APPROXIMATING THE EXPONENTIAL
FROM A LIE ALGEBRA TO A LIE GROUP

ELENA CELLEDONI AND ARIEH ISERLES

ABSTRACT. Consider a differential equation $y' = A(t,y)y$, $y(0) = y_0$ with $y_0 \in G$ and $A : \mathbb{R}^+ \times G \to \mathfrak{g}$, where $\mathfrak{g}$ is a Lie algebra of the matricial Lie group G. Every $B \in \mathfrak{g}$ can be mapped to G by the matrix exponential map $\exp(tB)$ with $t \in \mathbb{R}$.

   Most numerical methods for solving ordinary differential equations (ODEs) on Lie groups are based on the idea of representing the approximation y_n of the exact solution $y(t_n)$, $t_n \in \mathbb{R}^+$, by means of exact exponentials of suitable elements of the Lie algebra, applied to the initial value y_0. This ensures that $y_n \in G$.

   When the exponential is difficult to compute exactly, as is the case when the dimension is large, an approximation of $\exp(tB)$ plays an important role in the numerical solution of ODEs on Lie groups. In some cases rational or polynomial approximants are unsuitable and we consider alternative techniques, whereby $\exp(tB)$ is approximated by a product of simpler exponentials.

   In this paper we present some ideas based on the use of the Strang splitting for the approximation of matrix exponentials. Several cases of $\mathfrak{g}$ and G are considered, in tandem with general theory. Order conditions are discussed, and a number of numerical experiments conclude the paper.

1. INTRODUCTION

Consider the differential equation

$$(1.1) \qquad y' = A(t,y)y, \qquad y(0) \in G,$$

with

$$A : \mathbb{R}^+ \times G \to \mathfrak{g},$$

where G is a matricial Lie group and $\mathfrak{g}$ is the underlying Lie algebra. In this case $y(t) \in G$, and for sufficiently small t it can be written in the form

$$y(t) = e^{\sigma(t)}y(0),$$

where σ is a suitable function, $\sigma(t) \in \mathfrak{g}$. A Lie group is a differentiable manifold endowed with group structure. The Lie algebra $\mathfrak{g}$ of G, its tangent space at the identity, is a linear space, endowed with an anti-symmetric binary operation (called commutation) $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ and a natural mapping $\exp : \mathfrak{g} \to G$. In the case of matricial Lie groups and algebras the commutation is $[A, B] := AB - BA$, the standard matrix commutator, and exp is the matrix exponential.

Differential equations on Lie groups are well known as models of several problems arising in physics. Some of them can be directly written as differential equations on Lie groups; others are given with respect to manifolds that are invariant under the transitive action of a Lie group (homogeneous spaces).

Problems of mechanical systems and robotics are modeled using the special orthogonal group $\mathrm{SO}(m)$, problems of volume conservation and fluid dynamics can be expressed in terms of the special linear group $\mathrm{SL}(m)$, and the conservation of the symplectic form in Hamiltonian systems involves the symplectic group $\mathrm{Sp}(m)$.

Numerical methods considered here are designed so that approximations y_n of the solution $y(t_n)$, $t_n \in \mathbb{R}^+$, stay inside G. For this reason, here and in the remainder of the paper we call them Lie-group methods. In [10] the authors show how Lie-group methods can be extended to homogeneous spaces by means of a transitive action of a Lie group.

The solution of (1.1), $y(t) = \mathrm{e}^{\sigma(t)} y_0$, is such that $\sigma(t) \in \mathfrak{g}$ obeys the equation

$$\sigma' = A + \tfrac{1}{2}[A, \sigma] + \tfrac{1}{12}[[A, \sigma], \sigma] - \tfrac{1}{720}[[[[A, \sigma], \sigma], \sigma], \sigma] + \cdots$$

with $\sigma(0) = 0$.

Given a suitable approximation to σ in the Lie algebra, exponentiation produces an approximation of $y(t)$ in the Lie group G. That is why virtually all Lie group methods are realized by writing the approximation y_n of the solution $y(t_n)$, $t_n \in \mathbb{R}^+$, by means of exact exponentials of suitable elements of the Lie algebra, applied to the initial value y_0. Significant examples of such techniques are given by the method of Crouch and Grossman [2], based on rigid frames, the Runge Kutta-type methods of Munthe-Kaas [9], the method of iterated commutators of Zanna [18] and the method of Magnus series of Iserles and Nørsett [7].

All these methods need to evaluate (or suitably approximate) the matrix exponential once (and often repeatedly) in each time step. This can be a very challenging task when the dimension is large.

Suppose that we embed the equation (1.1) in $\mathbb{R}^{m \times m}$. Conventional methods to approximate $\exp(tB)$ may lead to a final approximation, $\tilde{y}_n \approx y_n$, that departs from G. Hence, such methods may obviate the advantages in the conservation of invariants and symmetries that have motivated Lie-group methods at the first place. The aim of this paper is to present and study approximations $F(tB) \approx \exp(tB)$ such that $\tilde{y}_n \in \mathrm{G}$, in other words to ensure that $F : \mathfrak{g} \to \mathrm{G}$.

Standard approaches for large matrices, e.g. polynomial or rational approximants, do not guarantee that in general $\tilde{y}_n$ belongs to a Lie group. Polynomial approximants are typically based on the construction of approximations of $\exp(tB)v$ ($v \in \mathbb{R}^m$) belonging to some Krylov subspace $K_n(B, v) = \mathrm{span}\{v, Bv, \ldots, B^{n-1}v\}$. If v is the i-th column of some matrix $V \in \mathrm{G}$ and $B \in \mathfrak{g}$, then $p_{n-1}(tB)V \in K_n(B, V) \cap \mathrm{G}$ if and only if $p_{n-1}(tB) \in \mathrm{G}$, but in general it is not true that $p_{n-1} : \mathfrak{g} \to \mathrm{G}$. Rational approximants, e.g. Padé approximants, replace the exponential by a suitable rational function: again, there is no guarantee that the outcome lies in G. An important category of Lie groups lends itself to suitable rational approximations. This follows from the following theorem on analytic approximants.

Theorem 1. *Consider the Lie group* $\mathrm{G} := \{y \in \mathrm{GL}(m) : yPy^{\mathrm{T}} = P\}$, *where P is a nonsingular $m \times m$ matrix and the corresponding Lie algebra* $\mathfrak{g} = \{x \in \mathfrak{gl}(m) : xP + Px^{\mathrm{T}} = 0\}$. *Let ϕ be a function analytic in a neighbourhood $\mathcal{U}_0$ of 0, with*

$\phi(0) = 1$ *and* $\phi'(0) = 1$. *If*

(1.2) $$\phi(z)\phi(-z) \equiv 1,$$

$\forall z \in \mathcal{U}_0$, *then*

(1.3) $$\forall x \in \mathfrak{g}, \quad \phi(tx) \in \mathrm{G}$$

for $t \in \mathbb{R}^+$ *sufficently small. Moreover if* $\exists\, x \in \mathfrak{g}$ *such that* $\sigma(x) \neq \{0\}$, *where* $\sigma(x)$ *denotes the spectrum of* x, *then* (1.3) *implies* (1.2).

Proof. Let ϕ be a function analytic in a neighbourhood $\mathcal{U}_0$ of 0, with $\phi(0) = 1$ and $\phi'(0) = 1$. Given $x \in \mathfrak{g}$, the condition $\phi(tx) \in \mathrm{G}$ means that

$$\phi(tx)P\phi(tx^{\mathrm{T}}) = P;$$

hence, by virtue of the analyticity of ϕ,

$$\phi(tx)\phi(tPx^{\mathrm{T}}P^{-1}) = I.$$

Note now that $x \in \mathfrak{g}$ if and only if $Px^{\mathrm{T}}P^{-1} = -x$, and hence

(1.4) $$\phi(tx)\phi(-tx) = I, \quad \forall x \in \mathfrak{g},$$

i.e. (1.4) is equivalent to (1.3). Since (1.2) implies (1.4), the first part of the theorem is proved. For proving the converse, suppose that (1.3), i.e. (1.4), holds. By considering the Schur decomposition of x it is easy to realize that

$$\lambda \in \sigma(x) \Rightarrow \phi(t\lambda)\phi(-t\lambda) = 1$$

for $t\lambda \in \mathcal{U}_0$. Suppose that $\exists\, x \in \mathfrak{g}$ such that $\exists \lambda \in \sigma(x)$, $\lambda \neq 0$. Then $\exists \tilde{t} > 0$, $\tilde{t} \in \mathbb{R}$, such that $t \in [0, \tilde{t}] \Rightarrow t\lambda \in \mathcal{U}_0$ and $\phi(t\lambda) \neq 1$ for $t \neq 0$. Then $\phi(z)\phi(-z) = 1$ for $z = t\lambda$ and $t \in [0, \tilde{t}]$; but then

$$\phi(z)\phi(-z) = 1$$

$\forall z \in \mathcal{U}_0$. $\square$

Concrete examples of such Lie groups are for instance $\mathrm{O}(m)$, in which case $P = I$, and $\mathrm{SP}(m)$, in which case

$$P = \begin{bmatrix} O & I \\ -I & O \end{bmatrix}$$

In the complex case, replacing the transpose with the conjugate transpose, the same result holds for $\mathrm{G} = \mathrm{U}(m)$. Diagonal Padé approximants are analytic functions that fulfill (1.2).

On the other hand this result is not true for every subgroup of $\mathrm{GL}(m)$, and in particular the following theorem is relevant to the case of the special linear algebra and group.

Theorem 2 ([3]). *The only function ϕ analytic in a neighbourhood of 0 with $\phi(0) = 1$ and $\phi'(0) = 1$, such that $\forall x \in \mathfrak{sl}(m)$ $\phi(tx) \in \mathrm{SL}(m)$, $m \geq 3$, for $t \in \mathbb{R}^+$ sufficently small, is the exponential map.*

Proof. Our proof differs from [3]. We show that there exists $D \in \mathfrak{sl}(m)$ such that $\phi(tD) \in \mathrm{SL}(m)$ only if ϕ is the exponential map. Consider a diagonal matrix D

belonging to $\mathfrak{sl}(m)$ but such that $D^k \notin \mathfrak{sl}(m)$ for $k \geq 2$. Then $\phi(tD) \in \mathrm{SL}(m)$ if and only if

$$(1.5) \qquad \det(\phi(tD)) = \prod_{i=1}^{m} \phi(td_i) \equiv 1,$$

where d_i for $i = 1, \ldots, m$ are the eigenvalues of D. Let us consider the function $\psi(z) := \log \phi(z)$, which is analytic in a neighbourhood of the origin, and let $\psi(z) = \sum_{k=0}^{\infty} \psi_k z^k$. Now from (1.5)

$$(1.6) \qquad \sum_{i=1}^{m} \psi(td_i) = 0.$$

Consequently

$$\sum_{k=0}^{\infty} \psi_k t^k \left(\sum_{i=1}^{m} d_i^k \right) = 0.$$

Then, since $\sum_{i=1}^{m} d_i = 0$ but $\mathrm{tr} D^k = \sum_{i=1}^{m} d_i^k \neq 0$ for $k \geq 2$, necessarily $\psi_k = 0$ for $k \geq 2$ and $\psi(z) = \psi_1 z$. This implies $\phi(z) = \exp(\psi_1 z)$. Moreover, $\phi'(0) = 1$ implies $\psi_1 = 1$. $\square$

It turns out that, for $B \in \mathfrak{sl}(m)$ any approximation of e^{tB} should be expressed by means of exact exponentials of elements in $\mathfrak{sl}(m)$ in order to belong to $\mathrm{SL}(m)$.

Note that the above restriction does not hold for every subgroup of $\mathrm{SL}(m)$, an example being the subgroup $\mathrm{SO}(m)$, whose elements are the orthogonal matrices with determinant equal to 1, since for any function ϕ fulfilling the hypothesis of Theorem 1, continuity of the determinant implies that $\phi(x) \in \mathrm{SO}(m) \ \forall x \in \mathfrak{so}(m)$.

In the following we will consider the idea of splitting the matrix $B \in \mathfrak{g}$ into a sum of low-rank matrices B_i belonging to $\mathfrak{g}$ for the approximation of e^{tB}.

Splitting methods have been considered by various authors in different contexts: for constructing symplectic methods [11], for constructing volume-preserving algorithms [3], and in the PDE context [12]. An extensive survey of the theory of splitting methods can be found in [8].

The idea is as follows: given an $m \times m$ matrix $B \in \mathfrak{g}$, we split it in the form

$$(1.7) \qquad B = \sum_{i=1}^{k} B_i,$$

the $\exp(tB)$ is then approximated by $F(tB)$, and

$$F(tB) \approx \exp(tB_1) \cdots \exp(tB_k).$$

In general the computation of each of the exponential terms $\exp(tB_i)$ for $i = 1, \ldots, k$ implies the same cost of computing $\exp(tB)$ directely. For general splittings (1.7) this procedure can hardly be competitive with direct calculations of $\exp(tB)$. However, we will show in this paper that splitting $B \in \mathfrak{g}$ into appropriate low-rank terms B_i belonging to $\mathfrak{g}$, such that the exponentials $\exp(tB_i)$ are easy to compute exactly and can be cheaply multiplied together, makes this approach competitive with classical techniques. The computational costs that we obtain are of the type Cm^3, where C is a constant depending on the order of the approximation and m is the dimension of B. Standard classical techniques such as diagonal Padé approximants normally require computational costs of the same order of magnitude (m^3), with different constants C depending on the order required.

However, the main motivation of our work is obtaining approximations to the exponential that, given a matricial Lie group G and its corresponding Lie algebra $\mathfrak{g}$, map elements of a neiborhood of the origin of $\mathfrak{g}$ to a neighborhood of the identity of G. Low-rank splitting methods fulfill this requirement, because, provided that $B_i \in \mathfrak{g}$, the approximation $F(tB)$ resides in the Lie group G.

If $\mathfrak{g} = \mathfrak{sl}(m)$ and $m \geq 3$, diagonal Padé approximants can be used only if the approximation is accurate to machine epsilon. The `Matlab` function *expm* implements a diagonal Padé approximation of exp with a scaling and squaring technique; running this function on randomly generated matrices, it is possible to verify that obtaining an approximation of $\exp(tB)$ to machine accuracy yields C in the range between 20 and 30, depending on the eigenvalues of B. In the case of the approximation of $\exp(tB)v$ and $v \in \mathbb{R}^m$ the costs of low-rank splitting methods drop down to Cm^2 and are then comparable to those obtained by polynomial approximants [6], [5]. In Section 5 we present the values of the constants we obtain with the techniques proposed in this paper for different orders. Under the above considerations the values obtained confirm that it is cheaper to use such approximations instead of the built-in function of `Matlab` when machine accuracy is not required in the final approximation.

In Section 2 the problem of how to perform the decomposition of $B \in \mathfrak{g}$ into low-rank matrices $B_i \in \mathfrak{g}$ will be considered in the cases $\mathfrak{g} = \mathfrak{so}(m)$ and $\mathfrak{g} = \mathfrak{sl}(m)$. In Section 3 we will discuss the order conditions and consider the so-called Strang splitting, which is the optimal second-order splitting within our framework. In Section 4 techniques for increasing the order of the methods will be discussed. In Section 5 some numerical experiments are presented.

2. LOW-RANK SPLITTINGS

Given an $m \times m$ matrix $B \in \mathfrak{g}$, we split it in the form

$$B = \sum_{i=1}^{k} B_i,$$

such that

1. $B_i \in \mathfrak{g}$, for $i = 1, \dots, k$.
2. Each $\exp(tB_i)$ is easy to evaluate exactly.
3. Products of such exponentials are cheap.

We approximate

$$(2.8) \qquad \exp(tB) \approx F(tB) = \exp(tB_1)\exp(tB_2) \cdots \exp(tB_k).$$

Similar types of approximation have been considered in [11], [13], [12], although not in the context of Lie group methods. It is well known that (2.8) is an approximation of order one. In [13], [12], order analysis is developed extensively. In Section 4 our observations on the order of the methods follow a similar path. Given a splitting $B = \sum_{i=1}^{k} B_i$, for obtaining approximations of order higher then one, in general the number of exponentials to be multiplied together is $r > k$.

Suppose now that $C = B_i$ is a low rank matrix, i.e.

$$(2.9) \qquad C = \sum_{l=1}^{p} \boldsymbol{\alpha}_l \boldsymbol{\beta}_l^{\mathrm{T}} = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}},$$

with $\boldsymbol{\alpha}_l, \boldsymbol{\beta}_l \in \mathbb{R}^m$, where $p \geq 1$ is a small integer and $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_p]$, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p]$ are $m \times p$ matrices.

Definition 1. Let $r_C := \text{rank}(C)$, $C \in \text{M}_{m,m}[\mathbb{R}]$. We say that the decomposition (2.9) of C is optimal if $p = r_C$.

The decomposition of C is optimal if and only if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are full-rank matrices. The following proposition gives an explicit formula for computing $\exp(tC)$.

Proposition 3. Setting $D = \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\alpha}$ and $C = \boldsymbol{\alpha} \boldsymbol{\beta}^{\mathrm{T}}$, the function $\exp(tC)$ can be calculated by the explicit formula

$$(2.10) \qquad\qquad \exp(tC) = I + t \boldsymbol{\alpha} \varphi(tD) \boldsymbol{\beta}^{\mathrm{T}},$$

where $\varphi(z) := (\mathrm{e}^z - 1)/z$ is an analytic function.

Proof. Note that for $n \geq 1$, $C^n = (\boldsymbol{\alpha} \boldsymbol{\beta}^{\mathrm{T}})^n = \boldsymbol{\alpha} (\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\alpha})^{n-1} \boldsymbol{\beta}^{\mathrm{T}} = \boldsymbol{\alpha} D^{n-1} \boldsymbol{\beta}^{\mathrm{T}}$. Then, since $\varphi(z) = \sum_{n=1}^{\infty} \frac{1}{n!} z^{n-1}$, we have

$$\exp(tC) = I + \boldsymbol{\alpha} \sum_{n=1}^{\infty} \frac{1}{n!} t^n D^{n-1} \boldsymbol{\beta}^{\mathrm{T}} = I + t \boldsymbol{\alpha} \varphi(tD) \boldsymbol{\beta}^{\mathrm{T}}.$$

$\square$

Note that $C = \boldsymbol{\alpha} \boldsymbol{\beta}^{\mathrm{T}}$ is an optimal decomposition if and only if D is nonsingular, whence $t\varphi(tD) = D^{-1}(e^{tD} - I)$. Under the hypothesis of Proposition 3, computing $\exp(tC)$ via the formula (2.10) is very cheap. We approximate each $\exp(tD)$ with a diagonal Padé approximant at the cost of $q \cdot p^3$ flops for a $p \times p$ matrix, where q is the degree of the polynomials in the rational function.

The cost of evaluating $\exp(tC)$ is

- $2p^2 m$ flops for computing D;
- $qp^3 + \frac{5}{3}p^3 + p$ flops for computing $D^{-1}(\exp(tD) - I)$;

and

- $2pm^2 + 2p^2 m + m$ flops for multiplying $\boldsymbol{\alpha}$, $D^{-1}(\exp(tD) - I)$ and $\boldsymbol{\beta}^{\mathrm{T}}$ and adding I.

In total, we get $2pm^2 + 4p^2 m + (q + 5/3)p^3 + p + m$ flops.

The cost of evaluating the product of r of such exponentials when every B_i can be expressed in the form (2.9) is

$$r[2p^2 m + (q + 5/3)p^3 + p] + 2p^2 m + 2pm^2 + m + (r-1)(m^2 + 4pm^2 + 2p^2 m)$$

$$= 4rmp^2 + 2(2r-1)pm^2 + (r-1)m^2 + m + r[(q + 5/3)p^3 + p].$$

We separate the cost of computing the first factor of the product. The factorization of $\exp(tC) - I$ given in terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in (2.10) allows significant saving in computations: the cost of multiplying an $m \times m$ matrix by $\exp(tB_i)$, given in the factored form (2.10), is $m^2 + 4pm^2 + 2p^2 m$ flops, which is lower than $2m^3$, the usual cost for a product of two $m \times m$ matrices.

Subject to the extra assumption that rp is a multiple of m, we can simplify the previous expression by considering only the leading term, and we get

$$2(2r-1)pm^2 + (r-1)m^2.$$

If $r = k$ and we take $kp = m$, this leads to a count of about $5m^3$ flops in the order-one approximation.

Note that in the case of differential equations whose solution evolves on an homogeneous space $\mathcal{M}$ of dimension m, subject to the transitive action of matricial Lie group $G \subset GL(m)$, the approximation of the exponential map is an easier computational issue, because in this case we are interested in finding $\exp(tB)v \approx F(tB)v$, where $v \in \mathcal{M}$ is m-dimensional. In this case the factors that, composed together, give $F(tB)$ can be applied directly to v at a cost of $m + 4pm + 2p^2$ flops, and in total

$$(2.11) \qquad r[2p^2 m + (q + 5/3)p^3 + p + m + 4pm + 2p^2]$$

flops, and the cost for $r = k$ and $kp = m$ drops down to about $5m^2$. In this context the computational costs of our approach are comparable to the costs obtained with polynomial approximations (e.g. Krylov subspace methods [6], [5]). In the following we will refer to the count of flops only in the former case, considering that the latter can be recovered by dividing by m.

In the next section we will show that it is possible to obtain even lower costs in the actual implementation of these methods.

We observe that the optimality both of the decomposition of the B_i's and of their sum B requires that kp is minimal.

In the following proposition we consider the problem of converting a nonoptimal decomposition in an optimal one.

Proposition 4. *Let $C = \alpha\beta^T$ be a decomposition of C of type* (2.9) *and suppose that $r_C = \text{rank}(C) < p$. It is possible to construct $\hat{\alpha}$ and $\hat{\beta}$ such that*

$$C = \hat{\alpha}\hat{\beta}^T$$

is an optimal decomposition.

Proof. We will prove the assertion by constructing $\hat{\alpha}$ and $\hat{\beta}$, distinguishing between two cases:

1. $r_\alpha := \text{rank}(\alpha) < p$;
2. $r_\alpha = p$ but $r_\beta := \text{rank}(\beta) < p$.

Case 1. Consider the factorization $\alpha = Q_\alpha R_\alpha$, where Q_α $m \times r_\alpha$ is a matrix with orthogonal columns and $R_\alpha = Q_\alpha^T \alpha$, is a $r_\alpha \times p$ matrix. Consider the $m \times r_\alpha$ matrix βR_α^T. If $\text{rank}(\beta R_\alpha^T) = r_\alpha$, then take $\hat{\alpha} := Q_\alpha$ and $\hat{\beta} := \beta R_\alpha^T$; otherwise redefine $\alpha := Q_\alpha$, $\beta := \beta R_\alpha^T$, $r_\beta := \text{rank}(\beta)$, and consider an analogous factorization for $\beta = Q_\beta R_\beta$ with Q_β an $m \times r_\beta$ matrix with orthogonal columns and $R_\beta = Q_\beta^T \beta$ an $r_\beta \times r_\alpha$ matrix. Now if $\text{rank}(\alpha R_\beta^T) = r_\beta$, take $\hat{\beta} := Q_\beta$ and $\hat{\alpha} := \alpha R_\beta^T$; otherwise redefine $\beta := Q_\beta$, $\alpha := \alpha R_\beta^T$, $r_\alpha := \text{rank}(\alpha)$, and repeat the procedure. It will terminate in a number of steps less than or equal to $r_C - r_\alpha$.

Case 2. Finally, if $r_\alpha := \text{rank}(\alpha) = p$, but $r_\beta := \text{rank}(\beta) < p$, start the process by factoring β, rather than α. $\qquad \square$

The proof of Proposition 4 is an algorithm for converting a nonoptimal decomposition of C into an optimal one. For computing the factorizations $\alpha = Q_\alpha R_\alpha$ and $\beta = Q_\beta R_\beta$ it is possible to use a modified Gram–Schmidt algorithm at the cost of mp^2 flops, or to use a Householder transformation at the cost of $2(mp^2 - p^3/3)$ flops. Subsequently, computing $\varphi(tD)$, it is possible to carry out such an algorithm and to use the inverse of D in (2.10), although, alternatively, one can consider a rational approximant for $\varphi(tD)$ or perform the spectral decomposition of D.

2.1. Splitting in a Lie algebra. The issue of this section is, given $B \in \mathfrak{g}$, to find a splitting of B into low-rank matrices $B_i \in \mathfrak{g}$. We give now an example of optimal and another of a nonoptimal decomposition of $B \in \mathfrak{so}[m, \mathbb{R}]$.

1. *Columnwise decomposition.* The simplest choice corresponds to $p = 2$, $k = m - 1$. Taking $B^{[0]} := B$, we let

$$B_1 = \boldsymbol{b}_1^{[0]} e_1^{\mathrm{T}} - e_1 \boldsymbol{b}_1^{[0]\mathrm{T}} \in \mathfrak{so}(m),$$

where $\boldsymbol{b}_1^{[0]}$ is the first column of $B^{[0]}$ and e_1 is the first vector of the canonical basis of $\mathbb{R}^m$; in general take $B^{[i]} := B^{[i-1]} - B_i$ and

$$B_{i+1} = \boldsymbol{b}_{i+1}^{[i]} e_{i+1}^{\mathrm{T}} - e_{i+1} \boldsymbol{b}_{i+1}^{[i]\mathrm{T}} \in \mathfrak{so}(m)$$

where $\boldsymbol{b}_{i+1}^{[i]}$ is the first nonzero column of $B^{[i]}$ and e_{i+1} is the $i + 1$-th vector of the canonical basis of $\mathbb{R}^m$. A different way of obtaining the same decomposition is

$$B_i = \frac{1}{2} \boldsymbol{b}_i e_i^{\mathrm{T}} - \frac{1}{2} e_i \boldsymbol{b}_i^{\mathrm{T}}, \qquad i = 1, 2, \ldots, m,$$

where $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ are the columns of B, and $k = m$ in this case.

Regardless of the sparsity of the vectors that define the decomposition, the computational costs given by (2.11), putting $r = k$ and $kp = 2m$, amount to $10m^3$. More careful analysis, based on the fact that inner products with respect to vectors of the canonical basis are very inexpensive and that the vectors $\boldsymbol{b}_{i+1}^{[i]}$ have $m - (i - 1)p/2$ nonzeros, results in the following count of flops for the product of the exponentials:

$$\sum_{i=1}^{k} \left\{ 2m[m - (i-1)p/2]p + 2p^2 m + m[m - (i-1)p/2] \right\}$$

After straightforward calculation this sums up to

$$(2p + 1)m[km - pk(k-1)/4] + 2kp^2 m.$$

For $kp = 2m$ we obtain a leading term of $2m^3 + m^3/p$ for the order-one method.

In the case $p = 2$ an extra saving in the multiplication by the $m \times 2$ matrix $\boldsymbol{\alpha}$ reduces the cost of the product to $2m^3$.

This decomposition is not optimal. Although in principle it is possible to convert it to an optimal one, it might be possible that the B_i's will no longer belong to $\mathfrak{g}$ in that case.

A similar strategy can be used for decomposing matrices $B \in \mathfrak{sl}(m)$. Taking $B^{[0]} := B - \mathrm{diag}(B)$, we let

$$B_1 = \boldsymbol{b}_1^{[0]} e_1^{\mathrm{T}} + e_1 \boldsymbol{a}_1^{[0]\mathrm{T}} \in \mathfrak{sl}(m),$$

where $\boldsymbol{b}_1^{[0]}$ is the first column of $B^{[0]}$, $\boldsymbol{a}_1^{[0]\mathrm{T}}$ is the first row and e_1 is the first vector of the canonical basis of $\mathbb{R}^m$; in general take $B^{[i]} := B^{[i-1]} - B_i$ and

$$B_{i+1} = \boldsymbol{b}_{i+1}^{[i]} e_{i+1}^{\mathrm{T}} + e_{i+1} \boldsymbol{a}_{i+1}^{[i]\mathrm{T}} \in \mathfrak{sl}(m),$$

where $\boldsymbol{b}_{i+1}^{[i]}$ is the first nonzero column of $B^{[i]}$, $\boldsymbol{a}_{i+1}^{[i]}$ is the first nonzero row and e_{i+1} is the $i + 1$-th vector of the canonical basis of $\mathbb{R}^m$. Then $B_m = \mathrm{diag}(B)$.

2. *Optimal decomposition.* More interesting is the case $p = 2$, $k = \lfloor m/2 \rfloor$. Subject to the assumption that $b_{2,1}^{[0]} = b_{2,1} \neq 0$, we commence by letting $B^{[0]} = B$ and

$$B_1 = \frac{1}{b_{2,1}^{[0]}} (\boldsymbol{b}_2^{[0]} \boldsymbol{b}_1^{[0]\,\mathrm{T}} - \boldsymbol{b}_1^{[0]} \boldsymbol{b}_2^{[0]\,\mathrm{T}}) \in \mathfrak{so}[m, \mathbb{R}].$$

Therefore

$$(B_1)_{k,l} = \frac{b_{2,k}^{[0]} b_{1,l}^{[0]} - b_{1,k}^{[0]} b_{2,l}^{[0]}}{b_{2,1}^{[0]}} \begin{cases} b_{1,l}, & k = 1, \\ b_{2,l}, & k = 2. \end{cases}$$

Next we set $B^{[1]} = B^{[0]} - B_1$ and denote its columns by $\boldsymbol{b}_1^{[1]}, \ldots, \boldsymbol{b}_n^{[1]}$. Note that $\boldsymbol{b}_1^{[1]} \boldsymbol{b}_2^{[1]} = \boldsymbol{0}$ and nonzero entries are confined to the lower $(m-2) \times (m-2)$ principal minor. We let

$$B_2 = \frac{1}{b_{4,3}^{[1]}} (\boldsymbol{b}_4^{[1]} \boldsymbol{b}_3^{[1]\,\mathrm{T}} - \boldsymbol{b}_3^{[1]} \boldsymbol{b}_4^{[1]\,\mathrm{T}}) \in \mathfrak{so}[m, \mathbb{R}]$$

and continue in a similar vein up to B_k, where $k = \lfloor m/2 \rfloor$. In this case the decomposition is optimal, but it could cause instability due to the division by the values $b_{2j,2j-1}^{[j-1]}$ for $j = 1, \ldots, k$. In order to overcome these problems the decomposition can be implemented using a pivot strategy that allows us, at every step, to divide by the element of largest absolute value.

The cost of computing such a decomposition can be estimated as about $m^3/2$ operations. The cost of computing the product of the exponentials resulting in $F(tB)$ is

$$\sum_{i=1}^{k} \left\{ 4m[m - (i-1)p]p + 2p^2 m + m[m - (i-1)p] \right\},$$

which leads to a leading term of $2kp^2m + (4p+1)m[mk - pk(k-1)/2]$, or, taking $kp = m$, $2m^3 + m^3/2p$. For example, $p = 2$ yields $2.25m^3$.

In the following section the optimal decomposition will be extended to the case of the Lie algebra $\mathfrak{sl}(m)$.

2.2. General bordered matrices. Let $B \in \mathfrak{so}[m, \mathbb{R}]$. In columnwise decomposition every matrix B_i reconstructs one row and one column of the matrix B, and two rows and columns in the optimal decomposition. We generalize these techniques to the case of p rows and columns for any $B \in \mathrm{M}_{m,m}[\mathbb{R}]$.

To this end we consider a bordered matrix

$$(2.12) \qquad\qquad C = \begin{bmatrix} E & F \\ G & O \end{bmatrix},$$

where $E \in \mathrm{M}_{r,r}[\mathbb{R}]$, $G \in \mathrm{M}_{m-r,r}[\mathbb{R}]$, $F \in \mathrm{M}_{r,m-r}[\mathbb{R}]$ and $O \in \mathrm{M}_{m-r,m-r}[\mathbb{R}]$. The interesting case is, of course, when r is considerably smaller than m. The matrix C might belong to a Lie algebra, although this will not play a central role in our discussion. If we rewrite C as a sum of rank-one matrices, we can apply formula (2.10) for computing $\exp(tC)$. Suppose that

$$G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, \qquad F = [\, F_1 \quad F_2 \,],$$

where $G_1, F_1 \in \mathrm{M}_{r,r}[\mathbb{R}]$.

Proposition 5. *Given a bordered matrix* (2.12), *if* $\det(F_1) \neq 0$ *it is possible to decompose*

$$C = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}} = \sum_{l=1}^{2r} \boldsymbol{\alpha}_l\boldsymbol{\beta}_l^{\mathrm{T}},$$

where $\boldsymbol{\alpha}_l$ *and* $\boldsymbol{\beta}_l$, $l = 1, 2, \ldots, 2r$, *are the columns of* $\boldsymbol{\alpha}$ *and of* $\boldsymbol{\beta}$, *respectively.*

Proof. Supposing that

$$\boldsymbol{\alpha} = \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \bar{\boldsymbol{\alpha}} \end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \bar{\boldsymbol{\beta}} \end{bmatrix},$$

where $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}} \in \mathrm{M}_{2r,2r}[\mathbb{R}]$, the proposition will follow from (for $\det \hat{\boldsymbol{\beta}} \neq 0$)

$$\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\beta}}^{\mathrm{T}} = \begin{bmatrix} E & F_1 \\ G_1 & O \end{bmatrix}, \qquad \hat{\boldsymbol{\alpha}}\bar{\boldsymbol{\beta}}^{\mathrm{T}} = \begin{bmatrix} F_2 \\ O \end{bmatrix},$$

$$\bar{\boldsymbol{\alpha}}\hat{\boldsymbol{\beta}}^{\mathrm{T}} = \begin{bmatrix} G_2 & O \end{bmatrix}, \qquad \bar{\boldsymbol{\alpha}}\bar{\boldsymbol{\beta}}^{\mathrm{T}} = O.$$

We let

$$\hat{\boldsymbol{\alpha}} = \begin{bmatrix} E & F_1 \\ G_1 & O \end{bmatrix} \hat{\boldsymbol{\beta}}^{-T},$$

$$\begin{bmatrix} E & F_1 \\ G_1 & O \end{bmatrix} \hat{\boldsymbol{\beta}}^{-T}\bar{\boldsymbol{\beta}}^{\mathrm{T}} = \begin{bmatrix} F_2 \\ O \end{bmatrix}, \qquad \hat{\boldsymbol{\beta}}^{-T}\bar{\boldsymbol{\beta}}^{\mathrm{T}} = \begin{bmatrix} O \\ F_1^{-1}F_2 \end{bmatrix}$$

$$\bar{\boldsymbol{\alpha}} = \begin{bmatrix} G_2 & O \end{bmatrix} \hat{\boldsymbol{\beta}}^{-T},$$

where the nonsingular matrix $\hat{\boldsymbol{\beta}}$ is arbitrary. An obvious choice is $\hat{\boldsymbol{\beta}} = I$. $\square$

Note an alternative course of action, namely to LU-factorize:

$$\begin{bmatrix} E & F_1 \\ G_1 & O \end{bmatrix} = LU$$

and set $\hat{\boldsymbol{\alpha}} = L$, $\hat{\boldsymbol{\beta}}^{\mathrm{T}} = U$, whence

$$(2.13) \qquad \boldsymbol{\alpha} = \begin{bmatrix} L \\ \begin{bmatrix} G_2 & O \end{bmatrix} U^{-1} \end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix} U^{\mathrm{T}} \\ L^{-1}\begin{bmatrix} F_2 \\ O \end{bmatrix} \end{bmatrix}$$

In this case, however, the decomposition of C is not optimal.

1. *Columnwise decomposition.* When we consider the decomposition of the matrix $B = \sum_{i=1}^{k} B_i$ such that every B_i is a bordered matrix, then putting $B_i = C$ and writing C as a sum of rank-one matrices for $i = 1, \ldots, k$, we obtain a nonoptimal decomposition of B as a sum of rank-one matrices. For $r = 1$ we have a generalization to the case of $B \in \mathrm{M}_{m,m}[\mathbb{R}]$ of the columnwise decomposition described in the previous section.

2. *Optimal decomposition.* An alternative way of proceeding in order to obtain an optimal decomposition for B is the following. Suppose that

$$B = \begin{bmatrix} E & F \\ G & H \end{bmatrix} = C + \begin{bmatrix} O & O \\ O & H \end{bmatrix},$$

and $\det E \neq 0$. Defining $B^{[0]} = B$, we consider

$$\boldsymbol{\alpha} = \begin{bmatrix} E \\ G \end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix} I \\ (E^{-1}F)^{\mathrm{T}} \end{bmatrix},$$

where $\alpha, \beta \in M_{m,r}[\mathbb{R}]$. Thus

$$B_1 = \alpha\beta^{\mathrm{T}} = \left[\begin{array}{cc} E & F \\ G & GE^{-1}F \end{array} \right],$$

and taking $B^{[1]} = B^{[0]} - B_1$ we have

$$B^{[1]} = \left[\begin{array}{cc} O & O \\ O & S \end{array} \right],$$

with $S := H - GE^{-1}F \in M_{m-r,m-r}[\mathbb{R}]$, so that we obtain an analogous problem of reduced dimension that we can reduce further using the same technique.

The case $r = 2$ corresponds to a generalization to the case of $B \in M_{m,m}[\mathbb{R}]$ of the optimal decomposition described in the previous section.

Note finally that if $B \in \mathfrak{so}[m, \mathbb{R}]$, then $GE^{-1}F \in \mathfrak{so}[m, \mathbb{R}]$; therefore also $B^{[1]} \in \mathfrak{so}[m, \mathbb{R}]$. Moreover, as we will see in the next section, in the case of $B \in \mathfrak{sl}[m, \mathbb{R}]$ the condition $B_i \in \mathfrak{sl}[m, \mathbb{R}]$ is not necessary for $\exp(B_1) \cdots \exp(B_k) \in \mathrm{SL}(m)$.

A different way of addressing the problem is to compute $\exp(tC)$ directly, using the technique based on the following observations.

Proposition 6. *Consider the bordered matrix C (2.12), define the $(2r) \times (2r)$ matrix*

$$P = \left[\begin{array}{cc} O & I \\ FG & E \end{array} \right],$$

and suppose that its exponential is

$$e^{tP} = \left[\begin{array}{cc} \tilde{\Xi}(t) & \tilde{\Sigma}(t) \\ \tilde{\Phi}(t) & \tilde{\Psi}(t) \end{array} \right].$$

Then if $\det(FG) \neq 0$, the exponential of C is

$$e^{tC} = \left[\begin{array}{cc} \Xi(t) & \Sigma(t) \\ \Phi(t) & \Psi(t) \end{array} \right],$$

where $\Xi = \tilde{\Psi}$, $\Sigma = \tilde{\Sigma}F$, $\Phi = G\tilde{\Sigma}$ and $\Psi = I + G(\tilde{\Psi} - I)(FG)^{-1}F$.

Proof. We denote powers and the exponential of C by

$$C^k = \left[\begin{array}{cc} \Gamma_k & \Theta_k \\ \Delta_k & \Omega_k \end{array} \right], \quad k \in \mathbb{Z}^+, \qquad e^{tC} = \left[\begin{array}{cc} \Xi(t) & \Sigma(t) \\ \Phi(t) & \Psi(t) \end{array} \right].$$

Multiplying C^k by C on the right, we derive the recurrences

$$\Gamma_{k+1} = \Gamma_k E + \Theta_k G,$$
$$\Delta_{k+1} = \Delta_k E + \Omega_k G,$$
$$\Theta_{k+1} = \Gamma_k F,$$
$$\Omega_{k+1} = \Delta_k F,$$

while multiplication by B on the left yields

$$\Gamma_{k+1} = E\Gamma_k + F\Delta_k,$$
$$\Delta_{k+1} = G\Gamma_k,$$
$$\Theta_{k+1} = E\Theta_k + F\Omega_k,$$
$$\Omega_{k+1} = G\Theta_k.$$

From all this we deduce the differential equations

$$(2.14) \qquad \Xi' = \Xi E + \Sigma G,$$

$$(2.15) \qquad \Phi' = \Phi E + \Psi G,$$

$$(2.16) \qquad \Sigma' = \Xi F,$$

$$(2.17) \qquad \Psi' = \Phi F,$$

$$(2.18) \qquad \Xi' = E\Xi + F\Phi,$$

$$(2.19) \qquad \Phi' = G\Xi,$$

$$(2.20) \qquad \Sigma' = E\Sigma + F\Psi,$$

$$(2.21) \qquad \Psi' = G\Sigma.$$

We commence by differentiating (2.14) and substituting (2.16). The outcome is

$$(2.22) \qquad \Xi'' - \Xi'E - \Xi FG = O,$$

given with the initial conditions $\Xi(0) = I$, $\Xi'(0) = E$. Alternatively, we may differentiate (2.18), substitute (2.19), and conclude that

$$(2.23) \qquad \Xi'' - E\Xi' - FG\Xi = O.$$

Let

$$P = \begin{bmatrix} O & I \\ FG & E \end{bmatrix}.$$

Setting

$$Y := \begin{bmatrix} \Xi \\ \Xi' \end{bmatrix},$$

we see from (2.23) that the function obeys the linear system $Y' = PY$, hence

$$Y = e^{tP} \begin{bmatrix} I \\ E \end{bmatrix}.$$

Thus, integrating,

$$\begin{bmatrix} \int_0^t \Xi(\tau)d\tau \\ \Xi(t) \end{bmatrix} = \begin{bmatrix} O \\ I \end{bmatrix} + (e^{tP} - I)P^{-1} \begin{bmatrix} I \\ E \end{bmatrix}.$$

However, it is easy to verify that

$$\begin{bmatrix} O & I \\ FG & E \end{bmatrix}^{-1} \begin{bmatrix} I \\ E \end{bmatrix} = \begin{bmatrix} O \\ I \end{bmatrix},$$

and we deduce that

$$\begin{bmatrix} \int_0^t \Xi(\tau)d\tau \\ \Xi(t) \end{bmatrix} = \exp\left(t \begin{bmatrix} O & I \\ FG & E \end{bmatrix} \right) \begin{bmatrix} O \\ I \end{bmatrix}.$$

Let

$$e^{tP} = \begin{bmatrix} \tilde{\Xi}(t) & \tilde{\Sigma}(t) \\ \tilde{\Phi}(t) & \tilde{\Psi}(t) \end{bmatrix}.$$

Therefore

$$\Xi = \tilde{\Psi}, \qquad \int_0^t \Xi(\tau)d\tau = \tilde{\Sigma}.$$

It follows from (2.16) and (2.19) respectively that

$$\Sigma = \tilde{\Sigma} F, \qquad \Phi = G \tilde{\Sigma}.$$

Finally, to derive Ψ, we integrate (2.21). Hence

$$\Psi = I + G \int_0^t \tilde{\Sigma}(\tau) \mathrm{d}\tau F.$$

Note however that

$$\int_0^1 e^{\tau P} \mathrm{d}\tau \begin{bmatrix} I \\ O \end{bmatrix} = (e^{tP} - I) P^{-1} \begin{bmatrix} I \\ O \end{bmatrix} = \begin{bmatrix} \tilde{\Xi} - I & \tilde{\Sigma} \\ \tilde{\Phi} & \tilde{\Psi} - I \end{bmatrix} \begin{bmatrix} O \\ (FG)^{-1} \end{bmatrix};$$

therefore

(2.24) $$\Psi = I + G(\tilde{\Psi} - I)(FG)^{-1} F.$$

$\square$

Finally, we demonstrate directly that (2.22) and (2.23) have the same solution. In principle, there is no need to do so, since the assertion follows from our construction. Yet, a direct proof is nice, and it sheds light on the importance of the initial condition $\Xi(0) = I$, $\Xi'(0) = C$. We thus substitute

$$\Xi(t) = \sum_{k=0}^{\infty} \frac{1}{k!} \Xi_k t^k$$

in (2.22) to derive the recurrence relation

(2.25) $$\Xi_{k+2} = E\Xi_{k+1} + R\Xi_k, \qquad k \in \mathbb{Z}^+.$$

Therefore

$$\Xi_0 = I,$$
$$\Xi_1 = E,$$
$$\Xi_2 = E^2 + R,$$
$$\Xi_3 = E^3 + (ER + RE),$$
$$\Xi_4 = E^4 + (E^2 R + ERE + RE^2) + R^2,$$

and so on. We claim that, in general,

(2.26) $$\Xi_k = \sum_{2l \le k} \sum_{\substack{i_0, i_1, \dots, i_l \ge 0 \\ i_0 + \cdots + i_l = k - 2l}} E^{i_0} R E^{i_1} R \cdots R E^{i_l}, \qquad k \in \mathbb{Z}^+.$$

The proof is by induction. Suppose thus that (2.26) is correct for k and $k+1$, and substitute in (2.25). Therefore

$$\Xi_{k+2} = \sum_{2l \le k+1} \sum_{\substack{i_0, i_1, \dots, i_l \ge 0 \\ i_0 + \cdots + i_l = k+1-2l}} E^{i_0+1} R E^{i_1} R \cdots R E^{i_l}$$

$$+ \sum_{2l \le k} \sum_{\substack{i_0, i_1, \dots, i_l \ge 0 \\ i_0 + \cdots + i_l = k - 2l}} R E^{i_0} R E^{i_1} R \cdots R E^{i_l}.$$

But

$$\sum_{\substack{2l \leq k+1 \\ \substack{i_0, i_1, \ldots, i_l \geq 0 \\ i_0 + \cdots + i_l = k+1-2l}}} E^{i_0+1} R E^{i_1} R \cdots R E^{i_l}$$

$$= \sum_{\substack{2l \leq k+1 \\ \substack{i_0 \geq 1, \, i_1, \ldots, i_l \geq 0 \\ i_0 + \cdots + i_l = k+1-2l}}} E^{i_0+1} R E^{i_1} R \cdots R E^{i_l}$$

and

$$\sum_{\substack{2l \leq k \\ \substack{i_0, i_1, \ldots, i_l \geq 0 \\ i_0 + \cdots + i_l = k-2l}}} R E^{i_0} R E^{i_1} R \cdots R E^{i_l}$$

$$= \sum_{\substack{2l \leq k \\ \substack{i_0 = 0, \, i_1, \ldots, i_{l+1} \geq 0 \\ i_0 + \cdots + i_{l+1} = k-2l}}} E^{i_0} R E^{i_1} R E^{i_2} R \cdots R E^{i_{l+1}}$$

$$= \sum_{\substack{2l \leq k+2 \\ \substack{i_0 = 0, \, i_1, \ldots, i_l \geq 0 \\ i_0 + \cdots + i_l = k+2-2l}}} E^{i_0} R E^{i_1} R \cdots R E^{i_l}.$$

Adding the two expressions affirms the validity of (2.26) for $k + 2$.

Substituting the expansion in (2.23) gives the recursion

$$\Xi_{k+2} = \Xi_{k+1} E + \Xi_k R, \qquad k \in \mathbb{Z}^+,$$

and comparison with (2.25) demonstrates that the two expansions are symmetric mirror images of each other. However, (2.26) is symmetric, hence its own mirror image. This concludes the proof.

Note that directly from the previous proposition it follows that the exponential of a bordered matrix can be espressed in the following way:

$$e^{tC} = \begin{bmatrix} O & O \\ O & I \end{bmatrix} + \boldsymbol{\alpha} \varphi \boldsymbol{\beta}^{\mathrm{T}},$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $m \times (2r)$ matrices and φ is a $(2r) \times (2r)$ matrix. More precisely,

$$\boldsymbol{\alpha} = \begin{bmatrix} I & O \\ O & G \end{bmatrix}, \qquad \boldsymbol{\beta} = \begin{bmatrix} I & O \\ O & F^{\mathrm{T}} \end{bmatrix},$$

and

$$\varphi = \begin{bmatrix} \tilde{\Psi} & \tilde{\Sigma} \\ \tilde{\Sigma} & (\tilde{\Psi} - I)(FG)^{-1} \end{bmatrix}.$$

As already pointed out, this formulation lets us save computations in performing the products between the computed exponentials of the terms of the splitting.

2.3. Implementation.

2.3. Implementation. We commence from the case $B \in \mathfrak{so}[m, \mathbb{R}]$ and set

$$
B_1 = \left[\begin{array}{cc|ccc}
0 & a_{1,2} & \frac{1}{2}a_{1,3} & \cdots & \frac{1}{2}a_{1,m} \\
a_{2,1} & 0 & \frac{1}{2}a_{2,3} & \cdots & \frac{1}{2}a_{2,m} \\
\hline
\frac{1}{2}a_{3,1} & \frac{1}{2}a_{3,2} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
\frac{1}{2}a_{m,1} & \frac{1}{2}a_{m,2} & 0 & \cdots & 0
\end{array} \right],
$$

$$
B_2 = \left[\begin{array}{cc|cc|ccc}
0 & 0 & \frac{1}{2}a_{1,3} & \frac{1}{2}a_{1,4} & 0 & \cdots & 0 \\
0 & 0 & \frac{1}{2}a_{2,3} & \frac{1}{2}a_{2,4} & 0 & \cdots & 0 \\
\hline
\frac{1}{2}a_{3,1} & \frac{1}{2}a_{3,2} & 0 & a_{3,4} & \frac{1}{2}a_{3,5} & \cdots & \frac{1}{2}a_{3,m} \\
\frac{1}{2}a_{4,1} & \frac{1}{2}a_{4,2} & a_{4,3} & 0 & \frac{1}{2}a_{4,5} & \cdots & \frac{1}{2}a_{4,m} \\
\hline
0 & 0 & \frac{1}{2}a_{5,3} & \frac{1}{2}a_{5,4} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \frac{1}{2}a_{m,3} & \frac{1}{2}a_{m,4} & 0 & \cdots & 0
\end{array} \right]
$$

and so on. Each B_i is skew-symmetric and, up to a trivial permutation, bordered with $r = 2$. Moreover, if m is even then $k = m/2$, while if m is odd then we are left at the end with $(m - 1)/2$ bordered matrices with $r = 2$ and a single one with $r = 1$. Hence, altogether $k = \lfloor (m + 1)/2 \rfloor$.

Next, let us consider $B \in \mathfrak{sl}[m, \mathbb{R}]$. The only impediment to a construction as above is that taking two rows and corresponding columns of B fails to produce a matrix in $\mathfrak{sl}[m, \mathbb{R}]$.

Note, however, that in this case if $B_i \in M_{m,m}[\mathbb{R}]$ are general matrices, then

$$
\det e^{tB_1} e^{tB_2} \cdots e^{tB_k} = \prod_{i=1}^{k} \det e^{tB_i} = \prod_{i=1}^{k} \exp(t \operatorname{tr} B_i)
$$

$$
= \exp \left(\sum_{i=1}^{k} t \operatorname{tr} B_i \right) = \exp(t \operatorname{tr} B) = 1,
$$

and therefore

$$
e^{tB_1} e^{tB_2} \cdots e^{tB_k} \in SL[m, \mathbb{R}].
$$

We conclude that the optimal decomposition can be performed when $B \in \mathfrak{sl}[m, \mathbb{R}]$, disregarding the fact that $B_1, \ldots, B_k$ may be outside the Lie algebra.

Alternatively, it is possible to consider

$$
B_1 = \begin{bmatrix} \beta_1 & a_{1,2} & \frac{1}{2}a_{1,3} & \cdots & \frac{1}{2}a_{1,m} \\ a_{2,1} & -\beta_1 & \frac{1}{2}a_{2,3} & \cdots & \frac{1}{2}a_{2,m} \\ \frac{1}{2}a_{3,1} & \frac{1}{2}a_{3,2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{2}a_{m,1} & \frac{1}{2}a_{m,2} & 0 & \cdots & 0 \end{bmatrix},
$$

$$
B_2 = \begin{bmatrix} 0 & 0 & \frac{1}{2}a_{1,3} & \frac{1}{2}a_{1,4} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{2}a_{2,3} & \frac{1}{2}a_{2,4} & 0 & \cdots & 0 \\ \frac{1}{2}a_{3,1} & \frac{1}{2}a_{3,2} & \beta_2 & a_{3,4} & \frac{1}{2}a_{3,5} & \cdots & \frac{1}{2}a_{3,m} \\ \frac{1}{2}a_{4,1} & \frac{1}{2}a_{4,2} & a_{4,3} & -\beta_2 & \frac{1}{2}a_{4,5} & \cdots & \frac{1}{2}a_{4,m} \\ 0 & 0 & \frac{1}{2}a_{5,3} & \frac{1}{2}a_{5,4} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \frac{1}{2}a_{m,3} & \frac{1}{2}a_{m,4} & 0 & \cdots & 0 \end{bmatrix},
$$

and so on, where $\beta_1, \beta_2, \ldots, \beta_{k-1}$ are arbitrary real scalars, $k\lfloor(m+3)/2\rfloor$, and

$$
B_m = \mathrm{diag}\,(a_{1,1} - \beta_1, a_{2,2} + \beta_1, a_{3,3} - \beta_2, a_{4,4} + \beta_2, \dots).
$$

Note that $B_1, B_2, \ldots, B_k \in \mathfrak{sl}[m, \mathbb{R}]$.

To choose $\beta_1, \ldots, \beta_{k-1}$, the most natural strategy is to let

$$
\beta_i = \frac{1}{2}(a_{2i-1,2i-1} - a_{2i,2i}), \qquad 2i \leq m,
$$

since this projects the vector $[\ a_{2i-1,2i-1} \quad a_{2i,2i}\ ]$ onto the one-dimensional subspace spanned by $[\ 1 \quad -1\ ]$. Note that, in that case,

$$
a_{2i-1,2i-1} - \beta_i = a_{2i,2i} + \beta_i = \tfrac{1}{2}(a_{2i-1,2i-1} + a_{2i,2i}).
$$

If m is odd, we let $\beta_{(m+1)/2} = a_{m,m}$.

3. ORDER CONDITIONS

Comparison of the Taylor expansions leads to the following order conditions for the approximation of exponentials by splitting methods:

$$
(3.27) \qquad p \geq 1: \qquad \sum_{i=1}^{k} B_i = B,
$$

$$
(3.28) \qquad p \geq 2: \qquad \frac{1}{2}\sum_{i=1}^{k} B_i^2 + \sum_{i=1}^{k-1}\sum_{l=i+1}^{k} = B_i B_l = \frac{1}{2}B^2,
$$

$$
p \geq 3: \qquad \frac{1}{6}\sum_{i=1}^{k} B_i^3 + \frac{1}{2}\sum_{i=1}^{k-1}\sum_{l=i+1}^{k} (B_i^2 B_l + B_i B_l^2)
$$

$$
(3.29)
$$

$$
+ \sum_{i=1}^{k-2}\sum_{l=i+1}^{k-1}\sum_{j=l+1}^{k} B_i = B_l B_j = \frac{1}{6}B^3.
$$

Substituting (3.27) in (3.28), we obtain, after brief algebra, the order-two condition

$$(3.30) \qquad \sum_{i=1}^{k-1} \sum_{l=i+1}^{k} [B_i, B_l] = O.$$

Given $B_1, \ldots, B_k$,

$$\exp(tB_1) \cdots \exp(tB_k) = \exp(C(t, B_1, \ldots, B_k)),$$

for t sufficiently small, where $C(t, B_1, \ldots, B_k)$ can be expressed as an expansion in powers of t whose coefficients are given in terms of iterated commutators of $B_1, \ldots, B_k$ via the Baker-Campbell-Hausdorff (BCH) formula, [16] p.114, [1]. Equation (3.30) and futher order conditions for the methods are a direct consequence of the BCH formula. However, after only a few terms in the expansion, the explicit expression for the coefficients becomes very complicated and in practice useless for computational purposes.

If $B = \sum_{i=1}^{k} B_i$, the sequence

$$\tfrac{1}{2} B_1, \ldots, \tfrac{1}{2} B_{k-1}, B_k, \tfrac{1}{2} B_{k-1}, \ldots, \tfrac{1}{2} B_1$$

satisfies order-two conditions and is known as Strang splitting [14].

3.1. Optimality of the Strang splitting.

Let us examine a splitting

$$B = \sum_{i=1}^{k} B_i$$

of order 1, assuming that the matrices B_i do not commute. We will consider splittings

$$(3.31) \qquad B = \sum_{l=1}^{r} \lambda_l B_{\omega(l)}$$

such that $\lambda_l \neq 0$ for $l = 1, \ldots, r$, and the map $\omega : I_r \to I_k$, where $I_k := \{1, \ldots, k\}$ and $I_r := \{1, \ldots, r\}$, is surjective and defines the splitting (3.31). For the Strang splitting we have $r = 2k - 1$, $\lambda_l = \frac{1}{2}$, $l \neq k$ and $\lambda_r = 1$. We wish to determine the smallest value of r such that the splitting (3.31) is of order 2.

Proposition 7. *If the splitting* (3.31) *is of order 2, then* $r \geq 2k - 1$.

Proof. Recall that a second-order splitting requires

$$(3.32) \qquad \sum_{l=1}^{r-1} \sum_{j=l+1}^{r} \lambda_l \lambda_j [B_{\omega(l)}, B_{\omega(j)}] = O.$$

Suppose that there exists a splitting (3.31) of order 2 with $r \leq 2k - 2$. Then for at least two indices $j, l \in I_r$ with $j \neq l$ it is true that if

$$\omega(j) = j^*, \quad \forall i \in I_r \setminus \{j\} \quad \omega(i) \neq j^*,$$
$$\omega(l) = l^*, \quad \forall i \in I_r \setminus \{l\} \quad \omega(i) \neq l^*,$$

then $j^* \neq l^*$. We will suppose without loss of generality that $j < l$. Then the term

$$\lambda_l \lambda_j [B_{\omega(l)}, B_{\omega(j)}] = \lambda_l \lambda_j [B_{l^*}, B_{j^*}] \neq 0$$

appears only once in the left hand side of (3.32), and the sum cannot be equal to zero. $\qquad \square$

Since for the Strang splitting we have $r = 2k - 1$ and

$$B_{\omega(l)} = B_{\omega(2k-l)} = \tfrac{1}{2} B_{k+1-l}, \quad l = 1, 2, \ldots, k-1, \qquad B_{\omega(k)} = B_1,$$

it fulfils the requirements of optimality stated in the previous proposition.

4. HIGHER ORDER SCHEMES AND OTHER COMPUTATIONAL ASPECTS

Suppose that

$$(4.1) \qquad\qquad F(t) = e^{tA} + E t^{p+1} + \mathcal{O}\big(t^{p+2}\big),$$

where $E \in \mathrm{M}_m[\mathbb{R}] \setminus \{O\}$. It is trivial to verify that the approximant

$$(4.2) \qquad\qquad G(t, \boldsymbol{\gamma}) := F(\gamma_0 tA) F(\gamma_1 tA) \cdots F(\gamma_r tA)$$

is of order $p + 1$ if and only if

$$(4.3) \qquad\qquad \sum_{l=0}^{r} \gamma_l = 1, \qquad \sum_{l=0}^{r} \gamma_l^{p+1} = 0.$$

Note that (4.3) has no solution if p is odd. On the other hand, it is easy to verify that $p = 2$, $r = 2$ has a real solution and that this is the least value of r that is adequate for $p = 2$. As a matter of fact, using Lagrange multipliers, it is possible to show that the solution of (4.3) which is least in the ℓ_2 norm is

$$\gamma_0 = \gamma_2 = \frac{2}{3} + \frac{\sqrt[3]{2}}{3} + \frac{\sqrt[3]{4}}{6}, \qquad \gamma_1 = -\frac{(1 + \sqrt[3]{2})^2}{3}.$$

This technique has been used by Yoshida [17] in the case of symplectic integrators for Hamiltonian systems. Here, as in Yoshida analysis, the following definition plays a central role.

Definition 2. We say that a mapping $\varphi_t : \mathcal{V} \to \mathcal{V}$, defined for all sufficiently small $|t|$ and an arbitrary Euclidean space $\mathcal{V}$, is *time-symmetric* if

$$\varphi_{-t} \circ \varphi_t = \mathrm{Id}.$$

Lemma 8. *The Strang splitting is time-symmetric.*

Proof. Let

$$F(t) = e^{\frac{1}{2} t B_k} e^{\frac{1}{2} t B_{k-1}} \cdots e^{\frac{1}{2} t B_2} e^{t B_1} e^{\frac{1}{2} t B_2} \cdots e^{\frac{1}{2} t B_{k-1}} e^{\frac{1}{2} t B_k}$$

be the Strang splitting. It is trivial to verify that $F(-t)F(t) = I$. Hence the splitting is a time-symmetric map. □

The above result is important by virtue of the following theorem.

Theorem 9. *If φ_t is time-symmetric and analytic (as a function of t), then*

$$\varphi_t X = e^{tY_0 + t^3 Y_1 + t^5 Y_2 + \cdots}, \qquad X \in \mathcal{V},$$

an expansion in odd powers of t.

In the special case $m = 2$ we can use the extended BCH formula (equation (12.12) in [11], p. 161), namely

$$e^{tX} e^{tY} e^{tX} = e^{W(t)},$$

where

$$W(t) = t(2X + Y) + \tfrac{1}{6}t^3[Y, [Y, X]] - \tfrac{1}{6}t^3[X, [X, Y]] + \tfrac{7}{360}t^5[X, [X, [X, [X, Y]]]]$$
$$- \tfrac{1}{360}t^5[Y, [Y, [Y, [Y, X]]]] + \tfrac{1}{90}t^5[X, [Y, [Y, [Y, X]]]] + \tfrac{1}{45}t^5[Y, [X, [X, [X, Y]]]]$$
$$- \tfrac{1}{60}t^5[X, [X, [Y, [Y, X]]]] + \tfrac{1}{30}t^5[Y, [Y, [X, [X, Y]]]] + \mathcal{O}(t^7)$$

to prove that, for the Strang splitting, $B = B_1 + B_2$, $X = \tfrac{1}{2}B_2$, $Y = B_1$,

$$W(t) = t(B_1 + B_2) + \tfrac{1}{24}t^3[2B_1 + B_2, [B_1, B_2]] + \tfrac{1}{5760}t^5\{-7[B_2, [B_2, [B_2, [B_1, B_2]]]]$$
$$- 8[B_1, [B_1, [B_1, [B_1, B_2]]]] + 16[B_2, [B_1, [B_1, [B_1, B_2]]]]$$
$$- 16[B_1, [B_2, [B_2, [B_1, B_2]]]] - 12[B_2, [B_2, [B_1, [B_1, B_2]]]]$$
$$- 48[B_1, [B_1, [B_2, [B_1, B_2]]]]\} + \mathcal{O}(t^7).$$

Let us suppose that

$$F(t) = \exp\left\{tB + \sum_{k=q}^{\infty} t^{2k+1}E_k\right\}$$

and

$$H(t) = F(\gamma_0 t)F(\gamma_1 t)F(\gamma_0 t),$$

where

$$2\gamma_0 + \gamma_1 = 1, \qquad 2\gamma_0^{2q+1} + \gamma_1^{2q+1} = 0.$$

Then, using again the extended BCH formula, we see that $H(t) = e^{V(t)}$, where

$$V = tB + t^{2q+3}\{(2\gamma_0^{2q+3} + \gamma_1^{2q+3})E_{q+1} + \tfrac{1}{6}\gamma_0\gamma_1(\gamma_0 + \gamma_1)(\gamma_0^{2q} - \gamma_1^{2q})[B, [B, E_q]]\}$$
$$+ \mathcal{O}(t^{2q+4}).$$

Therefore the outcome, $H(t)$, represents a method of order $2q + 2$. Moreover, it is trivial to observe that the time-symmetry of F implies that H is time-symmetric, since

$$H(-t) = F(-\gamma_0 t)F(-\gamma_1 t)F(-\gamma_0 t) = F^{-1}(\gamma_0 t)F^{-1}(\gamma_1 t)F^{-1}(\gamma_0 t)$$
$$= [F(\gamma_0 t)F(\gamma_1 t)F(\gamma_0 t)]^{-1} = H^{-1}(t).$$

Therefore this procedure can be iterated, increasing the order by two in each nested application and with perfect analogy to what Yoshida observed in the case of symplectic methods.

Note that Yoshida had two different techniques for increasing the order. Letting F stand again for the Strang splitting, the other method reads

$$(4.4) \qquad F(\gamma_r t)F(\gamma_{r-1} t) \cdots F(\gamma_1 t)F(\gamma_0 t)F(\gamma_1 t) \cdots F(\gamma_{r-1} t)F(\gamma_r t).$$

The coefficients $\gamma_0, \ldots, \gamma_r$ are chosen to maximize order. We already know that $r = 1$ leads to order 4. Likewise, $r = 3$ results in order 6 and $r = 7$ yields order 8. It is tempting to conjecture that $r = 2^{l-1} - 1$ gives order $2l$.

At any rate, our claim is that we need precisely $4(k - 1)r + 2k - 1$ exponentials to evaluate a Yoshida enhancement of type II for a Strang splitting into k matrices: Each Strang splitting requires $2k - 1$ exponentials, and we have $2r + 1$ Strang splittings. However, 'endpoint' matrices can be amalgamated,

$$e^{\gamma_i t B_1}e^{\gamma_{i+1} t B_1} = e^{(\gamma_i + \gamma_{i+1})B_1},$$

and this allows us to arrive at the lower count. This, as a matter of fact, is a phenomenon which is exclusive to our application, *precisely because our exponentials are exact*. It has no counterpart in the original Yoshida–Suzuki–Forest [17], [15], [4] application to separable Hamiltonian flows.

A moderate saving in the number of exponentials takes place also in the context of Yoshida type I: For order 4 we have $3(2k - 1) - 2 = 6k - 5$ exponentials – in this case type I and type II coincide. For order 6 we obtain $3(6k - 5) - 2 = 18k - 17$ exponentials, and for order 8 the count is $3(18k - 17) - 2 = 54k - 53$. In comparison, the number of exponentials for type II is $14k - 13$ ($r = 3$, order 6) and $30k - 29$ ($r = 7$, order 8). Type II is invariably better.

5. Numerical experiments

We present numerical experiments for computing $\exp(tB)$ with splitting methods in the cases $B \in \mathfrak{so}(m)$ and $B \in \mathfrak{sl}(m)$.

The methods are implemented both with and without the Yoshida technique for improving the order, and we restrict ourselves to the case in which the first and second Yoshida techniques coincide, i.e. $r = 1$ in (4.2) and (4.4) respectively.

The computations are made in `Matlab`. We evaluate the the error computing $\|\mathrm{e}^{-tB} F(tB) - I\|_{\mathrm{F}}$, where $F(tB)$ is the approximation with the splitting methods, the exact exponential e^{-tB} is computed by the `Matlab` built-in function *expm* and $\| \cdot \|_{\mathrm{F}}$ denotes the Frobenius norm.

The matrices are generated randomly using the `Matlab` function *randn* and scaling the Frobenius norm in such a way that $\|B\|_{\mathrm{F}} = 10$.

Note that in our experiments the absolute error in the Frobenius norm and $\|\mathrm{e}^{-tB} F(tB) - I\|_{\mathrm{F}}$ are almost the same, and the ratio of the latter to the relative error is a constant, so the three standard ways of measuring the error are equivalent.

We decompose $B = \sum_{i=1}^{k} B_i$ using the columnwise decomposition and the optimal decomposition in the case of $B \in \mathfrak{so}(m)$. The B_i's are obtained considering the sum of an even number of rank-one matrices; otherwise they cannot belong to $\mathfrak{so}(m)$. The exponentials of the B_i's are computed by formula (2.10).

In all figures the norm of the error is plotted (along the y-axis) to a logarithmic scale with respect to the time t. Figure 1 reports the results of the first test: we have considered a sparse matrix in $\mathfrak{so}(100)$, the B_i's are rank-2 matrices. In the plots the error norm is indicated with the symbols '$*$' and '$+$' for columnwise decomposition and with '$\circ$' and '$\times$' for the optimal decomposition. Note that we evaluate 50 rank-2 exponentials $\exp(tB_i)$ in the first case, and 99 in the second case. However, performing the optimal decomposition is more expensive than the columnwise one that is obtained directly in terms of the columns of the matrix B. The symbols '$+$' and '$\times$' pertain to the Strang splitting improved by the Yoshida technique, and '$*$', '$\circ$' refer to plain Strang splitting.

Figure 2 refers to the case of a full matrix in $\mathfrak{so}(100)$. The meaning of the symbols is the same as in the previous figure; the rank of the B_i's is 2. It turns out that the leading term of the error in the case of the optimal decomposition is larger in norm than that of the columnwise decomposition. Note that the B_i's of the columnwise decomposition are sparse matrices and each commutator of two different B_i's inherits the sparsity of the lesser sparse of the two; the B_i's of the optimal decomposition are instead full matrices. Since the leading term of the error is a matrix obtained by computing commutators between the B_i's, this seems to
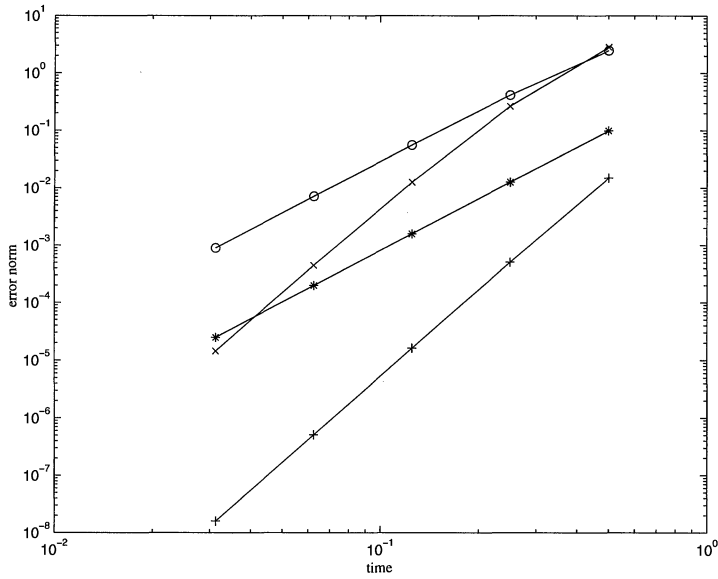
FIGURE 1. Exponential of a sparse matrix of $\mathfrak{so}(100)$.
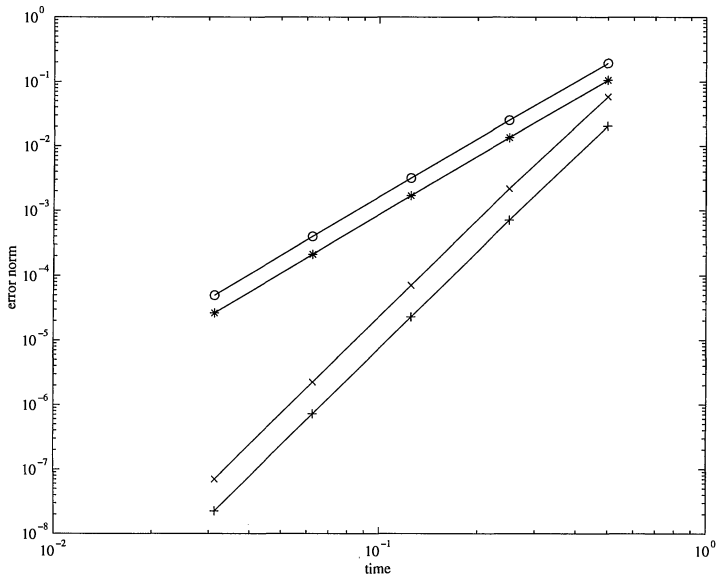

FIGURE 2. Exponential of a full matrix of $\mathfrak{so}(100)$.

indicate why the sparsity of the B_i's in the columnwise decomposition leads to a lower norm. Note that the difference between the two decompositions is not so evident in the previous test, in which, since B is sparse, the B_i's are sparse matrices for both approaches.
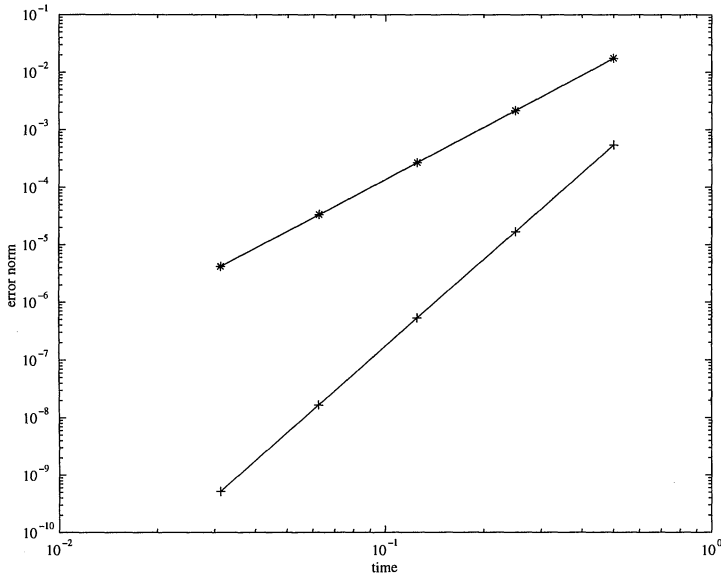
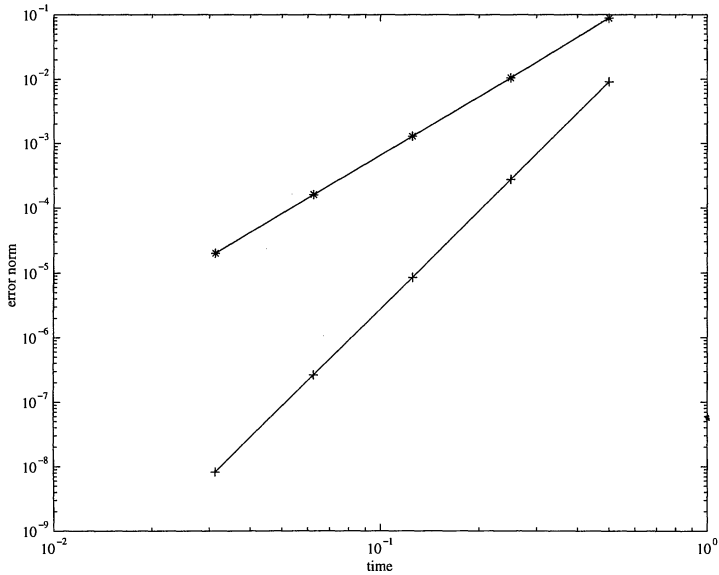FIGURE 3. Exponential with compact trajectory of a matrix of $\mathfrak{sl}(100)$.

FIGURE 4. Exponential of a matrix of $\mathfrak{sl}(100)$ with an arbitrary spectrum.

When $B \in \mathfrak{sl}(m)$ we take every B_i as a bordered matrix and we compute the exponential directly, using the result of Proposition 6, with $r = 2$ in our tests. Note that in our examples the matrix FG is always nonsingular. Where this is not the case, it would have been possible to perform one of the alternative decompositions (columnwise or optimal) described in subsection 2.2.

TABLE 1. $p = 2$

	order 1	order 2	order 4
Col. dec.	2.04	4.09	12.27
Opt. dec.	2.51	5.04	15.09
Basis dec.	2.10	4.11	12.25

TABLE 2. $p = 4$

	order 1	order 2	order 4
Col. dec.	2.38	4.72	14.04
Opt. dec.	3.04	6.02	17.94

In the first example of $\mathfrak{sl}$, Figure 3, we generate a full matrix whose exponential has a compact trajectory $\{\exp(tB)|t > 0\}$. Note that all such matrices have pure imaginary eigenvalues. We consider $B = V^{-1}SV$, where $S \in \mathfrak{so}(m)$ and V is a random $m \times m$ nonsingular matrix. Note that $V^{-1}\mathfrak{so}V$ is isomorphic to $\mathfrak{so}$ and that in this case diagonal Padé approximants map $V^{-1}\mathfrak{so}V$ to the subgroup $V^{-1}SOV$ of SL. We report the norm of the error for the Strang splitting ('$*$') and for the same technique with the Yoshida improvement of the order ('$+$').

The last figure, Figure 4, refers to a matrix in $\mathfrak{sl}$ with arbitrary eigenvalues. In this case using the diagonal Padé approximant gives an approximation that does not belong to SL, emphasizing the need to employ the techniques of this paper.

Finally we present Tables 1 and 2 displaying the number of flops employed by the methods of order 1, 2 and 4 for a full matrix $B \in \mathfrak{so}(100)$. The flops have been counted with the built-in Matlab function *flops*, and the results reported have been scaled by a factor m^3—in other words, for each value C reported in the tables, the real costs of the methods can be obtained by taking Cm^3.

Note that, in Table 1, together with *columwise decomposition* and *optimal decomposition* we report the results for a decomposition obtained in terms of the elements of the following basis of $\mathfrak{so}(m)$:

$$(5.5) \qquad e_i e_j^{\mathrm{T}} - e_j e_i^{\mathrm{T}}, \quad i = 1, \ldots, m, \ \ j = i + 1, \ldots, m.$$

The coordinates of $B \in \mathfrak{so}(m)$ with respect to this basis are the coefficients of B corresponding to the same couples of indices. The number of exponentials to be multiplied together in this case is $m(m-1)/2$. Since each exponential can be evaluated directly using an explicit formula and their product can be implemented very efficiently, computational costs in this case are analogous to the previous two. Based on this observation, we are at present investigating a new strategy of composition methods for the approximation of the exponential matrix that will be described in a forthcoming paper.

Table 1 refers to the case $p = 2$, and Table 2 to $p = 4$. It seems that considering rank-4 decompositions instead of rank-2 is not any cheaper. The number obtained can be compared with the analogous result obtained with the built-in function *expm* of Matlab, namely $20.98m^3$.

Based on the construction of low-rank splittings in a Lie algebra, in this paper we have proposed a new approach for approximating the matrix exponential. The importance of such an approximation of exp is that it is a local map from a Lie algebra to a Lie group. This feature turns out to be essential when Lie-group type integrators are concerned, and in some cases such techniques offer a valid alternative

to the exact calculation of the exponential. Finally, if we disregard the geometric advantages offered from the method and we consider simply the bounds obtained for the computational costs, and the counts of flops in the numerical experiments, we can conclude that the proposed technique is competitive with other classical approaches, i.e. Padé approximants, and, in the case of $\exp(tB)v$ and v, an m-dimensional vector with Krylov subspace techniques.

6. Acknowledgments

References

[1] R.V. Chacon and A.T. Fomenko. Recursion formulas for the Lie integral. *Advances in Math.*, 88:200–257, 1991. MR **93e**:22013

[2] P. E. Crouch and R. Grossman. Numerical Integration of Ordinary Differential Equations on Manifolds. *J. Nonlinear Sci.*, 3:1–33, 1993. MR **94e**:65069

[3] K. Feng and Z.-J. Shang. Volume-preserving algorithms for source-free dynamical systems. *Numer. Math.*, 71:451–463, 1995. MR **96g**:65065

[4] E. Forest. Sixth-order Lie group integrators. *J. of Comp. Phys.*, 99:209–213, 1992. MR **93b**:70003

[5] E. Gallopoulos and Y. Saad. Efficient solution of parabolic equations by Krylov approximation methods. *SIAM J. Sci. Statist. Comput.*, 13:1236–1264, 1992. MR **93d**:65085

[6] M. Hochbruck and Ch. Lubich. On Krylov Subspace Approximations to the Matrix Exponential Operator. *SIAM J. Numer. Anal.*, 34:1911–1925, 1997. MR **98h**:65018

[7] A. Iserles and S. P. Nørsett. On the Solution of Linear Differential Equations in Lie Groups. Roy. Soc. London Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357** (1999), 983–1019. CMP 99:13

[8] R. I. McLachlan. On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM J. Sci. Comput.*, 16:151–168, 1995. MR **95j**:65073

[9] H. Munthe-Kaas. Runge-Kutta Methods on Lie Groups. *BIT*, 38:92–111, 1998. MR **99f**:65117

[10] H. Munthe-Kaas and A. Zanna. Numerical integration of differential equations on homogeneous manifolds. In F. Cucker and M. Shub, editors, *Foundations of Computational Mathematics*, pages 305–315. Springer Verlag, 1997. MR **99i**:65082

[11] J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian Problems*. AMMC 7. Chapman & Hall, London, 1994. MR **95f**:65006

[12] M. Schatzman. Higher order alternate directions methods. Technical report, Université de Lyon, Laboratoire d'analyse numérique, 1997.

[13] Q. Sheng. *Solving partial differential equations by exponential splitting*. PhD thesis, Cambridge University, 1989.

[14] G. Strang. Accurate partial difference methods. II. Nonlinear problems. *Num. Math.*, 6:37–49, 1964. MR **29**:4215

[15] M. Suzuki. General theory of fractal path integrals with applications to many-body theories and statistical physics. *J. of Math. Phys.*, 32:400–407, 1991. MR **92k**:81096

[16] V. S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representation*. GTM 102. Springer-Verlag, 1984. MR **85e**:22001

[17] H. Yoshida. Construction of higher order symplectic integrators. *Physics Letters A*, 150:262–268, 1990. MR **91h**:70014

[18] A. Zanna. The method of iterated commutators for ordinary differential equations on Lie groups. Technical Report 1996/NA12, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, 1996.

DAMTP, Cambridge University, Silver Street, England CB3 9EW
Current address: Department of Mathematical Sciences, NTNU 7491 Trondheim, Norway
E-mail address: `elenac@math.ntnu.no`

DAMTP, Cambridge University, Silver Street, England CB3 9EW
E-mail address: `A.Iserles@damtp.cam.ac.uk`