ is called Hamiltonian vector field. The triple $(\mathbb{V}, \omega_{2n}, \mathcal{H})$ is referred to as Hamiltonian system. We denote the flow of a Hamiltonian system as the mapping $\varphi_t : \mathbb{R}^{2n} \times \mathcal{P} \rightarrow \mathbb{R}^{2n}$ that evolves the initial state $\mathbf{x}_0(\boldsymbol{\mu}) \in \mathbb{R}^{2n}$ to the corresponding solution $\mathbf{x}(t, \boldsymbol{\mu}; t_0, \mathbf{x}_0(\boldsymbol{\mu}))$ of Hamilton's equation

$$\varphi_t(\mathbf{x}_0, \boldsymbol{\mu}) := \mathbf{x}(t, \boldsymbol{\mu}; t_0, \mathbf{x}_0(\boldsymbol{\mu})),$$

where $\mathbf{x}(t, \boldsymbol{\mu}; t_0, \mathbf{x}_0(\boldsymbol{\mu}))$ indicates that it is the solution with the initial data $t_0, \mathbf{x}_0(\boldsymbol{\mu})$.

The two characteristic properties of Hamiltonian systems are (a) the preservation of the Hamiltonian function and (b) the symplecticity of the flow.

Proposition 1 (Preservation of the Hamiltonian). *The flow of Hamilton's equation φ_t preserves the Hamiltonian function \mathcal{H} .*

Proof. We prove the assertion by showing that the evolution over time is constant for any $\mathbf{x} \in \mathbb{R}^{2n}$ due to

$$\frac{d}{dt} \mathcal{H}(\varphi_t(\mathbf{x})) = (\nabla_{\mathbf{x}} \mathcal{H}(\varphi_t(\mathbf{x})))^\top \frac{d}{dt} \varphi_t(\mathbf{x}) \stackrel{(5)}{=} (\nabla_{\mathbf{x}} \mathcal{H}(\varphi_t(\mathbf{x})))^\top \mathbb{J}_{2n} \nabla_{\mathbf{x}} \mathcal{H}(\varphi_t(\mathbf{x})) \stackrel{(2)}{=} 0.$$

□

Proposition 2 (Symplecticity of the flow). *Let the Hamiltonian function be twice continuously differentiable in the first argument. Then, the flow $\varphi_t(\bullet, \boldsymbol{\mu}) : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ of a Hamiltonian system is a symplectic map.*

Proof. See [11, Chapter VI, Theorem 2.4]. □

2.3 Symplectic model order reduction for autonomous Hamiltonian systems

The goal of MOR [4] is to reduce the order, i.e. the dimension, of high dimensional systems. To this end, we approximate the high-dimensional state $\mathbf{x}(t) \in \mathbb{R}^{2n}$ with

$$\mathbf{x}(t, \boldsymbol{\mu}) \approx \mathbf{x}_{\text{rc}}(t, \boldsymbol{\mu}) = \mathbf{V} \mathbf{x}_{\text{r}}(t, \boldsymbol{\mu}), \quad \mathcal{V} = \text{colspan}(\mathbf{V})$$

with the reduced state $\mathbf{x}_{\text{r}}(t) \in \mathbb{R}^{2k}$, the reduced-order basis (ROB) $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$, the reconstructed state $\mathbf{x}_{\text{rc}}(t) \in \mathcal{V}$ and the reduced space $\mathcal{V} \subset \mathbb{R}^{2n}$. The restriction to even-dimensional spaces \mathbb{R}^{2n} and \mathbb{R}^{2k} is not necessary for MOR in general but is required for the symplectic MOR in the following. To achieve a computational advantage with MOR, the approximation should introduce a clear reduction of the order, i.e. $2k \ll 2n$.

For Petrov–Galerkin projection-based MOR techniques, the ROB \mathbf{V} is accompanied by a projection matrix $\mathbf{W} \in \mathbb{R}^{2n \times 2k}$ which is chosen to be biorthogonal to \mathbf{V} , i.e. $\mathbf{W}^\top \mathbf{V} = \mathbf{I}_{2k}$. The reduced-order model (ROM) is

derived with the requirement that the residual $\mathbf{r}(t, \boldsymbol{\mu})$ vanishes in the space spanned by the columns of the projection matrix, i.e. in our case

$$\mathbf{r}(t, \boldsymbol{\mu}) = \frac{d}{dt} \mathbf{x}_{rc}(t, \boldsymbol{\mu}) - \mathbf{X}_{\mathcal{H}}(\mathbf{x}_{rc}(t, \boldsymbol{\mu}), \boldsymbol{\mu}) \in \mathbb{R}^{2n}, \quad \mathbf{W}^T \mathbf{r}(t, \boldsymbol{\mu}) = \mathbf{0}_{2k \times 1}, \quad (6)$$

where $\mathbf{0}_{2k \times 1} \in \mathbb{R}^{2k}$ is the vector of all zeros. Due to the biorthogonality, this is equivalent to

$$\frac{d}{dt} \mathbf{x}_r(t, \boldsymbol{\mu}) = \mathbf{W}^T \mathbf{X}_{\mathcal{H}}(\mathbf{x}_{rc}(t, \boldsymbol{\mu}), \boldsymbol{\mu}) = \mathbf{W}^T \mathbb{J}_{2n} \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x}_{rc}(t, \boldsymbol{\mu}), \boldsymbol{\mu}), \quad \mathbf{x}_r(t_0, \boldsymbol{\mu}) = \mathbf{W}^T \mathbf{x}_0(\boldsymbol{\mu}). \quad (7)$$

In the context of symplectic MOR, the ROB is chosen to be a symplectic matrix (3) which we call a symplectic ROB. Additionally, the transposed projection matrix is the symplectic inverse $\mathbf{W}^T = \mathbf{V}^+$ and the projection in (7) is called a symplectic projection or symplectic Galerkin projection [19]. The (possibly oblique) projection reads

$$\mathbf{P} = \mathbf{V} (\mathbf{W}^T \mathbf{V})^{-1} \mathbf{W}^T = \mathbf{V} (\mathbf{V}^+ \mathbf{V})^{-1} \mathbf{V}^+ = \mathbf{V} \mathbf{V}^+.$$

In combination, this choice of \mathbf{V} and \mathbf{W} guarantees that the Hamiltonian structure is preserved by the reduction which is shown in the following proposition.

Proposition 3 (Reduced autonomous Hamiltonian system). *Let \mathbf{V} be a symplectic ROB with the projection matrix $\mathbf{W}^T = \mathbf{V}^+$. Then, the ROM (7) of a high-dimensional Hamiltonian system $(\mathbb{R}^{2n}, \omega_{2n}, \mathcal{H})$ is a Hamiltonian system $(\mathbb{R}^{2k}, \omega_{2k}, \mathcal{H}_r)$ on \mathbb{R}^{2k} with the canonical symplectic form ω_{2k} and the reduced Hamiltonian function $\mathcal{H}_r(\mathbf{x}_r, \boldsymbol{\mu}) = \mathcal{H}(\mathbf{V} \mathbf{x}_r, \boldsymbol{\mu})$ for all $\mathbf{x}_r \in \mathbb{R}^{2k}$.*

Proof. First, we remark that the symplectic inverse is a valid biorthogonal projection matrix since it fulfils $\mathbf{W}^T \mathbf{V} = \mathbf{V}^+ \mathbf{V} = \mathbf{I}_{2k}$. To derive the Hamiltonian form of the ROM in (7), we use the identity

$$\mathbf{W}^T \mathbb{J}_{2n} = \mathbf{V}^+ \mathbb{J}_{2n} \stackrel{(4)}{=} \mathbb{J}_{2k}^T \mathbf{V}^T \mathbb{J}_{2n} \mathbb{J}_{2n} = -\mathbb{J}_{2k}^T \mathbf{V}^T = \mathbb{J}_{2k} \mathbf{V}^T, \quad (8)$$

which makes use of the properties (2) of the Poisson matrix. It follows with (5), (7) and (8)

$$\frac{d}{dt} \mathbf{x}_r(t) = \mathbf{W}^T \mathbb{J}_{2n} \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x}_{rc}(t)) = \mathbb{J}_{2k} \mathbf{V}^T \nabla_{\mathbf{x}} \mathcal{H}(\mathbf{x}_{rc}(t)) = \mathbb{J}_{2k} \nabla_{\mathbf{x}_r} \mathcal{H}_r(\mathbf{x}_r(t))$$

where the last step follows from the chain rule. Thus, the evolution of the reduced state takes the form of Hamilton's equation and the resultant ROM is equal to the Hamiltonian system $(\mathbb{R}^{2k}, \omega_{2k}, \mathcal{H}_r)$. \square

Corollary 1 (Linear Hamiltonian system). *Hamilton's equation is a linear system in the case of a quadratic Hamiltonian $\mathcal{H}(\mathbf{x}, \boldsymbol{\mu}) = 1/2 \mathbf{x}^T \mathbf{H}(\boldsymbol{\mu}) \mathbf{x} + \mathbf{x}^T \mathbf{h}(\boldsymbol{\mu})$ with $\mathbf{H}(\boldsymbol{\mu}) \in \mathbb{R}^{2n \times 2n}$ symmetric and $\mathbf{h}(\boldsymbol{\mu}) \in \mathbb{R}^{2n}$*

$$\frac{d}{dt} \mathbf{x}(t, \boldsymbol{\mu}) = \mathbf{A}(\boldsymbol{\mu}) \mathbf{x}(t, \boldsymbol{\mu}) + \mathbf{b}(\boldsymbol{\mu}), \quad \mathbf{A}(\boldsymbol{\mu}) = \mathbb{J}_{2n} \mathbf{H}(\boldsymbol{\mu}), \quad \mathbf{b}(\boldsymbol{\mu}) = \mathbb{J}_{2n} \mathbf{h}(\boldsymbol{\mu}). \quad (9)$$

The evolution of the reduced Hamiltonian system reads

$$\begin{aligned} \frac{d}{dt} \mathbf{x}_r(t, \boldsymbol{\mu}) &= \mathbf{A}_r(\boldsymbol{\mu}) \mathbf{x}_r(t, \boldsymbol{\mu}) + \mathbf{b}_r(\boldsymbol{\mu}), & \mathbf{A}_r(\boldsymbol{\mu}) &= \mathbb{J}_{2k} \mathbf{H}_r(\boldsymbol{\mu}) \stackrel{(8)}{=} \mathbf{W}^T \mathbf{A}(\boldsymbol{\mu}) \mathbf{V}, & \mathbf{H}_r(\boldsymbol{\mu}) &= \mathbf{V}^T \mathbf{H}(\boldsymbol{\mu}) \mathbf{V}, \\ & & \mathbf{b}_r(\boldsymbol{\mu}) &= \mathbb{J}_{2k} \mathbf{h}_r(\boldsymbol{\mu}) \stackrel{(8)}{=} \mathbf{W}^T \mathbf{b}(\boldsymbol{\mu}) \mathbf{V}, & \mathbf{h}_r(\boldsymbol{\mu}) &= \mathbf{V}^T \mathbf{h}(\boldsymbol{\mu}). \end{aligned}$$

with the reduced Hamiltonian function $\mathcal{H}_r(\mathbf{x}_r, \boldsymbol{\mu}) = 1/2 \mathbf{x}_r^T \mathbf{H}_r(\boldsymbol{\mu}) \mathbf{x}_r + \mathbf{x}_r^T \mathbf{h}_r(\boldsymbol{\mu})$.

Remark 1. We emphasise that the reduction of linear Hamiltonian systems follows the pattern of the classical projection-based MOR approaches [10] to derive the reduced model with $\mathbf{A}_r = \mathbf{W}^T \mathbf{A} \mathbf{V}$ and $\mathbf{b}_r = \mathbf{W}^T \mathbf{b}$ which allows a straightforward implementation in existing frameworks.

Since the ROM is a Hamiltonian system, it preserves its Hamiltonian. Thus, it can be shown that the error in the Hamiltonian $e_{\mathcal{H}}(t, \boldsymbol{\mu}) = \mathcal{H}(\mathbf{x}(t, \boldsymbol{\mu}), \boldsymbol{\mu}) - \mathcal{H}_r(\mathbf{x}_r(t, \boldsymbol{\mu}), \boldsymbol{\mu})$ is constant [19]. Furthermore, there are a couple of results for the preservation of stability [15, Theorem 18], [19, Section 3.4.] under certain assumptions on the Hamiltonian function.

Remark 2 (Offline/online decomposition). *A central concept in the field of MOR for parametric systems is the so-called offline/online decomposition. The idea is to split the procedure in a possibly costly offline phase and a cheap online phase where the terms costly and cheap refer to the computational cost. In the offline phase, the ROM is constructed. The online phase is supposed to evaluate the ROM fast. The ultimate goal is to avoid any computations that depend on the high dimension $2n$ in the online phase.*

For a linear system, the offline/online decomposition can be achieved if $\mathbf{A}(\boldsymbol{\mu})$, $\mathbf{b}(\boldsymbol{\mu})$ and $\mathbf{x}_0(\boldsymbol{\mu})$ allow a parameter-separability condition [10].

For systems with non-linear parts, multiple approaches [3, 9] exist to enable an offline/online decomposition by introducing an approximation of the non-linear terms. This allows an online-efficient MOR of non-linear systems. For symplectic MOR, the symplectic discrete empirical interpolation method (SDEIM) was introduced [19, Section 5.2.] to preserve the symplectic structure throughout the approximation of the non-linear terms.

2.4 Finite-dimensional, non-autonomous Hamiltonian systems

Non-autonomous Hamiltonian systems can be redirected to the case of autonomous systems if differentiability with respect to the time is assumed for the Hamiltonian function. The concept of the extended phase space is used. We briefly introduce the approach and explain the link to the symplectic MOR.

Definition 5 (Finite-dimensional, non-autonomous Hamiltonian system). *Let $\mathcal{H} : \mathbb{R} \times \mathbb{R}^{2n} \times \mathcal{P} \rightarrow \mathbb{R}$ be a scalar-valued function that is continuously differentiable in the second argument. A non-autonomous (or time-dependent) Hamiltonian system $(\mathbb{R}^{2n}, \omega_{2n}, \mathcal{H})$ is of the form*

$$\dot{\mathbf{x}}(t, \boldsymbol{\mu}) = \mathbb{J}_{2n} \nabla_{\mathbf{x}} \mathcal{H}(t, \mathbf{x}(t, \boldsymbol{\mu}), \boldsymbol{\mu}). \quad (10)$$

We therefore call $\mathcal{H}(t, \mathbf{x})$ a time-dependent Hamiltonian function.

A problem for non-autonomous Hamiltonian systems occurs as the explicit time dependence of the Hamiltonian function introduces an additional variable, the time, and the carrier manifold becomes odd-dimensional. As mentioned in Section 2.1, symplectic vector spaces are always even-dimensional which is why a symplectic description is no longer possible. Different approaches are available to circumvent this issue.

As suggested in [16, Section 4.3], we use the methodology of the so-called symplectic extended phase space [12, Chap. VI, Sec. 10] to redirect the non-autonomous system to an autonomous system. The formulation is based on the extended Hamiltonian function $\mathcal{H}^e : \mathbb{R}^{2n+2} \rightarrow \mathbb{R}$ with

$$\mathcal{H}^e(\mathbf{x}^e) = \mathcal{H}(q^e, \mathbf{x}) + p^e, \quad \mathbf{x}^e = (q^e \ \mathbf{q} \ p^e \ \mathbf{p})^\top \in \mathbb{R}^{2n+2}, \quad \mathbf{x} = (\mathbf{q} \ \mathbf{p})^\top \in \mathbb{R}^{2n}, \quad q^e, p^e \in \mathbb{R}. \quad (11)$$

Technically, the time is added to the extended state \mathbf{x}^e with $q^e = t$ and the corresponding momentum $p^e = -\mathcal{H}(t, \mathbf{x}(t))$ is chosen such that the extended system is an autonomous Hamiltonian system.

This procedure requires the time-dependent Hamiltonian function to be differentiable in the time variable. Thus, it does for example not allow for the description of loads that are not differentiable in time in the context of mechanical systems. This might, e.g., exclude systems that model mechanical contact since loads that are not differentiable in time are required.

2.5 Symplectic model order reduction of non-autonomous Hamiltonian systems

For the MOR of the, now autonomous, extended system, only the original phase space variable $\mathbf{x} \in \mathbb{R}^{2n}$ is reduced. The time and the corresponding conjugate momentum q^e, p^e are not reduced. To preserve the Hamiltonian structure, a symplectic ROB $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ is used for the reduction of $\mathbf{x} \in \mathbb{R}^{2n}$ analogous to the autonomous case. The result is a reduced extended system which again can be written as a non-autonomous Hamiltonian system $(\mathbb{R}^{2k}, \omega_{2k}, \mathcal{H}_r)$ with the time-dependent Hamiltonian $\mathcal{H}_r(t, \mathbf{x}_r, \boldsymbol{\mu}) = \mathcal{H}(t, \mathbf{V} \mathbf{x}_r, \boldsymbol{\mu})$ for all $(t, \mathbf{x}_r) \in [t_0, t_{\text{end}}] \times \mathbb{R}^{2k}$.

An unpleasant side effect of the extended formulation is that the linear dependency on the additional state variable p^e (see (11)) implies that the Hamiltonian cannot have strict extrema. Thus, the stability results listed in [19] and [15] do not apply if there is a true time-dependence in the Hamiltonian $\mathcal{H}(t, \mathbf{x})$. Nevertheless, symplectic MOR in combination with a non-autonomous Hamiltonian system shows stable results in the numerical experiments.

Furthermore, it is important to note that only the extended Hamiltonian \mathcal{H}^e is preserved throughout the reduction. The time-dependent Hamiltonian $\mathcal{H}(\cdot, t)$ is not necessarily preserved throughout the reduction, i.e. $\mathcal{H}^e(\mathbf{x}^e(t)) = \mathcal{H}_r^e(\mathbf{x}_r^e(t))$ but potentially $\mathcal{H}(\mathbf{x}(t), t) \neq \mathcal{H}_r(\mathbf{x}_r(t), t)$.

3 Symplectic basis generation with the Proper Symplectic Decomposition (PSD)

We yet require a symplectic ROB for symplectic MOR. In the following, we pursue the approach of a ROB generated from a set of snapshots of the system. A snapshot is an element of the so-called solution manifold \mathcal{S} that is approximated with a low-dimensional surrogate $\widehat{\mathcal{S}}_{\mathbf{V}\mathbf{W}}$

$$\mathcal{S} := \{\mathbf{x}(t, \boldsymbol{\mu}) \mid t \in [t_0, t_{\text{end}}], \boldsymbol{\mu} \in \mathcal{P}\} \subset \mathbb{R}^{2n}, \quad \widehat{\mathcal{S}}_{\mathbf{V}\mathbf{W}} := \{\mathbf{V} \mathbf{x}_r(t, \boldsymbol{\mu}) \mid t \in [t_0, t_{\text{end}}], \boldsymbol{\mu} \in \mathcal{P}\} \approx \mathcal{S}.$$

In [19], the Proper Symplectic Decomposition (PSD) is proposed as a snapshot-based basis generation technique for symplectic ROBs. The idea is to derive the ROB from a minimization problem which is suggested in analogy to the very well established Proper Orthogonal Decomposition (POD, also Principal Component Analysis) [4].

Classically, the POD chooses the ROB \mathbf{V}_{POD} to minimize the sum over squared norms of all $n_s \in \mathbb{N}$ residuals $(\mathbf{I}_{2n} - \mathbf{V}_{\text{POD}}\mathbf{V}_{\text{POD}}^T)\mathbf{x}_i^s$ of the orthogonal projection $\mathbf{V}_{\text{POD}}\mathbf{V}_{\text{POD}}^T\mathbf{x}_i^s$ of the $1 \leq i \leq n_s$ single snapshots $\mathbf{x}_i^s \in \mathcal{S}$ measured in the 2-norm $\|\bullet\|_2$ with the constraint that the ROB \mathbf{V}_{POD} is orthogonal, i.e.

$$\underset{\mathbf{V}_{\text{POD}} \in \mathbb{R}^{2n \times 2k}}{\text{minimize}} \sum_{i=1}^{n_s} \left\| (\mathbf{I}_{2n} - \mathbf{V}_{\text{POD}}\mathbf{V}_{\text{POD}}^T) \mathbf{x}_i^s \right\|_2^2 \quad \text{subject to} \quad \mathbf{V}_{\text{POD}}^T \mathbf{V}_{\text{POD}} = \mathbf{I}_{2k}. \quad (12)$$

In contrast, the PSD requires the ROB to be symplectic instead of orthogonal which is expressed in the reformulated constraint. Furthermore, the orthogonal projection is replaced by the symplectic projection $\mathbf{V}\mathbf{V}^+\mathbf{x}_i^s$ which results in

$$\underset{\mathbf{V} \in \mathbb{R}^{2n \times 2k}}{\text{minimize}} \sum_{i=1}^{n_s} \left\| (\mathbf{I}_{2n} - \mathbf{V}\mathbf{V}^+) \mathbf{x}_i^s \right\|_2^2 \quad \text{subject to} \quad \mathbf{V}^T \mathbb{J}_{2n} \mathbf{V} = \mathbb{J}_{2k}. \quad (13)$$

We summarize this in a more compact (matrix-based) formulation in the following definition.

Definition 6 (Proper Symplectic Decomposition (PSD)). *Given n_s snapshots $\mathbf{x}_1^s, \dots, \mathbf{x}_{n_s}^s \in \mathcal{S}$, we denote the snapshot matrix as $\mathbf{X}_s = [\mathbf{x}_1^s, \dots, \mathbf{x}_{n_s}^s] \in \mathbb{R}^{2n \times n_s}$. Find a symplectic ROB $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ which minimizes*

$$\underset{\mathbf{V} \in \mathbb{R}^{2n \times 2k}}{\text{minimize}} \left\| (\mathbf{I}_{2n} - \mathbf{V}\mathbf{V}^+) \mathbf{X}_s \right\|_{\text{F}}^2 \quad \text{subject to} \quad \mathbf{V}^T \mathbb{J}_{2n} \mathbf{V} = \mathbb{J}_{2k}, \quad (14)$$

We denote the minimization problem (14) in the following as $\text{PSD}(\mathbf{X}_s)$, where \mathbf{X}_s is the given snapshot matrix.

The constraint in (14) ensures that the ROB \mathbf{V} is symplectic and thus, guarantees the existence of the symplectic inverse \mathbf{V}^+ . Furthermore, the matrix-based formulation (14) is equivalent to the vector-based formulation presented in (13) due to the properties of the Frobenius norm $\|\bullet\|_{\text{F}}$.

3.1 Symplectic, orthonormal basis generation

The foremost problem of the PSD is that there is no explicit solution procedure known so far due to the high nonlinearity and possibly multiple local optima. This is an essential difference to the POD as the POD allows to find a global minimum by solving an eigenvalue problem [4].

Current solution procedures for the PSD restrict to a certain subset of symplectic matrices and derive an optimal solution for this subset which might be suboptimal in the class of symplectic matrices. In the following, we show that this subclass almost exclusively restricts to symplectic, orthonormal ROB.

Definition 7 (Symplectic, orthonormal ROB). *We call a ROB $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ symplectic, orthonormal (also orthosymplectic, e.g. in [15]) if it is symplectic w.r.t. ω_{2n} and ω_{2k} and is orthonormal, i.e. the matrix \mathbf{V} has orthonormal columns*

$$\mathbf{V}^T \mathbb{J}_{2n} \mathbf{V} = \mathbb{J}_{2k} \quad \text{and} \quad \mathbf{V}^T \mathbf{V} = \mathbf{I}_{2k}.$$

In the following, we show an alternative characterization of a symplectic and orthonormal ROB. Therefore, we extend the results given e.g. in [18] for square matrices $\mathbf{Q} \in \mathbb{R}^{2n \times 2n}$ in the following Proposition 4 to the case of rectangular matrices $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$. This was also partially addressed in [19, Lemma 4.3.].

Proposition 4 (Characterization of a symplectic matrix with orthonormal columns). *The following statements are equivalent for any matrix $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$*

- (i) \mathbf{V} is symplectic with orthonormal columns,
- (ii) \mathbf{V} is of the form

$$\mathbf{V} = \begin{bmatrix} \mathbf{E} & \mathbb{J}_{2n}^T \mathbf{E} \end{bmatrix} =: \mathbf{V}_{\mathbf{E}} \in \mathbb{R}^{2n \times 2k}, \quad \mathbf{E} \in \mathbb{R}^{2n \times k}, \quad \mathbf{E}^T \mathbf{E} = \mathbf{I}_k, \quad \mathbf{E}^T \mathbb{J}_{2n} \mathbf{E} = \mathbf{0}_k, \quad (15)$$

- (iii) \mathbf{V} is symplectic and it holds $\mathbf{V}^T = \mathbf{V}^+$.

We remark that these matrices are characterized in [19] to be elements in $\text{Sp}(2k, \mathbb{R}^{2n}) \cap V_k(\mathbb{R}^{2n})$ where $\text{Sp}(2k, \mathbb{R}^{2n})$ is the symplectic Stiefel manifold and $V_k(\mathbb{R}^{2n})$ is the Stiefel manifold.

Proof. “(i) \implies (ii)”: Let $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ be a symplectic matrix with orthonormal columns. We rename the columns to $\mathbf{V} = [\mathbf{E} \quad \mathbf{F}]$ with $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_k]$ and $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_k]$. The symplecticity of the matrix written in terms of \mathbf{E} and \mathbf{F} reads

$$\mathbf{V}^T \mathbb{J}_{2n} \mathbf{V} = \begin{bmatrix} \mathbf{E}^T \mathbb{J}_{2n} \mathbf{E} & \mathbf{E}^T \mathbb{J}_{2n} \mathbf{F} \\ \mathbf{F}^T \mathbb{J}_{2n} \mathbf{E} & \mathbf{F}^T \mathbb{J}_{2n} \mathbf{F} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_k & \mathbf{I}_k \\ -\mathbf{I}_k & \mathbf{0}_k \end{bmatrix} \iff \begin{cases} \mathbf{E}^T \mathbb{J}_{2n} \mathbf{E} = \mathbf{F}^T \mathbb{J}_{2n} \mathbf{F} = \mathbf{0}_k, \\ -\mathbf{F}^T \mathbb{J}_{2n} \mathbf{E} = \mathbf{E}^T \mathbb{J}_{2n} \mathbf{F} = \mathbf{I}_k. \end{cases} \quad (16)$$

Expressed in terms of the columns $\mathbf{e}_i, \mathbf{f}_i$ of the matrices \mathbf{E}, \mathbf{F} , this condition reads for any $1 \leq i, j \leq k$

$$\mathbf{e}_i^\top \mathbb{J}_{2n} \mathbf{e}_j = 0, \quad \mathbf{e}_i^\top \mathbb{J}_{2n} \mathbf{f}_j = \delta_{ij}, \quad \mathbf{f}_i^\top \mathbb{J}_{2n} \mathbf{e}_j = -\delta_{ij}, \quad \mathbf{f}_i^\top \mathbb{J}_{2n} \mathbf{f}_j = 0,$$

and the orthonormality of the columns of \mathbf{V} implies

$$\mathbf{e}_i^\top \mathbf{e}_j = \delta_{ij}, \quad \mathbf{f}_i^\top \mathbf{f}_j = \delta_{ij}.$$

For a fixed $i \in \{1, \dots, k\}$, it is easy to show with $\mathbb{J}_{2n}^\top \mathbb{J}_{2n} = \mathbf{I}_{2n}$ that $\mathbb{J}_{2n} \mathbf{f}_i$ is of unit length

$$1 = \delta_{ii} = \mathbf{f}_i^\top \mathbf{f}_i = \mathbf{f}_i^\top \mathbb{J}_{2n}^\top \mathbb{J}_{2n} \mathbf{f}_i = \|\mathbb{J}_{2n} \mathbf{f}_i\|_2^2.$$

Thus, \mathbf{e}_i and $\mathbb{J}_{2n} \mathbf{f}_i$ are both unit vectors which fulfill $\mathbf{e}_i^\top \mathbb{J}_{2n} \mathbf{f}_i = \langle \mathbf{e}_i, \mathbb{J}_{2n} \mathbf{f}_i \rangle_{\mathbb{R}^{2n}} = 1$. By the Cauchy-Schwarz inequality, it holds $\langle \mathbf{e}_i, \mathbb{J}_{2n} \mathbf{f}_i \rangle = \|\mathbf{e}_i\| \|\mathbb{J}_{2n} \mathbf{f}_i\|$ if and only if the vectors are parallel. Thus, we infer $\mathbf{e}_i = \mathbb{J}_{2n} \mathbf{f}_i$ which is equivalent to $\mathbf{f}_i = \mathbb{J}_{2n}^\top \mathbf{e}_i$. Since this holds for all $i \in \{1, \dots, k\}$, we conclude that \mathbf{V} is of the form proposed in (15).

“(ii) \implies (iii)”: Let \mathbf{V} be of the form (15). Direct calculation yields

$$\mathbf{V}^\top \mathbb{J}_{2n} \mathbf{V} = \begin{bmatrix} \mathbf{E}^\top \\ \mathbf{E}^\top \mathbb{J}_{2n} \end{bmatrix} \mathbb{J}_{2n} \begin{bmatrix} \mathbf{E} & \mathbb{J}_{2n}^\top \mathbf{E} \end{bmatrix} = \begin{bmatrix} \mathbf{E}^\top \mathbb{J}_{2n} \mathbf{E} & \mathbf{E}^\top \mathbf{E} \\ -\mathbf{E}^\top \mathbf{E} & \mathbf{E}^\top \mathbb{J}_{2n} \mathbf{E} \end{bmatrix} \stackrel{(15)}{=} \begin{bmatrix} \mathbf{0}_k & \mathbf{I}_k \\ -\mathbf{I}_k & \mathbf{0}_k \end{bmatrix} = \mathbb{J}_{2k}$$

which shows that \mathbf{V} is symplectic. Thus, the symplectic inverse \mathbf{V}^+ exists. The following calculation shows that it equals the transposed \mathbf{V}^\top

$$\mathbf{V}^+ = \mathbb{J}_{2k}^\top \mathbf{V}^\top \mathbb{J}_{2n} = \mathbb{J}_{2k}^\top \begin{bmatrix} \mathbf{E}^\top \\ \mathbf{E}^\top \mathbb{J}_{2n} \end{bmatrix} \mathbb{J}_{2n} = \begin{bmatrix} -\mathbf{E}^\top \mathbb{J}_{2n} \\ \mathbf{E}^\top \end{bmatrix} \mathbb{J}_{2n} = \begin{bmatrix} -\mathbf{E}^\top \mathbb{J}_{2n} \mathbb{J}_{2n} \\ \mathbf{E}^\top \mathbb{J}_{2n} \end{bmatrix} = \begin{bmatrix} \mathbf{E}^\top \\ \mathbf{E}^\top \mathbb{J}_{2n} \end{bmatrix} = \mathbf{V}^\top.$$

“(iii) \implies (i)”: Let \mathbf{V} be symplectic with $\mathbf{V}^\top = \mathbf{V}^+$. Then, we know that \mathbf{V} has orthonormal columns since

$$\mathbf{I}_k = \mathbf{V}^+ \mathbf{V} = \mathbf{V}^\top \mathbf{V}.$$

□

Proposition 4 essentially limits the symplectic, orthonormal ROB \mathbf{V} to be of the form (15). Later in the current section, we see how to solve the PSD for ROB of this type. In Section 3.2, we are interested in ridding the ROB \mathbf{V} of this requirement to explore further solution methods of the PSD.

As mentioned before, the current solution procedures for the PSD almost exclusively restrict to the class of symplectic, orthonormal ROB introduced in Proposition 4. This includes the Cotangent Lift [19], the Complex SVD [19], partly the non-linear programming algorithm from [19] and the greedy procedure presented in [15]. We briefly review these approaches in the following proposition.

Proposition 5 (Symplectic, orthonormal basis generation). *The Cotangent Lift (CT), Complex SVD (cSVD) and the greedy procedure for symplectic basis generation all derive a symplectic and orthonormal ROB. The non-linear programming (NLP) admits a symplectic, orthonormal ROB if the coefficient matrix \mathbf{C} in [19, Algorithm 3] is symplectic and has orthonormal columns, i.e. it is of the form $\mathbf{C}_G = [\mathbf{G} \quad \mathbb{J}_{2k}^\top \mathbf{G}]$. The methods can be rewritten with $\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^\top \mathbf{E}]$, where the different formulations of \mathbf{E} read*

$$\mathbf{E}_{CT} = \begin{bmatrix} \Phi_{CT} \\ \mathbf{0}_{n \times k} \end{bmatrix}, \quad \mathbf{E}_{cSVD} = \begin{bmatrix} \Phi_{cSVD} \\ \Psi_{cSVD} \end{bmatrix}, \quad \mathbf{E}_{greedy} = [\mathbf{e}_1, \dots, \mathbf{e}_k], \quad \mathbf{E}_{NLP} = \widetilde{\mathbf{V}}_E \mathbf{G}$$

where

(i) $\Phi_{CT}, \Phi_{cSVD}, \Psi_{cSVD} \in \mathbb{R}^{n \times k}$ are matrices that fulfil

$$\Phi_{CT}^\top \Phi_{CT} = \mathbf{I}_k, \quad \Phi_{cSVD}^\top \Phi_{cSVD} + \Psi_{cSVD}^\top \Psi_{cSVD} = \mathbf{I}_k, \quad \Phi_{cSVD}^\top \Psi_{cSVD} = \Psi_{cSVD}^\top \Phi_{cSVD},$$

which is technically equivalent to $\mathbf{E}^\top \mathbf{E} = \mathbf{I}_k$ and $\mathbf{E}^\top \mathbb{J}_{2n} \mathbf{E} = \mathbf{0}_k$ (see (15)) for \mathbf{E}_{CT} and \mathbf{E}_{cSVD} ,

(ii) $\mathbf{e}_1, \dots, \mathbf{e}_k \in \mathbb{R}^{2n}$ are the basis vectors selected by the greedy algorithm,

(iii) $\widetilde{\mathbf{V}}_E \in \mathbb{R}^{2n \times 2k}$ is a ROB computed from CT or cSVD and $\mathbf{G} \in \mathbb{R}^{2k \times r}$, $r \leq k$, stems from the coefficient matrix $\mathbf{C}_G = [\mathbf{G} \quad \mathbb{J}_{2k}^\top \mathbf{G}]$ computed by the NLP algorithm.

Proof. All of the listed methods derive a symplectic ROB of the form $\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^\top \mathbf{E}]$ which satisfies (15). By Proposition 4, these ROB are each a symplectic, orthonormal ROB. □

In the following, we show that PSD Complex SVD is the solution of the PSD in the subset of symplectic, orthonormal ROB. This was partly shown in [19] which yet lacked the final step that, restricting to orthonormal, symplectic ROB, a solution of $\text{PSD}([\mathbf{X}_s \quad -\mathbb{J}_{2n}\mathbf{X}_s])$ solves $\text{PSD}(\mathbf{X}_s)$ and vice versa. This proves that the PSD Complex SVD is not only near optimal in this set but indeed optimal. Furthermore, the proof we show is alternative to the original and naturally motivates an alternative formulation of the PSD Complex SVD which we call the POD of \mathbf{Y}_s in the following. To begin with, we reproduce the definition of PSD Complex SVD from [19].

Definition 8 (PSD Complex SVD). *We define the complex snapshot matrix*

$$\mathbf{C}_s = [\mathbf{q}_1^s + i\mathbf{p}_1^s, \dots, \mathbf{q}_{n_s}^s + i\mathbf{p}_{n_s}^s] \in \mathbb{C}^{n \times n_s}, \quad \mathbf{x}_j^s = \begin{bmatrix} \mathbf{q}_j \\ \mathbf{p}_j \end{bmatrix} \text{ for all } 1 \leq j \leq n_s \quad (17)$$

which is derived with the imaginary unit i . The PSD Complex SVD is a basis generation technique that requires the auxiliary complex matrix $\mathbf{U}_{C_s} \in \mathbb{C}^{n \times k}$ to fulfil

$$\underset{\mathbf{U}_{C_s} \in \mathbb{C}^{n \times k}}{\text{minimize}} \|\mathbf{C}_s - \mathbf{U}_{C_s} (\mathbf{U}_{C_s})^* \mathbf{C}_s\|_{\text{F}}^2 \quad \text{subject to} \quad (\mathbf{U}_{C_s})^* \mathbf{U}_{C_s} = \mathbf{I}_k \quad (18)$$

and builds the actual ROB $\mathbf{V}_E \in \mathbb{R}^{2n \times 2k}$ with

$$\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^T \mathbf{E}], \quad \mathbf{E} = \begin{bmatrix} \text{Re}(\mathbf{U}_{C_s}) \\ \text{Im}(\mathbf{U}_{C_s}) \end{bmatrix}.$$

The solution of (18) is known to be based on the left-singular vectors of \mathbf{C}_s which can be explicitly computed with a complex version of the SVD.

We emphasize that we denote this basis generation procedure as PSD Complex SVD in the following to avoid confusions with the usual complex SVD algorithm.

Proposition 6 (Minimizing PSD in the set of symplectic, orthonormal ROB). *Given the snapshot matrix $\mathbf{X}_s \in \mathbb{R}^{2n \times n_s}$ we augment this with “rotated” snapshots to $\mathbf{Y}_s = [\mathbf{X}_s \quad \mathbb{J}_{2n}\mathbf{X}_s]$. We assume that $2k$ is such that we obtain a gap in the singular values of \mathbf{Y}_s , i.e. $\sigma_{2k}(\mathbf{Y}_s) > \sigma_{2k+1}(\mathbf{Y}_s)$. Then, minimizing the PSD in the set of symplectic, orthonormal ROB is equivalent to the following minimization problem*

$$\underset{\mathbf{V} \in \mathbb{R}^{2n \times 2k}}{\text{minimize}} \left\| (\mathbf{I}_{2n} - \mathbf{V}\mathbf{V}^T) [\mathbf{X}_s \quad \mathbb{J}_{2n}\mathbf{X}_s] \right\|_{\text{F}}^2 \quad \text{subject to} \quad \mathbf{V}^T \mathbf{V} = \mathbf{I}_{2k}. \quad (19)$$

Clearly, this is equivalent to the POD (12) applied to the snapshot matrix \mathbf{Y}_s . We, thus, call this procedure the POD of \mathbf{Y}_s in the following. A minimizer can be derived with the SVD as it is common for POD [4].

Proof. The proof proceeds in three steps: we show

- (i) that (\mathbf{u}, \mathbf{v}) is a pair of left- and right-singular vectors of \mathbf{Y}_s to the singular value σ if and only if $(\mathbb{J}_{2n}^T \mathbf{u}, \mathbb{J}_{2n_s}^T \mathbf{v})$ also is a pair of left- and right-singular vectors of \mathbf{Y}_s to the same singular value σ ,
- (ii) that a solution of the POD of \mathbf{Y}_s is a symplectic, orthonormal ROB, i.e. $\mathbf{V} = \mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^T \mathbf{E}]$,
- (iii) that the POD of \mathbf{Y}_s is equivalent to the PSD for symplectic, orthonormal ROB.

We start with the first step (i). Let (\mathbf{u}, \mathbf{v}) be a pair of left- and right-singular vectors of \mathbf{Y}_s to the singular value σ . We use that the left-singular (or right-singular) vectors of \mathbf{Y}_s are a set of orthonormal eigenvectors of $\mathbf{Y}_s \mathbf{Y}_s^T$ (or $\mathbf{Y}_s^T \mathbf{Y}_s$). To begin with, we compute

$$\begin{aligned} \mathbb{J}_{2n}^T \mathbf{Y}_s \mathbf{Y}_s^T \mathbb{J}_{2n} &= \mathbb{J}_{2n}^T (\mathbf{X}_s \mathbf{X}_s^T + \mathbb{J}_{2n} \mathbf{X}_s \mathbf{X}_s^T \mathbb{J}_{2n}^T) \mathbb{J}_{2n} = \mathbb{J}_{2n}^T \mathbf{X}_s \mathbf{X}_s^T \mathbb{J}_{2n} + \mathbf{X}_s \mathbf{X}_s^T = \mathbf{Y}_s \mathbf{Y}_s^T, \\ \mathbb{J}_{2n_s}^T \mathbf{Y}_s^T \mathbf{Y}_s \mathbb{J}_{2n_s} &= \mathbb{J}_{2n_s}^T \begin{bmatrix} \mathbf{X}_s^T \mathbf{X}_s & \mathbf{X}_s^T \mathbb{J}_{2n} \mathbf{X}_s \\ \mathbf{X}_s^T \mathbb{J}_{2n}^T \mathbf{X}_s & \mathbf{X}_s^T \mathbf{X}_s \end{bmatrix} \mathbb{J}_{2n_s} = \begin{bmatrix} \mathbf{X}_s^T \mathbf{X}_s & -\mathbf{X}_s^T \mathbb{J}_{2n}^T \mathbf{X}_s \\ -\mathbf{X}_s^T \mathbb{J}_{2n} \mathbf{X}_s & \mathbf{X}_s^T \mathbf{X}_s \end{bmatrix} = \mathbf{Y}_s^T \mathbf{Y}_s \end{aligned} \quad (20)$$

where we use $\mathbb{J}_{2n_s}^T = -\mathbb{J}_{2n_s}$. Thus, we can reformulate the eigenvalue problems of $\mathbf{Y}_s \mathbf{Y}_s^T$ and, respectively, $\mathbf{Y}_s^T \mathbf{Y}_s$ as

$$\begin{aligned} \sigma \mathbf{u} &= \mathbf{Y}_s \mathbf{Y}_s^T \mathbf{u} = \mathbb{J}_{2n} \mathbb{J}_{2n}^T \mathbf{Y}_s \mathbf{Y}_s^T \mathbb{J}_{2n} \mathbb{J}_{2n}^T \mathbf{u} & \xleftrightarrow{\mathbb{J}_{2n}^T \cdot} & \sigma \mathbb{J}_{2n}^T \mathbf{u} = \mathbb{J}_{2n}^T \mathbf{Y}_s \mathbf{Y}_s^T \mathbb{J}_{2n} \mathbb{J}_{2n}^T \mathbf{u} \stackrel{(20)}{=} \mathbf{Y}_s \mathbf{Y}_s^T \mathbb{J}_{2n}^T \mathbf{u} \\ \sigma \mathbf{v} &= \mathbf{Y}_s^T \mathbf{Y}_s \mathbf{v} = \mathbb{J}_{2n_s} \mathbb{J}_{2n_s}^T \mathbf{Y}_s^T \mathbf{Y}_s \mathbb{J}_{2n_s} \mathbb{J}_{2n_s}^T \mathbf{v} & \xleftrightarrow{\mathbb{J}_{2n_s}^T \cdot} & \sigma \mathbb{J}_{2n_s}^T \mathbf{v} = \mathbb{J}_{2n_s}^T \mathbf{Y}_s^T \mathbf{Y}_s \mathbb{J}_{2n_s} \mathbb{J}_{2n_s}^T \mathbf{v} \stackrel{(20)}{=} \mathbf{Y}_s^T \mathbf{Y}_s \mathbb{J}_{2n_s}^T \mathbf{v}. \end{aligned}$$

Thus, $(\mathbb{J}_{2n}^T \mathbf{u}, \mathbb{J}_{2n_s}^T \mathbf{v})$ is necessarily another pair of left- and right-singular vectors of \mathbf{Y}_s with the same singular value σ . We infer that the left-singular vectors \mathbf{u}_i , $1 \leq i \leq 2n$, ordered by the magnitude of the singular values in a descending order can be written as

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbb{J}_{2n}^T \mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbb{J}_{2n}^T \mathbf{u}_2 \quad \dots \quad \mathbf{u}_n \quad \mathbb{J}_{2n}^T \mathbf{u}_n] \in \mathbb{R}^{2n \times 2n}. \quad (21)$$

For the second step (ii), we remark that the solution of the POD is explicitly known to be any matrix which stacks in its columns $2k$ left-singular vectors of the snapshot matrix \mathbf{Y}_s with the highest singular value [4]. Due to the special structure (21) of the singular vectors for the snapshot matrix \mathbf{Y}_s , a minimizer of the POD of \mathbf{Y}_s necessarily adopts this structure. We are allowed to rearrange the order of the columns in this matrix and thus, the result of the POD of \mathbf{Y}_s can always be rearranged to the form

$$\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^T \mathbf{E}], \quad \mathbf{E} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_k], \quad \mathbb{J}_{2n}^T \mathbf{E} = [\mathbb{J}_{2n}^T \mathbf{u}_1 \quad \mathbb{J}_{2n}^T \mathbf{u}_2 \quad \dots \quad \mathbb{J}_{2n}^T \mathbf{u}_k].$$

Note that it automatically holds that $\mathbf{E}^T \mathbf{E} = \mathbf{I}_k$ and $\mathbf{E}^T (\mathbb{J}_{2n} \mathbf{E}) = \mathbf{0}_k$ since, in both products, we use the left-singular vectors from the columns of the matrix \mathbf{U} from (21) which is known to be orthogonal from properties of the SVD. Thus, (15) holds and we infer from Proposition 4 that the POD of \mathbf{Y}_s indeed is solved by a symplectic, orthonormal ROB.

For the final step (iii), we define the orthogonal projection operators

$$\mathbf{P}_{\mathbf{V}_E} = \mathbf{V}_E (\mathbf{V}_E)^T = \mathbf{E} \mathbf{E}^T + \mathbb{J}_{2n}^T \mathbf{E} \mathbf{E}^T \mathbb{J}_{2n}, \quad \mathbf{P}_{\mathbf{V}_E}^\perp = \mathbf{I}_{2n} - \mathbf{P}_{\mathbf{V}_E}.$$

Both are idempotent and symmetric, thus $(\mathbf{P}_{\mathbf{V}_E}^\perp)^T \mathbf{P}_{\mathbf{V}_E}^\perp = \mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{P}_{\mathbf{V}_E}^\perp = \mathbf{P}_{\mathbf{V}_E}^\perp$. Due to $\mathbb{J}_{2n} \mathbb{J}_{2n}^T = \mathbf{I}_{2n}$, it further holds

$$\mathbb{J}_{2n} (\mathbf{P}_{\mathbf{V}_E}^\perp)^T \mathbf{P}_{\mathbf{V}_E}^\perp \mathbb{J}_{2n}^T = \mathbb{J}_{2n} \mathbf{P}_{\mathbf{V}_E}^\perp \mathbb{J}_{2n}^T = \mathbb{J}_{2n} \mathbb{J}_{2n}^T - \mathbb{J}_{2n} \mathbf{E} \mathbf{E}^T \mathbb{J}_{2n}^T - \mathbb{J}_{2n} \mathbb{J}_{2n}^T \mathbf{E} \mathbf{E}^T \mathbb{J}_{2n} \mathbb{J}_{2n}^T = \mathbf{P}_{\mathbf{V}_E}^\perp = (\mathbf{P}_{\mathbf{V}_E}^\perp)^T \mathbf{P}_{\mathbf{V}_E}^\perp.$$

Thus, it follows

$$\left\| \mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{X}_s \right\|_{\mathbb{F}}^2 = \text{trace} \left(\mathbf{X}_s^T (\mathbf{P}_{\mathbf{V}_E}^\perp)^T \mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{X}_s \right) = \text{trace} \left(\mathbf{X}_s^T \mathbb{J}_{2n} (\mathbf{P}_{\mathbf{V}_E}^\perp)^T \mathbf{P}_{\mathbf{V}_E}^\perp \mathbb{J}_{2n}^T \mathbf{X}_s \right) = \left\| \mathbf{P}_{\mathbf{V}_E}^\perp \mathbb{J}_{2n}^T \mathbf{X}_s \right\|_{\mathbb{F}}^2$$

and with $\mathbf{Y}_s = [\mathbf{X}_s \quad \mathbb{J}_{2n}^T \mathbf{X}_s]$

$$2 \left\| \mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{X}_s \right\|_{\mathbb{F}}^2 = \left\| \mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{X}_s \right\|_{\mathbb{F}}^2 + \left\| \mathbf{P}_{\mathbf{V}_E}^\perp \mathbb{J}_{2n}^T \mathbf{X}_s \right\|_{\mathbb{F}}^2 = \left\| \mathbf{P}_{\mathbf{V}_E}^\perp [\mathbf{X}_s \quad \mathbb{J}_{2n}^T \mathbf{X}_s] \right\|_{\mathbb{F}}^2 = \left\| \mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{Y}_s \right\|_{\mathbb{F}}^2,$$

where we use in the last step that for two matrices $\mathbf{A} \in \mathbb{R}^{2n \times u}$, $\mathbf{B} \in \mathbb{R}^{2n \times v}$ for $u, v \in \mathbb{N}$, it holds $\|\mathbf{A}\|_{\mathbb{F}}^2 + \|\mathbf{B}\|_{\mathbb{F}}^2 = \|[\mathbf{A} \quad \mathbf{B}]\|_{\mathbb{F}}^2$ for the Frobenius norm $\|\bullet\|_{\mathbb{F}}$.

Since it is equivalent to minimize a function $f : \mathbb{R}^{2n \times 2k} \rightarrow \mathbb{R}$ or a multiple cf of it for any positive constant $c \in \mathbb{R}_{>0}$, minimizing $\|\mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{X}_s\|_{\mathbb{F}}^2$ is equivalent to minimizing $2 \|\mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{X}_s\|_{\mathbb{F}}^2 = \|\mathbf{P}_{\mathbf{V}_E}^\perp \mathbf{Y}_s\|_{\mathbb{F}}^2$. Additionally, for a ROB of the form $\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^T \mathbf{E}]$ the constraint of orthonormal columns is equivalent to the requirements in (15). Thus, to minimize the PSD in the class of symplectic, orthonormal ROBs is equivalent to the POD of \mathbf{Y}_s (19). \square

Remark 3. We remark that in the same fashion as the proof of step (iii) in Proposition 6, it can be shown that, restricting to symplectic, orthonormal ROBs, a solution of $\text{PSD}([\mathbf{X}_s \quad \mathbb{J}_{2n}^T \mathbf{X}_s])$ is a solution of $\text{PSD}(\mathbf{X}_s)$ and vice versa, which is one detail that was missing in [19] to show the optimality of PSD Complex SVD in the set of symplectic, orthonormal ROBs.

We next prove that PSD Complex SVD is equivalent to POD of \mathbf{Y}_s from (19) and thus, also minimizes the PSD in the set of symplectic, orthonormal bases. To this end, we repeat the optimality result from [19] and extend it with the results of the present paper.

Proposition 7 (Optimality of PSD Complex SVD). *Let $\mathbb{M}_2 \subset \mathbb{R}^{2n \times 2k}$ denote the set of symplectic bases with the structure $\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^T \mathbf{E}]$. The PSD Complex SVD solves $\text{PSD}([\mathbf{X}_s \quad -\mathbb{J}_{2n}^T \mathbf{X}_s])$ in \mathbb{M}_2 .*

Proof. See [19, Theorem 4.5]. \square

Proposition 8 (Equivalence of POD of \mathbf{Y}_s and PSD Complex SVD). *PSD Complex SVD is equivalent to the POD of \mathbf{Y}_s . Thus, PSD Complex SVD yields a minimizer of the PSD for symplectic, orthonormal ROBs.*

Proof. By Proposition 7, PSD Complex SVD minimizes (19) in the set \mathbb{M}_2 of symplectic bases with the structure $\mathbf{V}_E = [\mathbf{E} \quad \mathbb{J}_{2n}^T \mathbf{E}]$. Thus, (16) holds with $\mathbf{F} = \mathbb{J}_{2n}^T \mathbf{E}$ which is equivalent to the conditions on \mathbf{E} required in (15). By Proposition 4, we infer that \mathbb{M}_2 equals the set of symplectic, orthonormal bases.

Furthermore, we can show that, in the set \mathbb{M}_2 , a solution of $\text{PSD}([\mathbf{X}_s \quad -\mathbb{J}_{2n}^T \mathbf{X}_s])$ is a solution of $\text{PSD}(\mathbf{X}_s)$ and vice versa (see Remark 3). Thus, PSD Complex SVD minimizes the PSD for the snapshot matrix \mathbf{X}_s in the set of orthonormal, symplectic matrices and PSD Complex SVD and the POD of \mathbf{Y}_s solve the same minimization problem. \square

We emphasize that the computation of a minimizer of (19) via PSD Complex SVD requires less memory storage than the computation via POD of \mathbf{Y}_s . The reason is that the complex formulation uses the complex snapshot matrix $\mathbf{C}_s \in \mathbb{C}^{n \times n_s}$ which equals $2 \cdot n \cdot n_s$ floating point numbers while the solution with the POD of \mathbf{Y}_s method artificially enlarges the snapshot matrix to $\mathbf{Y}_s \in \mathbb{R}^{2n \times 2n_s}$ which are $4 \cdot n \cdot n_s$ floating point numbers. Still, the POD of \mathbf{Y}_s might be computationally more efficient since it is a purely real formulation and thereby does not require complex arithmetic operations.

3.2 Symplectic, non-orthonormal basis generation

In the next step, we want to give an idea how to leave the class of symplectic, orthonormal ROB. We call a basis generation technique symplectic, non-orthonormal if it is able to compute a symplectic, non-orthonormal basis.

In Proposition 5, we briefly showed that most existing symplectic basis generation techniques generate a symplectic, orthonormal ROB. The only exception is the NLP algorithm suggested in [19]. It is able to compute a non-orthonormal, symplectic ROB. The algorithm is based on a given initial guess $\mathbf{V}_0 \in \mathbb{R}^{2n \times 2k}$ which is a symplectic ROB e.g. computed with PSD Cotangent Lift or PSD Complex SVD. Nonlinear programming is used to leave the class of symplectic, orthonormal ROB and derive an optimized symplectic ROB $\mathbf{V} = \mathbf{V}_0 \mathbf{C}$ with the symplectic coefficient matrix $\mathbf{C} \in \mathbb{R}^{2k \times 2r}$ for some $r \leq k$. Since this procedure searches a solution spanned by the columns of \mathbf{V}_0 , it is not suited to compute a global optimum of the PSD which we are interested in the scope of this paper.

In the following, we present a new basis generation technique that is based on an SVD-like decomposition for matrices $\mathbf{B} \in \mathbb{R}^{2n \times m}$ presented in [21]. To this end, we introduce this decomposition in the following.

Proposition 9 (SVD-like decomposition [21]). *Any real matrix $\mathbf{B} \in \mathbb{R}^{2n \times m}$ can be decomposed as the product of a symplectic matrix $\mathbf{S} \in \mathbb{R}^{2n \times 2n}$, a sparse and potentially non-diagonal matrix $\mathbf{D} \in \mathbb{R}^{2n \times m}$ and an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{m \times m}$ with*

$$\mathbf{B} = \mathbf{S} \mathbf{D} \mathbf{Q}, \quad \mathbf{D} = \begin{pmatrix} p & q & p & m-2p-q \\ \Sigma_s & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Sigma_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{matrix} p \\ q \\ n-p-q \\ p \\ q \\ n-p-q \end{matrix}, \quad \Sigma_s = \text{diag}(\sigma_1^s, \dots, \sigma_p^s) \in \mathbb{R}^{p \times p}, \quad \sigma_i^s > 0 \quad \text{for } 1 \leq i \leq p. \quad (22)$$

with $p, q \in \mathbb{N}$ and $\text{rank}(\mathbf{B}) = 2p + q$ and where we indicate the block row and column dimensions in \mathbf{D} by small letters. The diagonal entries σ_i^s , $1 \leq i \leq p$, of the matrix Σ_s are related to the pairs of purely imaginary eigenvalues $\lambda_j(\mathbf{M}), \lambda_{p+j}(\mathbf{M}) \in \mathbb{C}$ of $\mathbf{M} = \mathbf{B}^T \mathbb{J}_{2n} \mathbf{B} \in \mathbb{R}^{m \times m}$ with

$$\lambda_j(\mathbf{M}) = -(\sigma_j^s)^2 i, \quad \lambda_{p+j}(\mathbf{M}) = (\sigma_j^s)^2 i, \quad 1 \leq j \leq p.$$

Remark 4 (Singular values). *We call the diagonal entries σ_i^s , $1 \leq i \leq p$, of the matrix Σ_s from Proposition 9 in the following the symplectic singular values. The reason is the following analogy to the classical SVD.*

The classical SVD decomposes $\mathbf{B} \in \mathbb{R}^{2n \times m}$ as $\mathbf{B} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ where $\mathbf{U} \in \mathbb{R}^{2n \times 2n}$, $\mathbf{V} \in \mathbb{R}^{m \times m}$ are each orthogonal matrices and $\mathbf{\Sigma} \in \mathbb{R}^{2n \times m}$ is a diagonal matrix with the singular values σ_i on its diagonal $\text{diag}(\mathbf{\Sigma}) = [\sigma_1, \dots, \sigma_r, 0, \dots, 0] \in \mathbb{R}^{\min(2n, m)}$, $r = \text{rank}(\mathbf{B})$. The singular values are linked to the real eigenvalues of $\mathbf{N} = \mathbf{B}^T \mathbf{B}$ with $\lambda_i(\mathbf{N}) = \sigma_i^2$. Furthermore, for the SVD, it holds due to the orthogonality of \mathbf{U} and \mathbf{V} , respectively, $\mathbf{B}^T \mathbf{B} = \mathbf{V}^T \mathbf{\Sigma}^2 \mathbf{V}$ and $\mathbf{B} \mathbf{B}^T = \mathbf{U}^T \mathbf{\Sigma}^2 \mathbf{U}$.

A similar relation can be derived for an SVD-like decomposition from Proposition 9. Due to the structure of the decomposition (22) and the symplecticity of \mathbf{S} , it holds

$$\mathbf{B}^T \mathbb{J}_{2n} \mathbf{B} = \mathbf{Q}^T \mathbf{D}^T \overbrace{\mathbf{S}^T \mathbb{J}_{2n} \mathbf{S}}{=\mathbb{J}_{2n}} \mathbf{D} \mathbf{Q} \quad \mathbf{D}^T \mathbb{J}_{2n} \mathbf{D} = \begin{pmatrix} p & q & p & m-2p-q \\ \mathbf{0} & \mathbf{0} & \Sigma_s^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\Sigma_s^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{matrix} p \\ q \\ p \\ m-2p-q \end{matrix}. \quad (23)$$

This analogy is why we call the diagonal entries σ_i^s , $1 \leq i \leq p$, of the matrix Σ_s symplectic singular values.

The idea for the basis generation now is to select $k \in \mathbb{N}$ pairs of columns of \mathbf{S} in order to compute a symplectic ROB. The selection should be based on the importance of these pairs which we characterize by the following proposition by linking the Frobenius norm of a matrix with the symplectic singular values.

Proposition 10. *Let $\mathbf{B} \in \mathbb{R}^{2n \times m}$ with an SVD-like decomposition $\mathbf{B} = \mathbf{S} \mathbf{D} \mathbf{Q}$ with $p, q \in \mathbb{N}$ from Proposition 9. The Frobenius norm of \mathbf{B} can be rewritten as*

$$\|\mathbf{B}\|_{\text{F}}^2 = \sum_{i=1}^{p+q} (w_i^s)^2, \quad w_i^s = \begin{cases} \sigma_i^s \sqrt{\|\mathbf{s}_i\|_2^2 + \|\mathbf{s}_{n+i}\|_2^2}, & 1 \leq i \leq p, \\ \|\mathbf{s}_i\|_2, & p+1 \leq i \leq p+q \end{cases} \quad (24)$$

where $\mathbf{s}_i \in \mathbb{R}^{2n}$ is the i -th column of \mathbf{S} for $1 \leq i \leq 2n$. In the following, we refer to each w_i^s as the weighted symplectic singular value.

Proof. We insert the SVD-like decomposition $\mathbf{B} = \mathbf{S}\mathbf{D}\mathbf{Q}$ and use the orthogonality of \mathbf{Q} to reformulate

$$\begin{aligned}\|\mathbf{B}\|_{\mathbb{F}}^2 &= \|\mathbf{S}\mathbf{D}\mathbf{Q}\|_{\mathbb{F}}^2 = \|\mathbf{S}\mathbf{D}\|_{\mathbb{F}}^2 = \text{trace}(\mathbf{D}^{\top}\mathbf{S}^{\top}\mathbf{S}\mathbf{D}) = \sum_{i=1}^p (\sigma_i^s)^2 \mathbf{s}_i^{\top} \mathbf{s}_i + \sum_{i=1}^q (\sigma_i^s)^2 \mathbf{s}_{n+i}^{\top} \mathbf{s}_{n+i} + \sum_{i=1}^q \mathbf{s}_{p+i}^{\top} \mathbf{s}_{p+i} \\ &= \sum_{i=1}^p (\sigma_i^s)^2 (\|\mathbf{s}_i\|_2^2 + \|\mathbf{s}_{n+i}\|_2^2) + \sum_{i=1}^q \|\mathbf{s}_{p+i}\|_2^2\end{aligned}$$

which is equivalent to (24). \square

It proves true in the following Proposition 11 that we can delete single addends w_i^s in (24) with the symplectic projection used in the PSD if we include the corresponding pair of columns in the ROB. This will be our selection criterion in the new basis generation technique that we denote PSD SVD-like decomposition.

Definition 9 (PSD SVD-like decomposition). *We compute an SVD-like decomposition (22) as $\mathbf{X}_s = \mathbf{S}\mathbf{D}\mathbf{Q}$ of the snapshot matrix $\mathbf{X}_s \in \mathbb{R}^{2n \times n_s}$ and define $p, q \in \mathbb{N}$ as in Proposition 9. In order to compute a ROB \mathbf{V} with $2k$ columns, find the k indices $i \in \mathcal{I}_{\text{PSD}} = \{i_1, \dots, i_k\} \subset \{1, \dots, p+q\}$ which have large contributions w_i^s in (24) with*

$$\mathcal{I}_{\text{PSD}} = \underset{\substack{\mathcal{I} \subset \{1, \dots, p+q\} \\ |\mathcal{I}|=k}}{\text{argmax}} \left(\sum_{i \in \mathcal{I}} (w_i^s)^2 \right). \quad (25)$$

To construct the ROB, we choose the k pairs of columns $\mathbf{s}_i \in \mathbb{R}^{2n}$ from \mathbf{S} corresponding to the selected indices \mathcal{I}_{PSD} such that

$$\mathbf{V} = [\mathbf{s}_{i_1}, \dots, \mathbf{s}_{i_k}, \mathbf{s}_{n+i_1}, \dots, \mathbf{s}_{n+i_k}] \in \mathbb{R}^{2n \times 2k}.$$

The special choice of the ROB is motivated by the following theoretical result which is very analogous to the results known for the classical POD in the framework of orthogonal projections.

Proposition 11 (Projection error by neglected weighted symplectic singular values). *Let $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ be a ROB constructed with the procedure described in Definition 9 to the index set $\mathcal{I}_{\text{PSD}} \subset \{1, \dots, p+q\}$ with $p, q \in \mathbb{N}$ from Proposition 9. The PSD functional can be calculated by*

$$\|(\mathbf{I}_{2n} - \mathbf{V}\mathbf{V}^+) \mathbf{X}_s\|_{\mathbb{F}}^2 = \sum_{i \in \{1, \dots, p+q\} \setminus \mathcal{I}_{\text{PSD}}} (w_i^s)^2, \quad (26)$$

which is the cumulative sum of the squares of the neglected weighted symplectic singular values.

Proof. Let $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ be a ROB constructed from an SVD-like decomposition $\mathbf{X}_s = \mathbf{S}\mathbf{D}\mathbf{Q}$ of the snapshot matrix $\mathbf{X}_s \in \mathbb{R}^{2n \times 2k}$ with the procedure described in Definition 9. Let $p, q \in \mathbb{N}$ be defined as in Proposition 9 and $\mathcal{I}_{\text{PSD}} = \{i_1, \dots, i_k\} \subset \{1, \dots, p+q\}$ be the set of indices selected with (25).

For the proof, we introduce a slightly different notation of the ROB \mathbf{V} . The selection of the columns \mathbf{s}_i of \mathbf{S} is denoted with the selection matrix $\mathbf{I}_{\mathcal{I}_{\text{PSD}}} \in \mathbb{R}^{2n \times 2k}$ based on

$$(\mathbf{I}_{\mathcal{I}_{\text{PSD}}})_{\alpha, \beta} = \begin{cases} 1, & \alpha = i_{\beta} \in \mathcal{I}_{\text{PSD}} \\ 0, & \text{else} \end{cases} \quad \text{for} \quad \begin{cases} 1 \leq \alpha \leq 2n, \\ 1 \leq \beta \leq k, \end{cases} \quad \mathbf{I}_{\mathcal{I}_{\text{PSD}}} = [\mathbf{I}_{\mathcal{I}_{\text{PSD}}}, \mathbb{J}_{2n}^{\top} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}]$$

which allows us to write the ROB as the matrix–matrix product $\mathbf{V} = \mathbf{S}\mathbf{I}_{\mathcal{I}_{\text{PSD}}}$. Furthermore, we can select the neglected entries with $\mathbf{I}_{2n} - \mathbf{I}_{\mathcal{I}_{\text{PSD}}} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top}$.

We insert the SVD-like decomposition and the representation of the ROB introduced in the previous paragraph in the PSD which reads

$$\|(\mathbf{I}_{2n} - \mathbf{V}\mathbf{V}^+) \mathbf{X}_s\|_{\mathbb{F}}^2 = \left\| (\mathbf{I}_{2n} - \mathbf{S}\mathbf{I}_{\mathcal{I}_{\text{PSD}}} \mathbb{J}_{2k}^{\top} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \mathbf{S}^{\top} \mathbb{J}_{2n}) \mathbf{S}\mathbf{D}\mathbf{Q} \right\|_{\mathbb{F}}^2 = \left\| \mathbf{S} (\mathbf{I}_{2n} - \mathbf{I}_{\mathcal{I}_{\text{PSD}}} \mathbb{J}_{2k}^{\top} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \overbrace{\mathbf{S}^{\top} \mathbb{J}_{2n} \mathbf{S}}^{=\mathbb{J}_{2n}}) \mathbf{D} \right\|_{\mathbb{F}}^2$$

where we use the orthogonality of \mathbf{Q} and the symplecticity of \mathbf{S} in the last step. We can reformulate the product of Poisson matrices and the selection matrix as

$$\mathbb{J}_{2k}^{\top} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \mathbb{J}_{2n} = \mathbb{J}_{2k}^{\top} \begin{bmatrix} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \\ \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \mathbb{J}_{2n} \end{bmatrix} \mathbb{J}_{2n} = \begin{bmatrix} \mathbf{0}_k & -\mathbf{I}_k \\ \mathbf{I}_k & \mathbf{0}_k \end{bmatrix} \begin{bmatrix} \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \mathbb{J}_{2n} \\ -\mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top} \end{bmatrix} = \mathbf{I}_{\mathcal{I}_{\text{PSD}}}^{\top}.$$

Thus, we can further reformulate the PSD as

$$\|(\mathbf{I}_{2n} - \mathbf{V}\mathbf{V}^+)\mathbf{X}_s\|_F^2 = \left\| \mathbf{S} \left(\mathbf{I}_{2n} - \mathbf{I}_{\mathcal{I}_{\text{PSD}}^{2k}} \left(\mathbf{I}_{\mathcal{I}_{\text{PSD}}^{2k}} \right)^\top \right) \mathbf{D} \right\|_F^2 = \sum_{i \in \{1, \dots, p+q\} \setminus \mathcal{I}_{\text{PSD}}} (w_i^s)^2$$

where w_i^s are the weighted symplectic singular values from (24). In the last step, we use that the resultant diagonal matrix in the braces sets all rows of \mathbf{D} with indices $i, n+i$ to zero for $i \in \mathcal{I}_{\text{PSD}}$. Thus, the last step can be concluded analogously to the proof of Proposition 10. \square

A direct consequence of Proposition 11 is that the decay of the PSD functional is proportional to the decay of the sum over the neglected weighted symplectic singular values w_i^s from (24). In the numerical example Section 4.2.1, we observe an exponential decrease of this quantities which induces an exponential decay of the PSD functional.

Remark 5 (Computation of the SVD-like decomposition). *To compute an SVD-like decomposition (22) of \mathbf{B} , several approaches exist. The original paper [21] derives a decomposition based on the product $\mathbf{B}^\top \mathbb{J}_{2n} \mathbf{B}$ which is not good for a numerical computation since errors can arise from cancellation. In [20], an implicit version is presented that does not require the computation of the full product $\mathbf{B}^\top \mathbb{J}_{2n} \mathbf{B}$ but derives the decomposition implicitly by transforming \mathbf{B} . Furthermore, [1] introduces an iterative approach to compute an SVD-like decomposition which computes parts of an SVD-like decomposition with a block-power iterative method. In the present case, we use the implicit approach [20].*

3.3 Interplay of non-orthonormal and orthonormal ROBs

We give further results on the interplay of non-orthonormal and orthonormal ROBs. The fundamental statement in the current section is the Orthogonal SR decomposition [7, 21].

Proposition 12 (Orthogonal SR decomposition). *For each matrix $\mathbf{B} \in \mathbb{R}^{2n \times m}$ with $m \leq n$, there exists a symplectic, orthogonal matrix $\mathbf{S} \in \mathbb{R}^{2n \times 2n}$, an upper triangular matrix $\mathbf{R}_{11} \in \mathbb{R}^{m \times m}$ and a strictly upper triangular matrix $\mathbf{R}_{21} \in \mathbb{R}^{m \times m}$ such that*

$$\mathbf{B} = \mathbf{S} \begin{bmatrix} \mathbf{R}_{11} \\ \mathbf{0}_{(n-m) \times m} \\ \mathbf{R}_{21} \\ \mathbf{0}_{(n-m) \times m} \end{bmatrix} = [\mathbf{S}_m \quad \mathbb{J}_{2n}^T \mathbf{S}_m] \begin{bmatrix} \mathbf{R}_{11} \\ \mathbf{R}_{21} \end{bmatrix}, \quad \begin{aligned} \mathbf{S}_i &= [\mathbf{s}_1, \dots, \mathbf{s}_i], \quad 1 \leq i \leq n, \\ \mathbf{S} &= [\mathbf{s}_1, \dots, \mathbf{s}_n, \mathbb{J}_{2n}^T \mathbf{s}_1, \dots, \mathbb{J}_{2n}^T \mathbf{s}_n]. \end{aligned}$$

We remark that a similar result can be derived for the case $m > n$ [21] but it is not introduced since we do not need it in the following.

Proof. Let $\mathbf{B} \in \mathbb{R}^{2n \times m}$ with $m \leq n$. We consider the QR decomposition

$$\mathbf{B} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{(2n-m) \times m} \end{bmatrix}$$

where $\mathbf{Q} \in \mathbb{R}^{2n \times 2n}$ is an orthogonal matrix and $\mathbf{R} \in \mathbb{R}^{2n \times m}$ is upper triangular. The original Orthogonal SR decomposition [7, Corollary 4.5.] for the square matrix states that we can decompose $\mathbf{Q} \in \mathbb{R}^{2n \times 2n}$ as a symplectic, orthogonal matrix $\mathbf{S} \in \mathbb{R}^{2n \times 2n}$, an upper triangular matrix $\tilde{\mathbf{R}}_{11} \in \mathbb{R}^{n \times n}$, a strictly upper triangular matrix $\tilde{\mathbf{R}}_{21} \in \mathbb{R}^{n \times n}$ and two (possibly) full matrices $\tilde{\mathbf{R}}_{12}, \tilde{\mathbf{R}}_{22} \in \mathbb{R}^{n \times n}$

$$\mathbf{Q} = \mathbf{S} \begin{bmatrix} \tilde{\mathbf{R}}_{11} & \tilde{\mathbf{R}}_{12} \\ \tilde{\mathbf{R}}_{21} & \tilde{\mathbf{R}}_{22} \end{bmatrix} \quad \text{and thus} \quad \mathbf{B} = \mathbf{S} \begin{bmatrix} \tilde{\mathbf{R}}_{11} & \tilde{\mathbf{R}}_{12} \\ \tilde{\mathbf{R}}_{21} & \tilde{\mathbf{R}}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{(2n-m) \times m} \end{bmatrix} = \mathbf{S} \begin{bmatrix} \tilde{\mathbf{R}}_{11} \\ \tilde{\mathbf{R}}_{21} \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{(n-m) \times m} \end{bmatrix}.$$

Since \mathbf{R} is upper triangular, it does preserve the (strictly) upper triangular pattern in $\tilde{\mathbf{R}}_{11}$ and $\tilde{\mathbf{R}}_{21}$ and we obtain the (strictly) upper triangular matrices $\mathbf{R}_{11}, \mathbf{R}_{21} \in \mathbb{R}^{m \times m}$ from

$$\begin{bmatrix} \mathbf{R}_{11} \\ \mathbf{0}_{(n-m) \times m} \\ \mathbf{R}_{21} \\ \mathbf{0}_{(n-m) \times m} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{R}}_{11} \\ \tilde{\mathbf{R}}_{21} \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{(n-m) \times m} \end{bmatrix}.$$

\square

Based on the Orthogonal SR decomposition, the following two propositions prove bounds for the projection errors of PSD which allows an estimate for the quality of the respective method. In both cases we require the basis size to satisfy $k \leq n$ or $2k \leq n$, respectively. This restriction is not limiting in the context of symplectic MOR as in all application cases $k \ll n$.

Proposition 13. Let $\mathbf{V} \in \mathbb{R}^{2n \times k}$ be a minimizer of POD with $k \leq n$ basis vectors and $\mathbf{V}_E \in \mathbb{R}^{2n \times 2k}$ be a minimizer of the PSD in the class of orthonormal, symplectic matrices with $2k$ basis vectors. Then, the orthogonal projection errors of \mathbf{V}_E and \mathbf{V} satisfy

$$\left\| (\mathbf{I}_{2n} - \mathbf{V}_E \mathbf{V}_E^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \leq \left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2.$$

Proof. The Orthogonal SR decomposition (see Proposition 12) guarantees that a symplectic, orthogonal matrix $\mathbf{S} \in \mathbb{R}^{2n \times 2k}$ and $\mathbf{R} \in \mathbb{R}^{2k \times k}$ exist with $\mathbf{V} = \mathbf{S} \mathbf{R}$. Since both matrices \mathbf{V} and \mathbf{S} are orthogonal and $\text{img}(\mathbf{V}) \subset \text{img}(\mathbf{S})$, we can show that \mathbf{S} yields a lower projection error than \mathbf{V} with

$$\begin{aligned} \left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2 &= \left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^\top) (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2 = \sum_{i=1}^{n_s} \left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^\top) (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^\top) \mathbf{x}_i^s \right\|_2^2 \\ &\leq \underbrace{\left\| \mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^\top \right\|_2^2}_{\leq 1} \sum_{i=1}^{n_s} \left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^\top) \mathbf{x}_i^s \right\|_2^2 \leq \left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \end{aligned}$$

Let $\mathbf{V}_E \in \mathbb{R}^{2n \times 2k}$ be a minimizer of the PSD in the class of symplectic, orthonormal ROBs. By definition of \mathbf{V}_E , it yields a lower projection error than \mathbf{S} . Since both ROBs are symplectic, orthonormal, we can exchange the symplectic inverse with the transposition (see Proposition 4, (iii)). This proves the assertion with

$$\left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \geq \left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \geq \left\| (\mathbf{I}_{2n} - \mathbf{V}_E \mathbf{V}_E^\top) \mathbf{X}_s \right\|_{\mathbb{F}}^2.$$

□

Proposition 13 proves that we require at most twice the number of basis vectors to generate a symplectic, orthonormal basis with an orthogonal projection error at least as small as the one of the classical POD. An analogous result can be derived in the framework of a symplectic projection which is proven in the following proposition.

Proposition 14. Assume there exists a minimizer $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ of the general PSD for a basis size $2k \leq n$ with potentially non-orthonormal columns. Let $\mathbf{V}_E \in \mathbb{R}^{2n \times 4k}$ be a minimizer of the PSD in the class of symplectic, orthogonal bases of size $4k$. Then, we know that the symplectic projection error of \mathbf{V}_E is less than or equal to the one of \mathbf{V} , i.e.

$$\left\| (\mathbf{I}_{2n} - \mathbf{V}_E \mathbf{V}_E^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \leq \left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2.$$

Proof. Let $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ be a minimizer of PSD with $2k \leq n$. By Proposition 12, we can determine a symplectic, orthogonal matrix $\mathbf{S} \in \mathbb{R}^{2n \times 4k}$ and $\mathbf{R} \in \mathbb{R}^{4k \times 2k}$ with $\mathbf{V} = \mathbf{S} \mathbf{R}$. Similar to the proof of Proposition 13, we can bound the projection errors. We require the identity

$$(\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^+) (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^+) = \mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^+ - \mathbf{V} \mathbf{V}^+ + \underbrace{\mathbf{S} \mathbf{S}^+ \mathbf{S} \mathbf{R}}_{=\mathbf{V}} \underbrace{\mathbb{J}_{2k}^\top \mathbf{R}^\top \mathbf{S}^\top \mathbb{J}_{2n}}_{=\mathbf{V}^+} = \mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^+.$$

With this identity, we proceed analogously to the proof of Proposition 13 and derive for a minimizer $\mathbf{V}_E \in \mathbb{R}^{2n \times 4k}$ of PSD in the class of symplectic, orthonormal ROBs

$$\begin{aligned} \left\| (\mathbf{I}_{2n} - \mathbf{V}_E \mathbf{V}_E^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2 &\leq \left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2 = \left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^+) (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \\ &\leq \underbrace{\left\| (\mathbf{I}_{2n} - \mathbf{S} \mathbf{S}^+) \right\|_2^2}_{\leq 1} \left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2 \leq \left\| (\mathbf{I}_{2n} - \mathbf{V} \mathbf{V}^+) \mathbf{X}_s \right\|_{\mathbb{F}}^2. \end{aligned}$$

□

Proposition 14 proves that we require at most twice the number of basis vectors to generate a symplectic, orthonormal basis with a symplectic projection error at least as small as the one of a (potentially non-orthonormal) minimizer of PSD.

4 Numerical results

The numerical experiments in the present paper are based on a two-dimensional plane strain linear elasticity model which is described by a Lamé–Navier equation

$$\rho_0 \frac{\partial^2}{\partial t^2} \mathbf{u}(\boldsymbol{\xi}, t, \boldsymbol{\mu}) - \mu_L \Delta_{\boldsymbol{\xi}} \mathbf{u}(\boldsymbol{\xi}, t, \boldsymbol{\mu}) + (\lambda_L + \mu_L) \nabla_{\boldsymbol{\xi}} (\operatorname{div}_{\boldsymbol{\xi}} (\mathbf{u}(\boldsymbol{\xi}, t, \boldsymbol{\mu}))) = \rho_0 \mathbf{g}(\boldsymbol{\xi}, t).$$

for $\boldsymbol{\xi} \in \Omega \subset \mathbb{R}^2$ and $t \in [t_0, t_{\text{end}}]$ with the density $\rho_0 \in \mathbb{R}_{>0}$, the Lamé constants $\boldsymbol{\mu} = (\lambda_L, \mu_L) \in \mathbb{R}_{>0}^2$, the external force $\mathbf{g} : \Omega \times [t_0, t_{\text{end}}] \rightarrow \mathbb{R}^2$ and Dirichlet boundary conditions on $\Gamma_u \subset \Gamma := \partial\Omega$ and Neumann boundary conditions on $\Gamma_t \subset \Gamma$. We apply non-dimensionalization (e.g. [13, Chapter 4.1]), apply the Finite Element Method (FEM) with first-order Lagrangian elements on a triangular mesh and rewrite the system as first-order system to arrive at the parametric linear system (9) with

$$\mathbf{x}(t, \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{q}(t, \boldsymbol{\mu}) \\ \mathbf{p}(t, \boldsymbol{\mu}) \end{bmatrix}, \quad \mathbf{H}(\boldsymbol{\mu}) = \begin{bmatrix} \mathbf{K}(\boldsymbol{\mu}) & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{M}^{-1} \end{bmatrix}, \quad \mathbf{h}(t) = \begin{bmatrix} -\mathbf{f}(t) \\ \mathbf{0}_{n \times 1} \end{bmatrix} \quad (27)$$

where $\mathbf{q}(t, \boldsymbol{\mu}) \in \mathbb{R}^n$ is the vector of displacement DOFs, $\mathbf{p}(t, \boldsymbol{\mu}) \in \mathbb{R}^n$ is the vector of linear momentum DOFs, $\mathbf{K}(\boldsymbol{\mu}) \in \mathbb{R}^{n \times n}$ is the stiffness matrix, $\mathbf{M}^{-1} \in \mathbb{R}^{n \times n}$ is the inverse of the mass matrix and $\mathbf{f}(t, \boldsymbol{\mu})$ is the vector of external forces.

We remark that a Hamiltonian formulation with the velocity DOFs $\mathbf{v}(t) = \frac{d}{dt} \mathbf{x}(t) \in \mathbb{R}^n$ instead of the linear momentum DOFs $\mathbf{p}(t)$ is possible if a non-canonical symplectic structure is used. Nevertheless, in [19, Remark 3.8.] it is suggested to switch to a formulation with a canonical symplectic structure for the MOR of Hamiltonian systems.

In order to solve the system (27) numerically with a time-discrete approximation $\mathbf{x}_i(\boldsymbol{\mu}) \approx \mathbf{x}(t_i, \boldsymbol{\mu})$ for each of $n_t \in \mathbb{N}$ time steps $t_i \in [t_0, t_{\text{end}}]$, $1 \leq i \leq n_t$, a numerical integrator is required. The preservation of the symplectic structure in the time-discrete system requires a so-called symplectic integrator [11, 5]. In the context of our work, the implicit midpoint scheme is used in all cases.

Remark 6 (Modified Hamiltonian). *We remark that, even though the symplectic structure is preserved by symplectic integrators, the Hamiltonian may be modified in the time-discrete system compared to the original Hamiltonian. In the case of a quadratic Hamiltonian (see Corollary 1) and a symplectic Runge-Kutta integrator, the modified Hamiltonian equals the original Hamiltonian since these integrators preserve quadratic first integrals. For further details, we refer to [11, Chapter IX.] or [14, Section 5.1.2 and 5.2].*

The model parameters are the first and second Lamé constants with $\boldsymbol{\mu} = (\lambda_L, \mu_L) \in \mathcal{P} = [35 \cdot 10^9, 125 \cdot 10^9] \text{ N/m}^2 \times [35 \cdot 10^9, 83 \cdot 10^9] \text{ N/m}^2$ which varies between cast iron and steel with approx. 12% chromium [17, App. E 1 Table 1]. The density is set to $\rho_0 = 7856 \text{ kg/m}^3$. The non-dimensionalization constants are set to $\lambda_L^c = \mu_L^c = 81 \cdot 10^9 \text{ N/m}^2$, $\xi^c = 1 \text{ m}$, $g^c = 9.81 \text{ m/s}^2$. The geometry is a simple cantilever beam clamped on the left side with a force applied to the right boundary. The time interval is chosen to be $t \in [t_0, t_{\text{end}}]$ with $t_0 = 0 \text{ s}$ and $t_{\text{end}} = 7.2 \cdot 10^{-2} \text{ s}$ which is one oscillation of the beam. For the numerical integration $n_t = 151$ time steps are used.

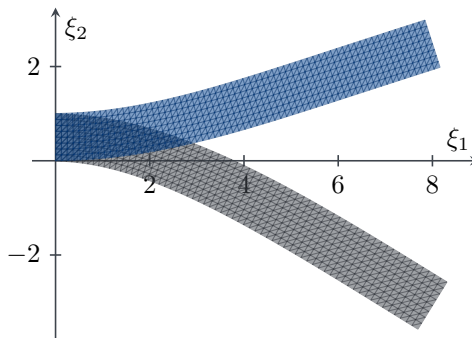


Figure 1: an exaggerated illustration of the displacements $\mathbf{q}(t, \boldsymbol{\mu})$ of the non-autonomous beam model (a) at the time with the maximum displacement (gray) and (b) at the final time (blue).

The symplectic MOR techniques examined are PSD Complex SVD (Definition 8), the greedy procedure [15] and the newly introduced PSD SVD-like decomposition (Definition 9). The MOR techniques that do not necessarily derive a symplectic ROB are called non-symplectic MOR techniques in the following. The non-symplectic MOR techniques investigated in the scope of our numerical results are the POD applied to the full state $\mathbf{x}(t, \boldsymbol{\mu})$ (POD full state) and a POD applied to the displacement $\mathbf{q}(t, \boldsymbol{\mu})$ and linear momentum states

method	solution	solution procedure	ortho-norm.	syml.
POD full	$\mathbf{V}_k = \mathbf{U}(:, 1:k)$	$\mathbf{U} = \text{SVD}(\mathbf{X}_s)$	✓	✗
POD separate	$\mathbf{V}_k = \begin{bmatrix} \mathbf{U}_p(:, 1:k) \\ \mathbf{U}_q(:, 1:k) \end{bmatrix}$	$\mathbf{U}_p = \text{SVD}([\mathbf{p}_1, \dots, \mathbf{p}_{n_s}])$ $\mathbf{U}_q = \text{SVD}([\mathbf{q}_1, \dots, \mathbf{q}_{n_s}])$	✓	✗
PSD cSVD	$\mathbf{V}_{2k} = [\mathbf{E}(:, 1:k) \quad \mathbb{J}_{2n}^\top \mathbf{E}(:, 1:k)]$	$\mathbf{E} = \begin{bmatrix} \Phi \\ \Psi \end{bmatrix}, \Phi + i\Psi = \text{cSVD}(\mathbf{C}_s)$ $\mathbf{C}_s = [\mathbf{p}_1 + i\mathbf{q}_1, \dots, \mathbf{p}_{n_s} + i\mathbf{q}_{n_s}]$	✓	✓
PSD greedy	$\mathbf{V}_{2k} = [\mathbf{E}(:, 1:k) \quad \mathbb{J}_{2n}^\top \mathbf{E}(:, 1:k)]$	\mathbf{E} from greedy algorithm	✓	✓
PSD SVD-like	$\mathbf{V}_{2k} = [\mathbf{s}_{i_1}, \dots, \mathbf{s}_{i_k}, \mathbf{s}_{n+i_1}, \dots, \mathbf{s}_{n+i_k}]$	$\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_{2n}]$ from (22), $\mathcal{I}_{\text{PSD}} = \{i_1, \dots, i_k\}$ from (25)	✗	✓

Table 1: Basis generation methods used in the numerical experiments in summary, where we use the MATLAB[®] notation to denote the selection of the first k columns of a matrix e.g. in $\mathbf{U}(:, 1:k)$.

$\mathbf{p}(t, \boldsymbol{\mu})$ separately (POD separate states). To summarize the basis generation methods, let us enlist them in Table 1 where $\text{SVD}(\bullet)$ and $\text{cSVD}(\bullet)$ denote the SVD and the complex SVD, respectively.

All presented experiments are generalization experiments, i.e. we choose 9 different training parameter vectors $\boldsymbol{\mu} \in \mathcal{P}$ on a regular grid to generate the snapshots and evaluate the reduced models for 16 random parameter vectors that are distinct from the 9 training parameter vectors. Thus, the number of snapshots is $n_s = 9 \cdot 151 = 1359$. The size $2k$ of the ROB \mathbf{V} is varied in steps of 20 with $2k \in \{20, 40, \dots, 280, 300\}$.

The software used for the numerical experiments is RBmatlab¹ which is an open-source library based on the proprietary software package MATLAB[®] and contains several reduced simulation approaches. An add-on to RBmatlab is provided² which includes all the additional code to reproduce the results of the present paper. The versions used in the present paper are RBmatlab 1.16.09 and MATLAB[®] 2017a.

4.1 Autonomous beam model

In the first model, we load the beam on the free end (far right) with a constant force which induces an oscillation. Due to the constant force, the discretized system can be formulated as an autonomous Hamiltonian system. Thus, the Hamiltonian is constant and its preservation in the reduced models can be analysed. All other reduction results are very similar to the non-autonomous case and thus, are exclusively presented for the non-autonomous case in the following Section 4.2.

4.1.1 Preservation over time of the modified Hamiltonian in the reduced model

In the following, we investigate the preservation of the Hamiltonian of our reduced models. With respect to Remark 6, we mean the preservation over time of the modified Hamiltonian. Since the Hamiltonian is quadratic in our example and the implicit midpoint is a symplectic Runge-Kutta integrator, the modified Hamiltonian equals the original which is why we speak of “the Hamiltonian” in the following.

We present in Fig. 2 the count of the total 240 simulations which show a preservation (over time) of the reduced Hamiltonian in the reduced model. The solution \mathbf{x}_r of a reduced simulation preserves the reduced Hamiltonian over time if $(\mathcal{H}_r(\mathbf{x}_r(t_i), \boldsymbol{\mu}) - \mathcal{H}_r(\mathbf{x}_r(t_0), \boldsymbol{\mu})) / \mathcal{H}_{\text{rel}}(\boldsymbol{\mu}) < 10^{-10}$ for all discrete times $t_i \in [t_0, t_{\text{end}}]$, $1 \leq i \leq n_t$ where $\mathcal{H}_{\text{rel}}(\boldsymbol{\mu}) > 0$ is a parameter-dependent normalization factor. The heat map shows that no simulation in the non-symplectic case preserves the Hamiltonian whereas the symplectic methods all preserve the Hamiltonian which is what was expected from theory.

¹<https://www.morepas.org/software/rbmatlab/>

²<https://doi.org/10.5281/zenodo.2578078>

In Fig. 3, we exemplify the non-constant evolution of the reduced Hamiltonian for three non-symplectic bases generated by POD separate states with different basis sizes and one selected test parameter $(\lambda, \mu) \in \mathcal{P}$. It shows that in all three cases, the Hamiltonian starts to grow exponentially.

POD full state	0/240
POD separate states	0/240
PSD complex SVD	240/240
PSD greedy	240/240
PSD SVD-like	240/240

Figure 2: Heat map which shows the preservation of the reduced Hamiltonian in the reduced model in x of y cases (x/y) .

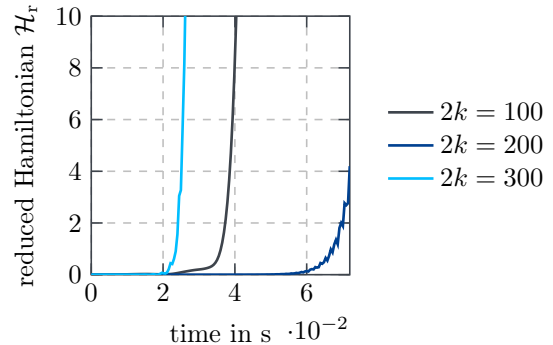


Figure 3: Evolution of the reduced Hamiltonian for POD separate states for a selected parameter $(\lambda, \mu) \in \mathcal{P}$.

4.2 Non-autonomous beam model

The second model is similar to the first one. The only difference is that the force on the free (right) end of the beam is loaded with a time-varying force. The force is chosen to act in phase with the beam. The time dependence of the force necessarily requires a non-autonomous formulation which requires in the framework of the Hamiltonian formulation a time-dependent Hamiltonian function which we introduced in Section 2.4.

We use the model to investigate the quality of the reduction for the considered MOR techniques. To this end, we investigate the projection error, i.e. the error on the training data, the orthogonality and symplecticity of the ROB and the error in the reduced model for the test parameters.

4.2.1 Projection error of the snapshots and singular values

The projection error is the error on the training data collected in the snapshot matrix \mathbf{X}_s , i.e.

$$e_{l_2}(2k) = \left\| (\mathbf{I}_{2n} - \mathbf{V}\mathbf{W}^T)\mathbf{X}_s \right\|_{\mathbb{F}}^2, \quad \begin{array}{l} \text{POD : } \mathbf{W}^T = \mathbf{V}^T, \\ \text{PSD : } \mathbf{W}^T = \mathbf{V}^+ (= \mathbf{V}^T \text{ for orthosymplectic ROB's, Proposition 4}). \end{array}$$

It is a measure for the approximation qualities of the ROB based on the training data. Fig. 4 (left) shows this quantity for the considered MOR techniques and different ROB sizes $2k$. All basis generation techniques show an exponential decay. As expected from theory, POD full state minimizes the projection error for the orthonormal basis generation techniques (see Table 1). PSD SVD-like decomposition shows a lower projection error than the other PSD methods for $2k \geq 80$ and yields a similar projection error for $k \leq 60$. Concluding this experiment, one might expect the full-state POD to yield decent results or even the best results. The following experiments prove this expectation to be wrong.

The decay of (a) the classical singular values σ_i , (b) the symplectic singular values σ_i^s (see Remark 4) and (c) the weighted symplectic singular values w_i^s (see (24)) sorted by the magnitude of the symplectic singular values is displayed in Fig. 4 (right). All show an exponential decrease. The weighting introduced in (24) for w_i^s does not influence the exponential decay rate of σ_i^s . The decrease in the classical singular values is directly linked to the exponential decrease of the projection error of POD full state due to properties of the Frobenius norm (see [4]). A similar result was deduced in the scope of the present paper for PSD SVD-like decomposition and the PSD functional (see Proposition 11).

4.2.2 Orthonormality and symplecticity of the bases

To verify the orthonormality and the symplecticity numerically, we consider the two functions

$$o_{\mathbf{V}}(2k) = \left\| \mathbf{V}^T \mathbf{V} - \mathbf{I}_{2k} \right\|_{\mathbb{F}}, \quad s_{\mathbf{V}}(2k) = \left\| \mathbb{J}_{2k}^T \mathbf{V}^T \mathbb{J}_{2n} \mathbf{V} - \mathbf{I}_{2k} \right\|_{\mathbb{F}} \quad (28)$$

which are zero / numerically zero if and only if the basis is orthonormal or symplectic, respectively. In Fig. 5, we show both values for the considered basis generation techniques and RB sizes.

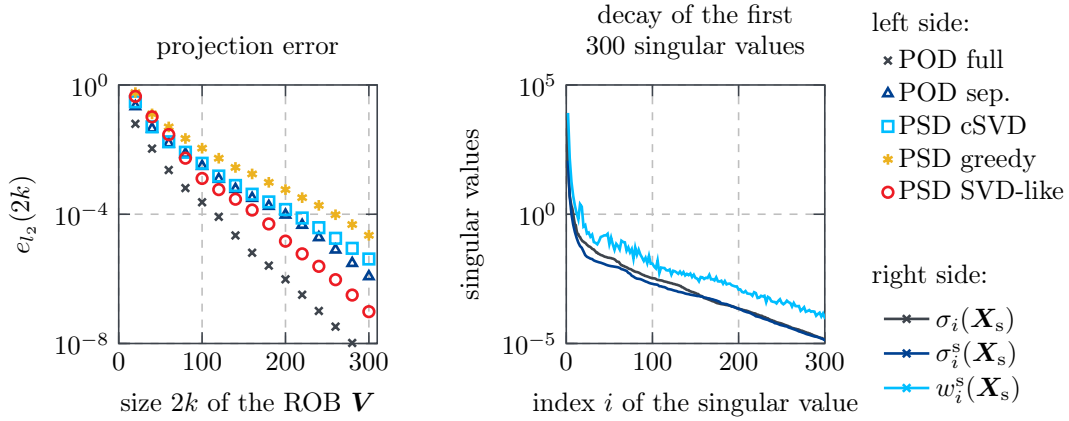


Figure 4: Projection error (left) and decay of the singular values from Remark 4 and (24) (right).

The orthonormality of the bases is in accordance with the theory. All procedures compute symplectic bases except for PSD SVD like-decomposition. PSD greedy shows minor loss in the orthonormality which is a known issue for the \mathbb{J}_{2n} -orthogonalization method used (modified symplectic Gram-Schmidt procedure with re-orthogonalization [2]). But no major impact on the reduction results could be attributed to this deficiency in the scope of this paper.

Also the symplecticity (or \mathbb{J}_{2n} -orthogonality) of the bases behaves as expected. All PSD methods generate symplectic bases whereas the POD methods do not. A minor loss of symplecticity is recorded for PSD SVD-like decomposition which is objected to the computational method that is used to compute an SVD-like decomposition. Further research on algorithms for the computation of an SVD-like decomposition should improve this result. Nevertheless, no major impact on the reduction results could be attributed to this deficiency in the scope of this paper.

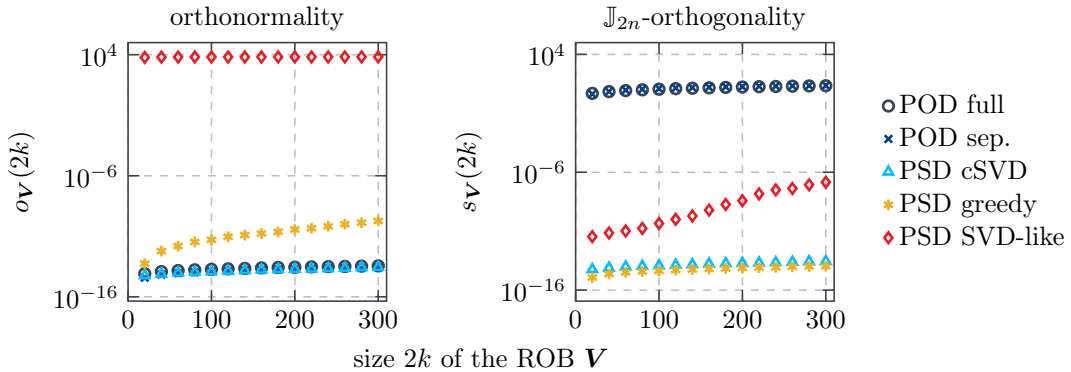


Figure 5: The orthonormality (left) and the \mathbb{J}_{2n} -orthogonality (right) from (28).

4.2.3 Relative error in the reduced model

We investigate the error introduced by MOR in the reduced model. The error is measured in the relative ∞ -norm $\|\bullet\|_\infty$ in time and space

$$\bar{e}(2k, \boldsymbol{\mu}) := \frac{\max_{i \in \{1, \dots, n_t\}} \|\mathbf{x}(t_i, \boldsymbol{\mu}) - \mathbf{V} \mathbf{x}_r(t_i, \boldsymbol{\mu})\|_\infty}{\max_{i \in \{1, \dots, n_t\}} \|\mathbf{x}(t_i, \boldsymbol{\mu})\|_\infty}, \quad (29)$$

where $2k$ indicates the size of the ROB $\mathbf{V} \in \mathbb{R}^{2n \times 2k}$ and $\boldsymbol{\mu} \in \mathcal{P}$ is one of the test parameters. To display the results for all 16 test parameters at once, we use box plots in Fig. 6. The box represents the 25%-quartile, the median and the 75%-quartile. The whiskers indicate the range of data points which lay within 1.5 times the interquartile range (IQR). The crosses show outliers. For the sake of a better overview, we truncated relative errors above $10^0 = 100\%$.

The experiments show that the non-symplectic MOR techniques show a strongly non-monotonic behaviour for an increasing basis size. For many of the basis sizes, there exists a parameter which shows crude approximation results which lay above 100% relative error. The POD full state is unable to produce results with a relative error below 2%.

On the other hand, the symplectic MOR techniques show an exponentially decreasing relative error. Furthermore, the IQRs are much lower than for the non-symplectic methods. We stress that the logarithmic scale of the y axis distorts the comparison of the IQRs – but only in favour of the non-symplectic methods. The low IQRs for the symplectic methods show that the symplectic MOR techniques derive a reliable reduced model that yields good results for any of the 16 randomly chosen test parameters. Furthermore, none of the systems shows an error above 0.19% – for PSD SVD-like decomposition this bound is 0.018%, i.e. one magnitude lower.

In the set of the considered symplectic, orthogonal MOR techniques, PSD greedy shows the best result for most of the considered ROB sizes. This superior behaviour of PSD greedy in comparison to PSD complex SVD is unexpected since PSD greedy showed inferior results for the projection error in Section 4.2.1. This was also observed in [15].

Within the set of investigated symplectic MOR techniques, PSD SVD-like decomposition shows the best results followed by PSD greedy and PSD complex SVD. While the two orthonormal procedures show comparable results, PSD SVD like-decomposition shows an improvement in the relative error. Comparing the best result of either PSD greedy or PSD complex SVD with the worst result of PSD SVD-like decomposition considering the 16 different test parameters for a fixed basis size – which is pretty much in favour of the orthonormal basis generation techniques –, the improvement of PSD SVD-like decomposition ranges from factor 3.3 to 11.3 with a mean of 6.7.

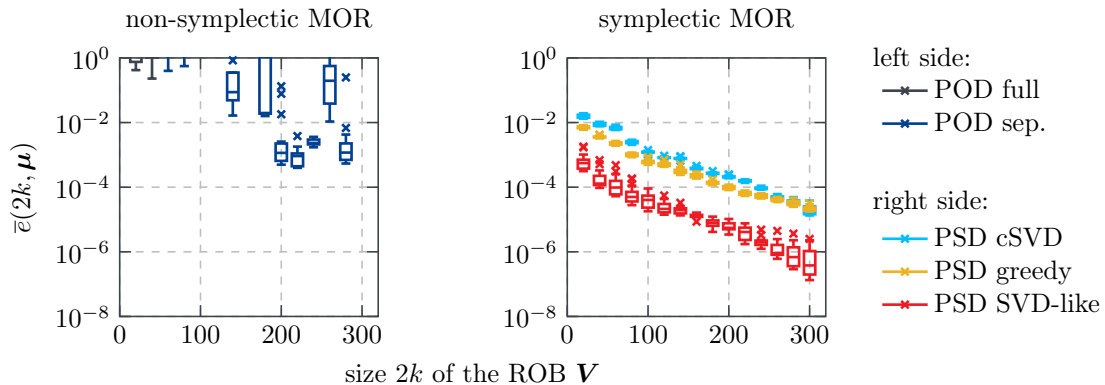


Figure 6: Relative error in the reduced model.

5 Summary and conclusions

We gave an overview of autonomous and non-autonomous Hamiltonian systems and the structure-preserving model order reduction (MOR) techniques for these kinds of systems [19, 15, 16]. Furthermore, we classified the techniques in orthonormal and non-orthonormal procedures based on the capability to compute a symplectic, (non-)orthonormal reduced order basis (ROB). To this end, we introduced a characterization of rectangular, symplectic matrices with orthonormal columns. Based thereon, an alternative formulation of the PSD Complex SVD [19] was derived which we used to prove the optimality with respect to the PSD functional in the set of orthonormal, symplectic ROBs. As a new method, we presented a symplectic, non-orthonormal basis generation procedure that is based on an SVD-like decomposition [21]. First theoretical results show that the quality of approximation can be linked to a quantity we referred to as weighted symplectic singular values.

The numerical examples show advantages for the considered linear elasticity model for symplectic MOR if a symplectic integrator is used. We were able to reduce the error introduced by the reduction with the newly introduced non-orthonormal method.

We conclude that non-orthonormal methods are able to derive bases with a lower error for both, the training and the test data. Yet, it is still unclear if the newly introduced method computes the global optimum of the PSD functional. Further work should investigate if a global optimum of the PSD functional can be computed with an SVD-like decomposition.

Funding— This research was partly funded by the German Research Foundation (DFG) grant number HA5821/5-1 and within the GRK 2198/1.

Acknowledgments— We thank the German Research Foundation (DFG) for funding this work. The authors thank Dominik Wittwar for inspiring discussions.

References

- [1] S. Agoujil, A. H. Bentbib, and A. Kanber. An iterative method for computing a symplectic svd-like decomposition. *Computational and Applied Mathematics*, 37(1):349–363, Mar 2018.
- [2] E.S. Al-Aidarous. Symplectic Gram-Schmidt Algorithm with Re-Orthogonalization. *Journal of King Abdulaziz University: Science*, 23(1):11–20, 2011.
- [3] Maxime Barrault, Yvon Maday, Ngoc Cuong Nguyen, and Anthony T. Patera. An ‘empirical interpolation’ method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathématique*, 339(9):667–672, 2004.
- [4] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox. *Model Reduction and Approximation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017.
- [5] A. Bhatt and B. Moore. Structure-preserving exponential runge–kutta methods. *SIAM Journal on Scientific Computing*, 39(2):A593–A612, 2017.
- [6] Patrick Buchfink. Structure-Preserving Model Order Reduction of Hamiltonian Systems for Linear Elasticity. *ARGESIM Report*, 55:35–36, 2018. Mathmod 2018 Extended Abstracts.
- [7] Angelika Bunse-Gerstner. Matrix factorizations for symplectic qr-like methods. *Linear Algebra and its Applications*, 83:49 – 77, 1986.
- [8] Ana Cannas da Silva. *Lectures on Symplectic Geometry*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [9] S. Chaturantabut and D. C. Sorensen. Discrete empirical interpolation for nonlinear model reduction. In *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 4316–4321, 2009.
- [10] Bernard Haasdonk and Mario Ohlberger. Efficient Reduced Models and A-Posteriori Error Estimation for Parametrized Dynamical Systems by Offline/Online Decomposition. *Math. Comput. Model. Dyn. Syst.*, 17(2):145–161, 2011.
- [11] Ernst Hairer, Gerhard Wanner, and Christian Lubich. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Berlin Heidelberg, 2006.
- [12] C. Lanczos. *The Variational Principles of Mechanics*. Dover Books on Physics. Dover Publications, 1970.
- [13] Hans Petter Langtangen and Geir K. Pedersen. *Scaling of Differential Equations*. Springer Publishing Company, Incorporated, 1st edition, 2016.
- [14] Benedict Leimkuhler and Sebastian Reich. *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2005.
- [15] Babak Maboudi Afkham and Jan S. Hesthaven. Structure Preserving Model Reduction of Parametric Hamiltonian Systems. *SIAM J. Sci. Comput*, 39(6):A2616–A2644, 2017.
- [16] Babak Maboudi Afkham and Jan S. Hesthaven. Structure-preserving model-reduction of dissipative hamiltonian systems. *Journal of Scientific Computing*, Feb 2018.
- [17] Matthias Oechsner, Karl-Heinz Kloos, B. Pyttel, Christina Berger, Michael Kübler, Andreas K. Müller, Karl-Heinz Habig, and Mathias Woydt. *Anhang E: Diagramme und Tabellen*, pages 324–357. Springer Berlin Heidelberg, 2014.
- [18] Chris Paige and Charles Van Loan. A schur decomposition for hamiltonian matrices. *Linear Algebra and its Applications*, 41:11 – 32, 1981.
- [19] Liaian Peng and Kamran Mohseni. Symplectic Model Reduction of Hamiltonian Systems. *SIAM J. Sci. Comput*, 38(1):A1–A27, 2016.
- [20] H. Xu. A numerical method for computing an svd-like decomposition. *SIAM Journal on Matrix Analysis and Applications*, 26(4):1058–1082, 2005.
- [21] Hongguo Xu. An svd-like matrix decomposition and its applications. *Linear Algebra and its Applications*, 368:1 – 24, 2003.