Krylov projection methods for linear Hamiltonian systems

Lu Li · Elena Celledoni

Abstract We study geometric properties of Krylov projection methods for large and sparse linear Hamiltonian systems. We consider in particular energy-preservation. We discuss the connection to structure preserving model reduction. We illustrate the performance of the methods by applying them to Hamiltonian PDEs.

1 Introduction

Large and sparse linear Hamiltonian systems arise in many fields of science and engineering, examples are models in network dynamics [1] and the semi-discretization of Hamiltonian partial differential equations (PDEs), like the wave equation [2,3] and Maxwell's equations [4,5]. In the context of Hamiltonian PDEs, the energy conservation law often plays a crucial role in the proof of existence and uniqueness of solutions [6]. Energy-preservation under numerical discretization can be advantageous as it testifies correct qualitative behaviour of the numerical solution, and it is also useful to prove convergence of numerical schemes [7]. There is an extensive literature on energy-preserving methods for ordinary differential equations (ODEs) [8,9,10,11], but these methods need to be implemented efficiently to be competitive for large and sparse systems arising in numerical PDEs. Krylov projection methods are attractive for discrete PDE problems because they are iterative, accurate and they allow for

Elena Celledoni
Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway
Tel.: +47 73593541
E-mail: elena.celledoni@ntnu.no

Lu Li
Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway
Tel.: +47 73591650
E-mail: lu.li@ntnu.no

restart and preconditioning strategies. But their structure preserving properties are not completely understood and should be further studied. This paper is a contribution in this direction.

It is well known that integration methods cannot be simultaneously symplectic and energy-preserving for general Hamiltonian systems [12]. However, the situation changes when we restrict to linear systems. An example is the midpoint rule which is symplectic and is also energy-preserving on linear problems; this is because the energy is quadratic for linear problems and the midpoint rule preserves all quadratic invariants. The midpoint method is implicit and requires the solution of one linear system of algebraic equations at each time step. The structure preserving properties are then retained only to the precision of the linear iterative solver. In this paper, we investigate preservation of geometric properties in Krylov projection methods for the exponential function. These are popular methods for the solution of discrete linear time dependent PDEs [13,14], but because of the Krylov projection, structure is only preserved to the accuracy of the method. On the other hand, we show that some of these Krylov projection methods can be energy-preserving to a higher level of precision, and can preserve several first integrals simultaneously. We finally discuss the connections to structure-preserving model reduction and variational principles. In particular, we consider modified Hamilton's principle as the natural variational formulation for projection methods based on block J-orthogonal basis. Previous work in the context of structure preserving Krylov projection methods can be found in [15,16] and for Hamiltonian eigenvalue problems for example in [17].

The structure of this paper is as follows. We discuss symplecticity in section 2. Section 3 is devoted to the preservation of first integrals. Section 4 is devoted to projection methods based on block J-orthogonal bases and their connection to structure preserving model reduction. In Section 5, the geometric properties of the considered methods are illustrated by numerical examples.

2 Krylov projection and symplecticity

Consider a linear Hamiltonian initial value problem of the form

$$\dot{y} = f(y) = JH\,y, \quad y(t_0) = y_0, \qquad\qquad J = J_m = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}, \qquad (1)$$

where $y(t) \in \mathbb{R}^{2m}$, $H \in \mathbb{R}^{2m \times 2m}$ is symmetric, $H^T = H$, $y_0 \in \mathbb{R}^{2m}$, and I_m is the $m \times m$ identity matrix. In what follows we denote by A the product $A = JH$. The matrix J is skew-symmetric, $J^T = -J$, and it defines a symplectic inner product[1] on $\mathbb{R}^{2m}$, $\omega(x,y) := x^T Jy$. Considering the energy function $\mathcal{H}(y) := \frac{1}{2} y^T Hy$, we have the gradient of $\mathcal{H}$ is $\nabla \mathcal{H}(y) = Hy$. The vector field

---

[1] A symplectic inner product on a vector space is a nondegenerate skew-symmetric bilinear form [18].

of equation (1) is a Hamiltonian vector field, i.e. $\omega(f(y), v) = \nabla\mathcal{H}(y)^T v$, $\forall v \in \mathbb{R}^{2m}$. From this it follows that the flow map,

$$\varphi_t : \mathbb{R}^{2m} \to \mathbb{R}^{2m}, \qquad y_0 \mapsto y(t),$$

is a symplectic map [19], i.e. it satisfies

$$\Psi_{y_0}(t)^T J \Psi_{y_0}(t) = J, \qquad \text{where} \quad \Psi_{y_0}(t) := \frac{\partial\varphi_t(y_0)}{\partial y_0}. \qquad (2)$$

A non-constant function $\mathcal{I}(y)$ is a first integral of the ODE $\dot{y} = f(y)$, if $\frac{d\mathcal{I}(y)}{dt}\mid_{y=y(t)} = \nabla\mathcal{I}(y)\dot{y} = \nabla\mathcal{I}(y)f(y) = 0$ for all y. So $\mathcal{I}(y)$ is constant along the solution trajectory: $\mathcal{I}(y(t)) - \mathcal{I}(y(t_0)) = \int_{t_0}^{t} \nabla\mathcal{I}(y)\dot{y}dt = 0$. The energy function $\mathcal{H}(y)$ is a first integral of (1). An approximation one-step method for (1) is said to be energy-preserving if $\mathcal{H}$ is constant along the numerical solution, and symplectic if the numerical one-step method (numerical flow map)

$$\phi_h : \mathbb{R}^{2m} \to \mathbb{R}^{2m}, \qquad y_0 \mapsto \tilde{y} \approx y(t_0 + h)$$

is such that

$$\frac{\partial\phi_h(y_0)}{\partial y_0}^T J \frac{\partial\phi_h(y_0)}{\partial y_0} = J,$$

[19].

The idea of Krylov projection methods is to build numerical approximations for (1) in the Krylov subspace:

$$\mathcal{K}_r(A, y_0) := \text{span}\{y_0, Ay_0, \cdots, A^{r-1}y_0\},$$

which is a subspace of $\mathbf{R}^{2m}$ of dimension $r << 2m$. Let us consider even dimension $r = 2n$. A basis of $\mathcal{K}_{2n}(A, y_0)$ is constructed. The most well known Krylov projection method is the one based on the Arnoldi algorithm [20] generating an orthonormal basis for $\mathcal{K}_{2n}(A, y_0)$. The method gives rise to a $2m \times 2n$ matrix V_{2n} with orthonormal columns, and to an upper Hessenberg $2n \times 2n$ matrix T_{2n} such that $I_{2n} = V_{2n}{}^T V_{2n}$, and $T_{2n} = V_{2n}^T A V_{2n}$. The approximation of $y(t)$ is

$$y_A(t) := V_{2n}z(t), \quad \text{where} \quad \dot{z} = T_{2n}z, \quad z(0) = z_0 = V_{2n}^T y_0. \qquad (3)$$

We will denote this method by Arnoldi projection method (APM). Consider J_{2n} and the corresponding symplectic inner product in $\mathbb{R}^{2n}$, $\omega(u, v) := u^T J_{2n}v$. If $n < m$, unless we make further assumptions on H, the projected system (3) is not a Hamiltonian system in $\mathbb{R}^{2n}$, this is because $J_{2n}^{-1}T_{2n} = J_{2n}^{-1}V_{2n}^T JHV_{2n}$ is in general not symmetric and $J_{2n}^{-1}T_{2n}z$ is in general not the gradient of some energy function.

Instead of using an orthonormal basis, one can construct a J-orthogonal basis for $\mathcal{K}_{2n}(A, y_0)$ using the symplectic Lanczos algorithm [21]. The matrix S_{2n} whose columns are the vectors of this J-orthogonal basis satisfies

$$S_{2n}^T J S_{2n} = J_{2n}.$$

We will denote the corresponding Krylov projection method by Symplectic Lanczos projection method (SLPM). The projected system for SLPM is analog to (3), with V_{2n} replaced by S_{2n}, T_{2n} by $J_{2n}S_{2n}^T H S_{2n}$ and an appropriate z_0 (see Section 3.3). This projected system is a Hamiltonian system. But for $n < m$, the approximation $y_S(t) := S_{2n}z(t)$ is not symplectic. In fact, y_S is the solution of the system

$$\dot{y}_S = (S_{2n}J_{2n}S_{2n}^T)\,H\,y_S, \quad y_S(t_0) = y_0, \tag{4}$$

and (4) is a Poisson system with Poisson structure given by the matrix $S_{2n}J_{2n}S_{2n}^T$ which is skew-symmetric and depends on the initial condition[2]. For $n = m$, $J_{2m} = J$, and $y_S = y$. However, the case $n < m$ is the most relevant for the use of the method in practice. In spite of not preserving the symplectic inner product ω, SLPM clearly shares important structural properties with the exact solution of (1) and is energy-preserving, see Section 3.3.

The symplectic Lanczos algorithm is not the only way to obtain a J-orthogonal basis of the Krylov subspace. We will consider block J-orthogonal bases in Section 4 and show that they can be viewed as techniques of structure preserving model reduction, in the spirit of [22]. We propose one Krylov algorithm based on these ideas.

3 Preservation of first integrals and energy

We first present a result about the first integrals for a general linear Hamiltonian system. Recall that two first integrals F and G of an ODE are said to be in involution if their Jacobi bracket $\{F, G\} := (\nabla F)^T J \nabla G$ vanishes [19].

Proposition 1 *For $A = JH$ where J is skew-symmetric and invertible, and H is symmetric and invertible, the system $\dot{y} = Ay$, $y(t_0) = y_0$ has the following first integrals in involution, $\mathcal{H}_k(y) = \frac{1}{2}y^T H A^{2k}y$, for $k = 0, 1, \ldots$. The Hamiltonian of the system is $\mathcal{H} = \mathcal{H}_0$.*

Proof We consider the derivative of $\mathcal{H}_k$ along solution trajectories $y(t)$

$$\begin{aligned}
\frac{d}{dt}\mathcal{H}_k(y(t)) &= \frac{1}{2}\left[\dot{y}^T H(JH)^{2k}y + y^T H(JH)^{2k}\dot{y}\right] \\
&= \frac{1}{2}\left[-y^T HJH(JH)^{2k}y + y^T H(JH)^{2k}JHy\right] \\
&= \frac{1}{2}\left[-y^T H(JH)^{2k+1}y + y^T H(JH)^{2k+1}y\right] = 0,
\end{aligned}$$

so $\mathcal{H}_k(y)$, $k = 0, \ldots$, are preserved: $\mathcal{H}_k(y(t)) = \mathcal{H}_k(y_0)$. The integrals are in involution because their Poisson bracket is zero,

$$\begin{aligned}
\{\mathcal{H}_k, \mathcal{H}_p\} &= (\nabla\mathcal{H}_k)^T J\nabla\mathcal{H}_p = (A^{2k}y)^T HJH(A^{2p}y) \\
&= y^T((JH)^{2k})^T HJH(JH)^{2p}y = y^T H(JH)^{2(k+p)+1}y = 0,
\end{aligned}$$

---

[2] A Poisson system in $\mathbb{R}^d$ is a system of the type $\dot{y} = \Omega\,\nabla\mathcal{H}(y)$, where Ω is a skew-symmetric matrix, not necessarily invertible and can depend on y. Ω must satisfy the Jacobi identity, [19]. In our case, Ω depends on y_0.

where we have used the skew-symmetry of $H(JH)^{2(k+p)+1}$.

In what follows, we will discuss the preservation of the first integrals of Proposition 1 when applying Krylov projection methods.

3.1 Preservation of first integrals for the APM

It can be observed from numerical simulations that the APM fails in general to preserve energy when applied to Hamiltonian systems, Figure 1 (left), Section 5, but structure-preserving properties can be ensured for such method via a simple change of inner product. Assume that H is symmetric and positive definite so that $\langle \cdot, \cdot \rangle_H := \langle \cdot, H \cdot \rangle$ defines an inner product. We modify the Arnoldi algorithm by replacing the usual inner product $\langle \cdot, \cdot \rangle$ by $\langle \cdot, \cdot \rangle_H$. We then show that the numerical solution given by this method preserves to machine accuracy certain first integrals. The modified Arnoldi algorithm (see Algorithm 1a) generates a H-orthonormal basis, which is stored in the $2m \times n$ matrix V_n, satisfying $V_n^T H V_n = I_n$. This algorithm generates a skew-symmetric tridiagonal matrix T_n such that

$$AV_n = V_n T_n + w_{n+1} e_n^T, \quad w_{n+1} = h_{n+1,n} v_{n+1},$$
$$V_n{}^T H V_n = I_n, \quad V_n{}^T H w_{n+1} = 0.$$

In what follows, we consider the Krylov projection method

$$y_H := V_n z, \quad \text{where } z \text{ satisfies} \quad \dot{z} = T_n z, \quad z(t_0) = V_n{}^T H y_0.$$

Proposition 2 *The numerical approximation y_H for the solution y of (1) preserves the following first integrals:*

$$\bar{\mathcal{H}}_k(y_H) = \frac{1}{2} y_H^T H V_n (T_n)^{2k} V_n{}^T H y_H \tag{5}$$

for all $k = 0, 1, \dots$.

Proof We observe that $T_n = V_n{}^T H J H V_n$ is skew-symmetric. So the ODE system for z has first integrals: $\mathcal{I}_k(z) = \frac{1}{2} z^T (T_n)^{2k} z$, for all $k = 0, 1, \dots$. Therefore $\bar{\mathcal{H}}_k(y_H) = \frac{1}{2} y_H^T H V_n (T_n)^{2k} V_n{}^T H y_H = \frac{1}{2} z^T (T_n)^{2k} z$ and $\frac{d}{dt} \bar{\mathcal{H}}_k(y_H(t)) = \frac{d}{dt} \mathcal{I}_k(z(t)) = 0$, so the first integrals are preserved.

Remark 1 If n is even, the above Krylov projection method induces a projected problem which is conjugate to a Hamiltonian system, i.e., it can be written in the form (1) via change of variables. Since H_n is skew-symmetric, H_n can be factorized as $H_n = U_n J_n D_n U_n^{-1}$ where D_n is diagonal. Then, H_n can be transformed to a Hamiltonian matrix by a similarity transformation using U_n.

3.2 Hamiltonian system with $JA = AJ$

We now consider J given by (1). Assume that A and J commute, then A is skew-symmetric, and the Hamiltonian system (1) has two Hamiltonian structures, one associated to A with Hamiltonian $\frac{1}{2}y^T y$, the other to J with Hamiltonian $\frac{1}{2}y^T H y$. The APM with Euclidean inner product $\langle \cdot, \cdot \rangle$ preserves modified first integrals. To proceed, we first give the following result.

Proposition 3 *Suppose A is a Hamiltonian matrix. Then J and A commute if and only if the matrix A is skew-symmetric.*

Proof Suppose A is a Hamiltonian matrix and $A = JH$, where J and H are defined as in equation (1). Then the fact that J and A commute implies that $JJH = JHJ$, i.e. $-H = JHJ$ and by multiplying J^{-1} from right side, we get $-(JH)^T = JH$, namely $A^T = -A$. On the other hand, the fact that A is skew-symmetric implies that $(JH)^T = -JH$ and using this we get $JA = JJH = -J(JH)^T = JHJ = AJ$.

The first integrals of the system (1) are given by the following proposition.

Proposition 4 *If $JA = AJ$, the Hamiltonian system (1) has the following first integrals in involution, $\mathcal{H}_k(y) = \frac{1}{2}y^T A^{2k} y$ for $k = 0, 1, \ldots$, and the fisrt integrals are in involution with the Hamiltonian $\mathcal{H}(y) = \frac{1}{2}y^T H y$.*

Proof From Proposition 3 we know that A is skew-symmetric. Then Proposition 1 holds with J replaced by A, and H replaced by the identity matrix. The integrals are in involution with the Hamiltonian $\mathcal{H}(y) = \frac{1}{2}y^T H y$ in fact

$$\{\mathcal{H}_k, \mathcal{H}\} = y^T A^{2k} J H y = y^T A^{2k} A y = 0, \quad k = 0, \ldots.$$

Remark 2 By a direct application of Proposition 2, the APM to the Hamiltonian system (1), under the assumption $JA = AJ$, gives a numerical approximation $y_A := V_n z$ which preserves the following modified first integrals

$$\bar{\mathcal{H}}_k(y_A) := \frac{1}{2}y_A^T V_n (H_n)^{2k} V_n^T y_A, \quad k = 0, 1, \ldots. \tag{6}$$

We next prove that the Hamiltonian of (1) is bounded by y_A under the assumption that J and A commute.

Proposition 5 *Assume the APM is applied to (1). Under the assumption $JA = AJ$, the energy $\mathcal{H}(y) = \frac{1}{2}y^T J^{-1} A y$, is bounded along the numerical solution.*

Proof This result follows directly from Remark 2 with $k = 0$, i.e.,

$$\frac{1}{2}y_A^T J^{-1} A y_A \leq \frac{1}{2}y_A^T y_A \|J^{-1}A\|_2 = \frac{1}{2}y_0^T y_0 \|J^{-1}A\|_2.$$

Proposition 5 explains the good behaviour of the APM in [23].

3.3 Symplectic Lanczos projection method

We now consider the symplectic Lanczos projection method (SLPM). Krylov subspace methods based on the symplectic Lanczos algorithm are widely used for the computation of eigenvalues of large and sparse Hamiltonian matrices [24, 25, 26]. For their use in the approximation of linear Hamiltonian systems see [27], [14].

Given $A \in R^{2m,2m}$ and the starting vector $y_0 \in R^{2m}$, the symplectic Lanczos method generates a sequence of matrices

$$S_{2n} = [v_1, ..., v_n, w_1, ...w_n] \quad \text{satisfying} \quad AS_{2n} = S_{2n}T_{2n} + r_{n+1}e_{2n}^T, \qquad (7)$$

where T_{2n} is a tridiagonal Hamiltonian matrix, and $r_{n+1} = \zeta_{n+1}v_{n+1}$ is J-orthogonal with respect to the columns of S_{2n}. Since S_{2n} has J-orthogonal columns, i.e., $S_{2n}{}^T J S_{2n} = J_{2n}$, we know that

$$T_{2n} = J_{2n}^{-1} S_{2n}^T J A S_{2n} = J_{2n} S_{2n}^T H S_{2n}, \qquad (8)$$

and the projected system is a Hamiltonian system, where $z_0 = J_{2n}^{-1} S_{2n}^T J y_0$. Moreover, we have

$$\mathcal{H}_S(z) = \frac{1}{2} z^T J_{2n}^{-1} T_{2n} z \equiv \frac{1}{2} z_0^T J_{2n}^{-1} T_{2n} z_0. \qquad (9)$$

Proposition 6 *The SLPM is an energy-preserving method for* (1).

Proof The result follows by computing the Hamiltonian of (1) along numerical trajectories $y_S = S_{2n} z$, $\mathcal{H}(y_S) = \frac{1}{2} y_S{}^T J^{-1} A y_S$, and then using (7) and (9).

4 Projection methods based on block J-orthogonal basis

We now consider a general strategy for Krylov projection methods to obtain J-orthogonal bases, this will lead automatically to energy preserving methods for (1). In what follows we will use the notation $(q^T, p^T)^T = y$ and write H in block form, and rewrite (1) accordingly:

$$\begin{aligned}\dot{q} &= H_{12}^T q + H_{22} p, \\ \dot{p} &= -H_{11} q - H_{12} p,\end{aligned} \qquad H = \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{bmatrix}. \qquad (10)$$

Assume that we can construct two matrices with linearly independent columns $V_n \in \mathbb{R}^{m \times n}$ and $W_n \in \mathbb{R}^{m \times n}$ such that $V_n{}^T W_n = I_n$. Then the matrix

$$S_{2n} := \begin{bmatrix} V_n & 0 \\ 0 & W_n \end{bmatrix} \qquad (11)$$

has J-orthogonal columns. We will approximate y by the following projection method: $y \approx y_B$ defined by

$$y_B = S_{2n} z, \quad \text{where } z \text{ satisfies} \quad \dot{z} = J_{2n} S_{2n}^T J^{-1} A S_{2n} z, \quad z(t_0) = z_0, \quad (12)$$

and for the SLPM $z_0 = J_{2n}^{-1} S_{2n}^T J y_0$.

Proposition 7 *If $y_0 = S_{2n}z(t_0)$, then the energy of the original Hamiltonian system* (1) *will be preserved by the numerical solution* (11)-(12).

Proof Notice that $\mathcal{H}(S_{2n}\,z) = \frac{1}{2}z^T S_{2n}^T J^{-1} A S_{2n}\,z$ is constant with respect to t because z is the solution of a Hamiltonian system with energy $\mathcal{E}(z) = \frac{1}{2}z^T(S_{2n}^T J^{-1} A S_{2n})\,z$. The result then follows directly from the fact that $\mathcal{E}(z) \equiv \mathcal{E}(z_0) = \mathcal{H}(y_0)$.

We here propose one strategy to construct S_{2n} as in (11) with $W_n^T V_n = I_n$ and $V_n = W_n$. Let K_n be the Krylov matrix $2m \times n$, and consider the first m rows of K_n and the last m separately:

$$K_n := [y_0, Ay_0, \ldots, A^{n-1}y_0], \qquad K_n = \begin{bmatrix} K_n^q \\ K_n^p \end{bmatrix}.$$

We then find an orthonormal basis V_n for span$\{K_n^q, K_n^p\} \subset \mathbb{R}^m$ by either a QR-factorisation (algorithm 1b in the Appendix [3]) or a Gram-Schmidt process.

4.1 Structure preserving model reduction using Krylov subspaces

In this section we consider Hamilton's phase space variational principle (also called modified Hamilton's principle) [28, Ch. 8-5] which is the fundament of the projection methods based on block J-orthogonal basis. Since our system (1) is given in the form of an Hamiltonian system, it is natural to use the phase space variational principle, which is formulated in terms of the variables p and q and the Hamiltonian function $H(q, p)$, rather than the classical Hamilton's principle which is formulated in terms of q and $\dot{q}$ and the Lagrangian function $L(q, \dot{q})$. Following [22], we restrict the phase space variational principle to low dimensional subspaces of $\mathbb{R}^m$ and derive the projected Hamiltonian system taking variations on the low dimensional subspaces.

Assume $[q^T, p^T]^T := y$ and q and p are m-dimensional vectors belonging to $\mathbf{R}^m$ and its dual respectively, and that the Hamiltonian $\mathcal{H} : \mathbf{R}^m \times (\mathbf{R}^m)^* \to \mathbf{R}$ is $\mathcal{H}(q, p) := \mathcal{H}(y)$.[4] Considering the action functional $\mathcal{S} : \mathbf{R}^m \times (\mathbf{R}^m)^* \to \mathbf{R}$

$$\mathcal{S}(q, p) := \int_{t_0}^{t_{end}} \left(p^T \dot{q} - \mathcal{H}(q, p) \right) \, dt, \tag{13}$$

Hamilton's phase space variational principle states that

$$\delta \mathcal{S} = 0$$

for fixed $q_0 = q(t_0)$ and $q_{end} = q(t_{end})$, and it is equivalent to Hamilton's equations (1), [28, Ch. 8-5] . By projecting q and p separately on appropriate subspaces span$\{V_n\} \subset \mathbf{R}^m$ and span$\{W_n\} \subset (\mathbf{R}^m)^*$, i.e., $q \approx V_n \hat{q}$ and

---

[3]   Notice that to obtain a stable algorithm it is an advantage to replace the Krylov matrix with an orthonormal matrix obtained by the Arnoldi algorithm.

[4]   The duality pairing between $\mathbf{R}^m$ and $(\mathbf{R}^m)^*$ is here simply $\langle p, q \rangle := p^T q$.

$p \approx W_n \hat{p}$, one restricts the variational principle to $\text{span}\{V_n\} \times \text{span}\{W_n\}$: $\hat{\mathcal{S}}(\hat{q}, \hat{p}) := \mathcal{S}(V_n \hat{q}, W_n \hat{p})$. Taking variations

$$0 = \delta \hat{\mathcal{S}}(\hat{q}, \hat{p}) = \delta \int_{t_0}^{t_{end}} (W_n \hat{p})^T V_n \dot{\hat{q}} - \mathcal{H}(V_n \hat{q}, W_n \hat{p}) dt$$

for fixed endpoints $\hat{q}_0 = \hat{q}(t_0)$ and $\hat{q}_{end} = \hat{q}(t_{end})$, we obtain the Hamiltonian equations associated to this reduced variational principle

$$\begin{aligned} \dot{\hat{p}} &= -V_n^T H_{12} W_n \hat{p} - V_n^T H_{11} V_n \hat{q}, \\ \dot{\hat{q}} &= W_n^T H_{22} W_n \hat{p} + W_n^T H_{12}^T V_n \hat{q}, \end{aligned} \tag{14}$$

which coincide with the system for z in (12). This explains the connection of the projection methods based on block J-orthogonal basis, (11) and (12), with the techniques of structure preserving model reduction derived in [22] and here modified for the phase space variational principle.

Assuming additional structure on H, we will show in the next section that the usual APM applied to the resulting system falls in the same class of projection methods based on block J-orthogonal basis and is a structure preserving model reduction method in the spirit of [22]. Model reduction methods for general second order systems obtained projecting the differential systems onto Krylov subspaces using an Arnoldi or a Lanczos process have been previously studied [29].

4.2 Special case $H_{1,2} = O$, $H_{2,2} = I$.

This special case is directly related to the setting in [22]. Denoting $y = (q^T, p^T)^T$, we consider the corresponding Hamiltonian system

$$\dot{y} = Ay \quad \text{with} \quad A = \begin{bmatrix} 0 & I \\ -H_{11} & 0 \end{bmatrix}. \tag{15}$$

and we notice that $p = \dot{q}$ in this case. The action functional (13) from the previous section is the integral of the Lagrangian density function

$$L(q(t), \dot{q}(t)) = \frac{1}{2} \dot{q}(t)^T \dot{q}(t) - \frac{1}{2} q(t)^T H_{11} q(t), \tag{16}$$

and in this case because $\dot{q} = p$ the phase space variational principle coincides with Hamilton's principle. Let V_n be the basis of the Krylov subspace $\mathcal{K}_n(-H_{11}, p_0)$ obtained via the Arnoldi algorithm. The reduced Lagrangian becomes

$$L(\hat{q}(t), \dot{\hat{q}}(t)) = \frac{1}{2} \dot{\hat{q}}(t)^T \dot{\hat{q}}(t) - \frac{1}{2} \hat{q}(t)^T V_n^T H_{11} V_n \hat{q}(t), \tag{17}$$

and the corresponding Hamiltonian equations are

$$\begin{aligned} \dot{\hat{q}} &= \hat{p}, \\ \dot{\hat{p}} &= -V_n^T H_{11} V_n \hat{q}(t). \end{aligned} \tag{18}$$

By solving (18), we obtain $(\hat{q}^T, \hat{p}^T)^T$ and then can construct the model reduction approximation $((V_n\hat{q})^T, (V_n\hat{p})^T)^T \approx (q^T, p^T)^T$.

Proposition 8 *When applied to* (15) *with* $y_0 = (0, p_0^T)^T$, *the model reduction procedure outlined in* (16)-(18) *coincides with the APM.*

Proof Let $\mathbf{e}_1, \mathbf{e}_2 \in \mathbf{R}^2$ be the two vectors of the canonical basis in $\mathbf{R}^2$. Denote by $\otimes$ the Kronecker tensor product. We have

$$\mathcal{K}_{2n}(A, y_0) = \operatorname{span}\{\mathbf{e}_1 \otimes p_0, \mathbf{e}_2 \otimes p_0, \mathbf{e}_1 \otimes (-H_{11}p_0), \mathbf{e}_2 \otimes (-H_{11})p_0, \dots\}.$$

Denote by $\mathbb{U}_{2n} \in \mathbb{R}^{2m \times 2n}$ the orthogonal matrix generated by the usual Arnoldi algorithm with matrix A, vector $y_0 = (0, p_0^T)^T$ and Euclidean inner product. Then $\mathbb{U}_{2n}$ is given by

$$\mathbb{U}_{2n} = \begin{bmatrix} 0 & v_1 & 0 & v_2 & 0 & \dots & 0 & v_n \\ v_1 & 0 & v_2 & 0 & v_3 & \dots & v_n & 0 \end{bmatrix},$$

and satisfies

$$\mathbb{U}_{2n}{}^T A \mathbb{U}_{2n} = \Pi_{2n} \begin{bmatrix} 0 & I_n \\ -V_n^T H_{11} V_n & 0 \end{bmatrix} \Pi_{2n}{}^T \quad \text{and} \quad \mathbb{U}_{2n}\Pi_{2n} = \begin{bmatrix} V_n & O \\ O & V_n \end{bmatrix},$$

where $v_1, v_2, \dots v_n$ are the columns of V_n and Π_{2n} is a $2n \times 2n$ permutation matrix. After a permutation of the variables $w = \Pi_{2n}{}^T z$, the projected system by APM $\dot{z} = \mathbb{U}_{2n}{}^T A \mathbb{U}_{2n} z$, $z(t_0) = \mathbb{U}_{2n}{}^T y_0$ can be rewritten in the form (11)-(12).

5 Numerical Examples

In this section, several numerical examples are presented to illustrate the behavior of the following methods:

– APM: Arnoldi projection method using Euclidean inner product, Section 3;
– APMH: Arnoldi projection method using the inner product $\langle\cdot,\cdot\rangle_H$, Section 3;
– SLPM: symplectic Lanczos projection method, Section 3.3;
– BJPM: block J-orthogonal projection method with QR factorization, Section 4.1.

These methods are applied to solve randomly generated linear Hamiltonian systems, and linear systems arising from the discretization of Hamiltonian PDEs. If not stated otherwise, the dimension of the original space is set to be $2m = 400$ and the dimension of the Krylov subspace is chosen to be $2n = 4$, for all Krylov methods compared. The reference solution is computed using the Cayley transformation (midpoint rule) with step-size 0.004. The solution of the projected system (3) is obtained with the same method and the same step-size used for the reference solution. All the errors in energy and in first integrals are relative errors.

To obtain a desired global error accuracy on $[0, T]$ for large T, we either use a sufficiently large dimension of the Krylov subspace or use a time-stepping (restart) procedure. More precisely, this entails subdividing $[0, T]$ into subintervals $[t_r, t_{r+1}]$ and performing the projection on each subinterval recomputing the basis of the Krylov subspace with starting vector $y_r \approx y(t_r)$. In the experiments, we use subintervals of size $t_{r+1} - t_r = 0.2$. The restart procedure is of practical interest because it allows to use a Krylov subspace of low dimension. In exact arithmetic the first integrals would be preserved exactly, however, due to the propagation of roundoff errors, a small linear drift in the preserved quantities is observed. The numerical experiments show that the drift in the energy error can be lessened by applying the restart technique. However, the restart compromises the preservation of the first integrals of Propositions 2 and 4 for APM and APMH simply because the basis V_n is recomputed on each subinterval.

5.1 Randomly generated Hamiltonian matrices

5.1.1 Case $JA = AJ$: APM

In the experiment considered in Figure 1 (left), $H = J^{-1}A$ is block diagonal, symmetric and positive definite but with no particular extra structure. There is a clear drift in the energy for the APM, and the energy is preserved for the APMH and SLPM. Similar experiments show that the global error of APMH and SLPM is bounded, while the global error of the APM is not (these errors are not reported here). If we apply the APM to an example where $JA = AJ$, the first integrals are preserved and the energy error and global error are bounded, see [23].
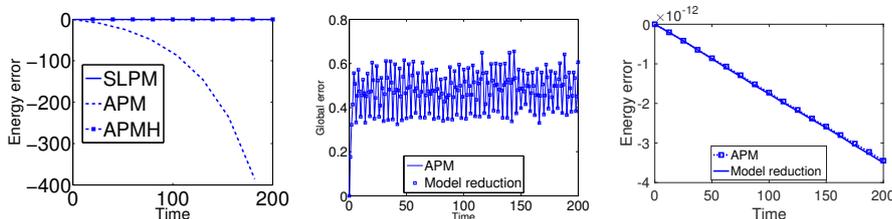
Fig. 1: **Left**: Methods without the restart applied to a block diagonal matrix A. Energy drift for APM and energy conservation for SLPM and APMH. **Middle**: Global error of APM and Model reduction versus time. **Right**: Relative energy error of APM and Model reduction versus time.

5.1.2 Case $H_{1,2} = O$, $H_{2,2} = I$: Model reduction

In this numerical test, we consider a Hamiltonian matrix A of the special form (15) with an initial vector of the form $y_0 = (0, p_0^T)^T$ and we apply the APM to

this system. For comparison, we use the model reduction procedure described in Section 4.2: we generate the orthogonal matrix V_n using the Arnoldi algorithm with matrix $-H_{11}$ and vector p_0. The methods behave as predicted, see Figure 1 (the middle and right figures). The experiment confirms that the APM in this case behaves as the model reduction method and preserves the energy. A small linear drift is observed at the level of roundoff and we will consider this error propagation in the next subsection.

5.1.3 Full matrices: Comparison of APMH, SLPM, BJPM

In this subsection, we consider a randomly generated, full Hamiltonian matrix $A = JH$. In Figure 2, we report numerical results for the methods APMH, SLPM and BJPM without restart. The left panel of Figure 2, reports the relative energy error for the methods. The right panel of Figure 2, illustrates the convergence of the methods: the global error at $T = 2$ decreases when the dimension of the Krylov subspace increases.

APMH is the method that better preserves the energy, but a linear error growth in time at the level of roundoff can be observed for all the methods and also in the error of the first integrals for APMH. To examine this propagation of roundoff errors, we compare the relative energy error and the error in the Cayley transformation as a function of time, see middle panel of Figure 2. For $t_k = \Delta t\, k$, we denote with $\mathrm{Cay}(t_k T_n) := \left((I - \frac{\Delta t}{2} T_n)^{-1}(I + \frac{\Delta t}{2} T_n)\right)^k$ the Cayley transformation approximating $\exp(t_k T_n)$. The error in the Cayley transformation is measured by verifying the orthogonality of the matrix $\mathrm{Cay}(t_k T_n)$. After one step ($t_1 = \Delta t$), this error is close to machine accuracy, $\|(\mathrm{Cay}(\Delta t T_n))^T \mathrm{Cay}(\Delta t T_n)) - I\|_2 = 1.1224e - 16$, but we see that $\|(\mathrm{Cay}(t_k T_n))^T \mathrm{Cay}(t_k T_n)) - I\|_2$ grows with t_k and comparably to the relative energy error. Likely, this error is the main cause of the roundoff error propagation in the energy. In this experiment, we have chosen $\Delta t := 2^{-s}$, where s is the minimum positive integer such that $2^{-s}\|T_n\|_1 \leq \frac{1}{2}$, see for example [30].

In Figure 3 (left), we see that the roundoff error drift is mitigated by applying the restart technique. In the right figure, we observe that for the methods with restart, the global error behaves well and stops increasing after a certain time.

5.2 Hamiltonian PDEs

In this section we apply the methods to the wave equations and the Maxwell's Equations.

5.2.1 Wave equation

We consider the 2D wave equations

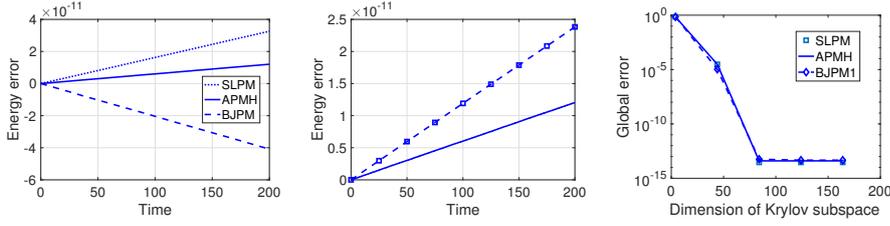$$\dot{\phi} = \psi, \qquad \dot{\psi} = \triangle\phi, \tag{19}$$

Fig. 2: Krylov projection methods applied to full matrices. **Left**: energy error for SLPM, APMH and BJPM, methods without restart. In this experiment $\Delta t := 2^{-s}$, where $s > 0$ is such that $2^{-s} \|T_n\|_1 \leq \frac{1}{2}$. **Middle**: relative energy error of APMH (solid line), reference line $\|(\text{Cay}(t_k T_n))^T \text{Cay}(t_k T_n)) - I\|_2$ (dotted square). Error in orthogonality and skew-symmetry: $\|V_n^T H V_n - I\|_2 = 2.8728e - 16$, $\|T_n^T + T_n\|_2 = 0$ and $\|(\text{Cay}(\Delta t T_n))^T \text{Cay}(\Delta t T_n)) - I\|_2 = 1.1224e - 16$. Same step-size as left panel. **Right**: global error versus dimension of the Krylov subspace.
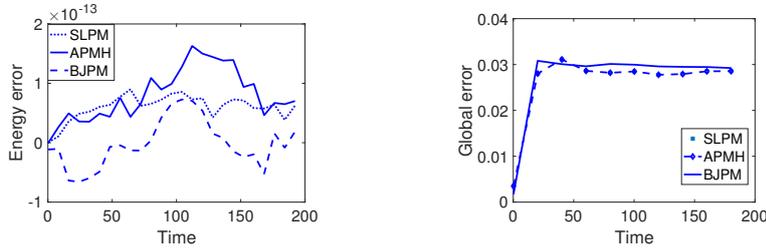
Fig. 3: Left, energy error, right, global error, methods with restart.

on $[0, 1] \times [0, 1]$ with homogeneous Dirichlet boundary conditions $\phi(t, 0, y) = \phi(t, 1, y) = \phi(t, x, 0) = \phi(t, x, 1) = 0$ and a randomly generated initial vector. Semi-discretizing on an equispaced grid $x_i = i \, \Delta x$ and $y_j = j \, \Delta y$, $\Delta x = \Delta y$, $i, j = 0, \ldots, N$ and assuming $u(x_i, y_j) \approx U_{i,j}$, we obtain a system

$$\dot{U} = AU, \quad U(0) = U_0, \qquad A = \begin{bmatrix} 0 & I \\ G & 0 \end{bmatrix} \tag{20}$$

with G the discrete 2D Laplacian obtained by using central differences. This is a Hamiltonian system with energy $\mathcal{H} = \frac{1}{2} U^T J A U \equiv \frac{1}{2} U(0)^T J A U(0)$. We perform experiments with all the Krylov projection methods discussed in this paper. Figure 4a shows that all the methods are energy-preserving.

5.2.2 1D *Maxwell's equations*

We consider 1D Maxwell's equations

$$\begin{aligned} \partial_t E &= \partial_x B, \\ \partial_t B &= \partial_x E \end{aligned} \tag{21}$$

(a) Energy error　　　　　　　　　　　(b) Global error-APMH

Fig. 4: Figure 4a: energy error for Wave equation in 2d, methods with restart are considered and the dimension of the problem is 392, namely $N = 15$. Figure 4b: energy error for Maxwell's equations in 1d, method with restart is considered.

for $x \in [0,1]$ and $t > 0$ with boundary conditions $E(0,t) = E(1,t) = 0$, $B_x(0,t) = B_x(1,t) = 0$ and initial conditions $E(x,0) = \sin(\pi x)$ and $B(x,0) = \cos(\pi x)$. After semi-discretization with $E(x_i,t) \approx E_i(t)$ and $B(x_i,t) \approx B_i(t)$, $i = 0, \ldots, N$, we get a system of ODEs

$$\dot{U} = \bar{S}DU, \quad U(0) = U_0, \tag{22}$$

where $U = [E_1, ..., E_{N-1}, B_0, ..., B_N]^T$ and

$$\bar{S} = \frac{1}{2h} \begin{bmatrix} 0_{N-1,N+1} & G \\ & \\ -G^T & 0_{N+1,N-1} \end{bmatrix}, \qquad G = \begin{bmatrix} -2 & 0 & 1 & & & \\ & -1 & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 0 & 1 \\ & & & & -1 & 0 & 2 \end{bmatrix}$$

and $D = \text{diag}(I_{N-1}, \frac{1}{2}, I_{N-1}, \frac{1}{2})$. Equation (22) fits the framework of section 3, with $\bar{S}$ skew-symmetric and D symmetric and positive definite, therefore APMH can be applied to this problem. The numerical approximation of U obtained applying the APMH preserves the first integrals $\mathcal{H}_k(\bar{U})$ of Proposition 2. The tables about preservations of first integrals asre not reported here. In Figure 4b, we consider the maxwell equation with fixed and given initial value in (21) and also the restart technique is used. We observe that the energy is preserved well.

5.3 Numerical results for 3D Maxwell's equations

We consider 3D Maxwell's equations in CGS units for the electromagnetic field in a vacuum

$$\begin{aligned} \partial_t E &= -c\nabla \times B, \\ \partial_t B &= c\nabla \times E. \end{aligned} \tag{23}$$

The boundary conditions are zero and the initial conditions are randomly generated for both fields. We consider $c = 1$. We get the following Hamiltonian system after semi-discretization:

$$\dot{U} = AU, \qquad U(0) = U_0, \qquad A = \begin{bmatrix} 0 & -G_1 \\ G_1 & 0 \end{bmatrix}, \qquad (24)$$

where $U = [E^x_{1,1,1}, ..., E^z_{N-1,N-1,N-1}, B^x_{1,1,1}, ..., B^z_{N-1,N-1,N-1}]^T$ and G_1, symmetric and of the size $3(N-1)^3$, is the discretization of the curl operator $\nabla\times$.

Remark 3 The matrix A is skew-symmetric in equation (24). Therefore the APMH with $J = A, H = I$ applied to the system (24), equals the APM and preserves the first integrals $\mathcal{H}_k(\bar{U})$ of Proposition 2.

Remark 4 Equation (24) can be rewritten as a Hamiltonian equation $\dot{U} = JHU$, with $H = J^{-1}A$ a symmetric matrix. Therefore we can also apply SLPM and BJPM to system (23) and the energy $\mathcal{H}(U) = \frac{1}{2}U^T J^{-1} AU$ is preserved. However, APMH cannot be used here because H is not a positive definite matrix, and the inner product $\langle\cdot,\cdot\rangle_H$ is degenerate. This can lead to instabilities and both global error and energy error might blow up during the iteration.
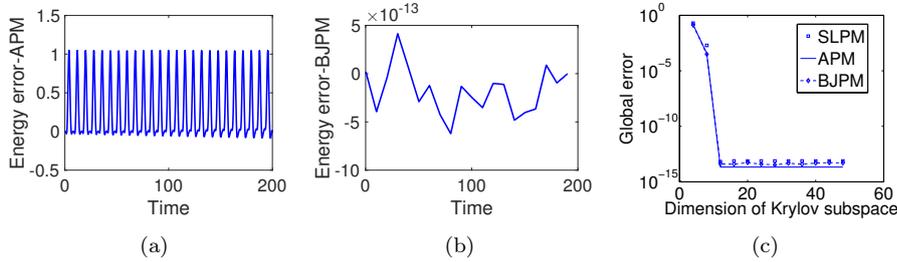

(a) (b) (c)

Fig. 5: The dimension of the problem is 384, namely $N = 5$. In Figure 5b the methods with the restart technique are used. Figure 5a corresponds to the energy error considered as in Remark 3, while figure 5b to the energy error considered in Remark 4. In Figure 5c we consider L^2 norm of the global error at $t = T = 2$ as a function of the dimension of the Krylov subspace.

Figure 5a shows that the energy error of APM is bounded as stated in Remark 3.

Figure 5b shows that the energy $\mathcal{H}(U) = \frac{1}{2}U^T J^{-1} AU$ is preserved for BJPM as stated in Remark 4. In Figure 5c, we report convergence plots for the methods. As the dimension of the Krylov subspace increases, the global error decreases very fast for all the methods. All the methods converge well also for larger end time, such as $T = 200$. Also in this example we observed a small linear growth in the error of the first integrals, due to the propagation of round-off errors (figures are not presented here).

Acknowledgment

References

1. A. Van der Schaft, D. Jeltsema, Port-Hamiltonian systems theory: An introductory overview, Foundations and Trends in Systems and Control **1**(2-3), 173 (2014)
2. K. Feng, M.z. Qin, in *Numerical Methods for Partial Differential Equations* (Springer, 1987), pp. 1–37
3. R. McLachlan, Symplectic integration of Hamiltonian wave equations, Numerische Mathematik **66**(1), 465 (1993)
4. J.E. Marsden, A. Weinstein, The Hamiltonian structure of the Maxwell-Vlasov equations, Physica D: Nonlinear Phenomena **4**(3), 394 (1982)
5. Y. Sun, P. Tse, Symplectic and multisymplectic methods for Maxwell's equations, Journal of Computational Physics **230**(5), 2076 (2010). DOI 10.1016/j.jcp.2010.12.006
6. E. Taylor Michael, *Partial Differential Equations I. Basic Theory*, vol. 115 (Springer-Verlag New York, 2011)
7. R.D. Richtmyer, K.W. Morton, *Difference Methods for Initial-value Problems* (Interscience Publishers John Wiley & Sons, Inc., Academia Publishing House of the Czechoslovak Acad, 1967)
8. R.A. LaBudde, D. Greenspan, Energy and momentum conserving methods of arbitrary order for the numerical integration of equations of motion, Numerische Mathematik **25**(4), 323 (1975)
9. R.I. McLachlan, G. Quispel, N. Robidoux, Geometric integration using discrete gradients, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **357**(1754), 1021 (1999)
10. L. Brugnano, F. Iavernaro, D. Trigiante, Hamiltonian boundary value methods (energy preserving discrete line integral methods), Journal of Numerical Analysis, Industrial Applied Mathematics **5**(1), 17 (2010)
11. E. Celledoni, V. Grimm, R.I. McLachlan, D. McLaren, B. Owren, D. O'Neale, G. Quispel, Preserving energy resp. dissipation in numerical PDEs using the "Average Vector Field" method, Journal of Computational Physics **231**(20), 6770 (2012)
12. Z. Ge, J. Marsden, Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators, Physics Letter A **133**, 134 (1988)
13. M.A. Botchev, J.G. Verwer, Numerical integration of damped Maxwell equations, SIAM Journal on Scientific Computing **31**(2), 1322 (2009)
14. T. Eirola, A. Koskela, Krylov integrators for hamiltonian systems, BIT Numerical Mathematics pp. 1–20 (2018). DOI 10.1007/s10543-018-0732-y. URL https://doi.org/10.1007/s10543-018-0732-y
15. L. Lopez, V. Simoncini, Preserving geometric properties of the exponential matrix by block Krylov subspace methods, BIT Numerical Mathematics **46**(4), 813 (2006)
16. A. Archid, A.H. Bentbib, Approximation of the matrix exponential operator by a structure-preserving block Arnoldi-type method, Applied Numerical Mathematics **75**, 37 (2014). DOI 10.1016/j.apnum.2012.11.008. URL https://doi.org/10.1016/j.apnum.2012.11.008
17. P. Benner, V. Mehrmann, H. Xu, A numerically stable structure-preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils, Numerische Mathematik **78**(3), 329 (1998)
18. V.I. Arnol'd, B. Dubrovin, A. Kirillov, I. Krichever, *Dynamical Systems IV: Symplectic Geometry and Its Applications*, vol. 4 (Springer Science & Business Media, 2001)
19. E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration: Structure-preserving Algorithms for Ordinary Differential Equations*, vol. 31 (Springer Science & Business Media, 2006)
20. W.E. Arnoldi, The principle of minimized iterations in the solution of the matrix eigenvalue problem, Quarterly of Applied Mathematics **9**(1), 17 (1951)

21. P. Benner, H. Faßbender, M. Stoll, A Hamiltonian Krylov–Schur-type method based on the symplectic Lanczos process, Linear Algebra and its Applications **435**(3), 578 (2011)
22. S. Lall, P. Krysl, J.E. Marsden, Structure-preserving model reduction for mechanical systems, Physica D: Nonlinear Phenomena **184**(1), 304 (2003)
23. E. Celledoni, L. Li, in *Proceedings of the ECMI conference* (2016), pp. 663–559
24. P. Benner, H. Faßbender, An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem, Linear Algebra and Its Applications **263**, 75 (1997)
25. D.S. Watkins, On Hamiltonian and symplectic Lanczos processes, Linear Algebra and Its Applications **385**, 23 (2004)
26. H. Faßbender, A detailed derivation of the parameterized SR algorithm and the symplectic Lanczos method for Hamiltonian matrices, Preprint (2006)
27. S. Agoujil, A. Bentbib, A. Kanber, A structure preserving approximation method for hamiltonian exponential matrices, Applied Numerical Mathematics **62**(9), 1126 (2012)
28. H. Goldstein, *Classical Mechanics* (Pearson Education India, 2011)
29. P. Benner, V. Mehrmann, D.C. Sorensen, *Dimension reduction of large-scale systems*, vol. 35 (Springer, 2005)
30. R.C. Ward, Numerical computation of the matrix exponential with accuracy estimate, SIAM J. on Num. Anal. **14**, 600 (1977)

6 Appendix

6.1 Algorithms

(a) Arnoldi's algorithm with modified inner product

1: Input: a matrix $J \in \mathbb{R}^{m \times m}$, $H \in \mathbb{R}^{m \times m}$, a vector $b \in \mathbb{R}^m$, a number $n \in \mathbb{N}$ and a tolerance $\iota \in \mathbb{R}$.
2: $A = JH$
3: $v_1 = \dfrac{b}{\langle b, b \rangle_H^{\frac{1}{2}}}$
4: **for** $j = 1 : n$ **do**
5:     compute $w_j = Av_j$
6:     **for** $k = 1 : 2$ **do**
7:         **for** $i = 1 : j$ **do**
8:             $h_{i,j} = \langle v_i, w_j \rangle_H$
9:             $w_j = w_j - h_{i,j} v_i$
10:         **end for**
11:     **end for**
12:     $h_{j+1,j} = \langle w_j, w_j \rangle_H^{\frac{1}{2}}$
13:     **if** $h_{j+1,j} < \iota$ **then**
14:         Stop
15:     **end if**
16:     $v_{j+1} = w_j / h_{j+1,j}$
17: **end for**
18: Output: $T_n, V_n, v_{n+1}, h_{n+1,n}$.

(b) Algorithm to generate V_n (by QR factorization)

1: Matrix $A \in \mathbb{R}^{2m \times 2m}$, vector $b \in \mathbb{R}^{2m}$, number $n \in \mathbb{N}$.
2: $v = b$
3: $K_n = v$
4: **for** $i = 1 : n - 1$ **do**
5:     $v = Av$
6:     $K_n = [K_n, v]$
7: **end for**
8: $K_n^q = K_n(1 : m, :)$
9: $K_n^p = K_n(m + 1 : 2m, :)$
10: $[Q, R] = qr([K_n^q, K_n^p])$
11: $V_n = Q(:, 1 : k), \quad k = \text{rank}([K_n^q, K_n^p]) \leq 2n$
12: Output V_n.