SIAM J. Numer. Anal.
Vol. 54, No. 2, pp. 1020–1038

DISCRETIZED DYNAMICAL LOW-RANK APPROXIMATION IN THE PRESENCE OF SMALL SINGULAR VALUES[*]

EMIL KIERI[†], CHRISTIAN LUBICH[‡], AND HANNA WALACH[‡]

Abstract. Low-rank approximations to large time-dependent matrices and tensors are the subject of this paper. These matrices and tensors either are given explicitly or are the unknown solutions of matrix and tensor differential equations. Based on splitting the orthogonal projection onto the tangent space of the low-rank manifold, novel time integrators for obtaining approximations by low-rank matrices and low-rank tensor trains were recently proposed. By standard theory, the Lie–Trotter and Strang projector-splitting methods are first and second order accurate, respectively, but the usual error bounds break down when the low-rank approximation has small singular values. This happens when the singular values of the solution decay without a distinct gap or when the effective rank of the solution is overestimated. On the other hand, the integrators are exact when given time-dependent matrices or tensors are already of the prescribed rank. We provide an error analysis which unifies these properties. We show that in cases where the exact solution is an ε-perturbation of a low-rank matrix or tensor train, the error of the projector-splitting integrator is favorably bounded in terms of ε and the stepsize, independently of the smallness of the singular values. Such a result does not hold for any standard integrator. Numerical experiments illustrate the theory.

Key words. tensor train, low-rank approximation, tensor differential equations, splitting integrator

AMS subject classifications. 15A18, 65L05, 65L70

DOI. 10.1137/15M1026791

1. Introduction.

Low-rank approximations to matrices and tensors are a basic tool in data and model reduction; see, e.g., [4, 5]. In this paper we are concerned with the low-rank approximation of *time-dependent* matrices and tensors, which either are given explicitly by their increments or are the unknown solutions of differential equations. In [11] such time-dependent problems and their numerical treatment were first studied for matrices. Differential equations for the factors of a low-rank factorization similar to the singular value decomposition were derived and their approximation properties were studied. Extensions to time-dependent tensors in various tensor formats were given in [2, 12, 18, 19]; see also [15] for a review of dynamical low-rank approximation.

The approach yields differential equations on low-rank matrix and tensor manifolds, which need to be solved numerically. Recently, very efficient integrators based on splitting the projection onto the tangent space of the low-rank manifold have been proposed and studied for matrices and for tensors in the tensor-train format in [16] and [17], respectively. The objective of the present paper is to show that *these projector-splitting integrators are insensitive to the presence of small singular values*

[*] Received by the editors June 18, 2015; accepted for publication (in revised form) March 1, 2016; published electronically April 5, 2016.

http://www.siam.org/journals/sinum/54-2/M102679.html

[†] Division of Scientific Computing, Department of Information Technology, Uppsala University, 751 05 Uppsala, Sweden (emil.kieri@it.uu.se). The research reported in this paper was carried out while this author was visiting Universität Tübingen with support from Anna Maria Lundins stipendiefond vid Smålands nation i Uppsala.

[‡] Mathematisches Institut, Universität Tübingen, D-72076 Tübingen, Germany (lubich@na.uni-tuebingen.de, walach@na.uni-tuebingen.de). The research of the third author was supported by a grant from DFG through the GRK 1838.

in the low-rank approximation, a property that is not shared by any standard integrator such as explicit or implicit Runge–Kutta methods, whose behavior deteriorates when singular values become small.

The presence of small singular values in the low-rank approximation of a large matrix is very common. Unless the matrix has a distinct gap in the distribution of its singular values, truncating all the smallest singular values below a tolerance ε yields a remaining matrix of reduced rank that still has singular values of magnitude $O(\varepsilon)$. Even if there is a distinct gap in the singular value distribution such that two groups of large and negligibly small singular values, respectively, are formed, it is typically not known a priori at which rank the former group ends. We are also in a time-dependent setting where the distribution of singular values may change over time. Underestimating the effective rank means we neglect a significant part of the matrix, which leads to poor accuracy, but overestimating the effective rank yields an approximation with small singular values. A similar situation arises in the approximation of tensors.

In the matrix case, we are concerned with time-dependent matrices $A(t) \in \mathbb{C}^{m\times n}$, $t_0 \le t \le T$, for large m and n. These matrices either are known explicitly or are the unknown solution of a matrix differential equation

$$\dot{A}(t) = F(t, A(t)), \qquad A(t_0) = A_0. \tag{1.1}$$

We seek an approximate solution $Y(t)$ to (1.1) on the manifold $\mathcal{M}_r$ of complex rank-r $m \times n$-matrices. To construct an evolution equation for $Y(t) \in \mathcal{M}_r$ we project the right-hand side of the differential equation onto the tangent space $\mathcal{T}_{Y(t)}\mathcal{M}_r$ of $\mathcal{M}_r$ at the current approximation $Y(t)$. This can be interpreted as a Galerkin method on the solution-dependent tangent space. In the context of quantum physics, such a procedure is known as the Dirac–Frenkel time-dependent variational principle [13, 14] and can be traced back to [3] for a special application. Denoting the orthogonal projection onto $\mathcal{T}_{Y(t)}\mathcal{M}_r$ by $P(Y(t))$, we get the differential equation for $Y(t)$ on the manifold $\mathcal{M}_r$,

$$\dot{Y}(t) = P(Y(t))F(t, Y(t)), \qquad Y(t_0) = Y_0 \in \mathcal{M}_r. \tag{1.2}$$

In the case where $A(t)$ is instead given explicitly, the dynamical low-rank approximation $Y(t)$ is still determined by a differential equation, where the right-hand side is of the same form with $F(t, Y) = \dot{A}(t)$ independent of Y. A first difficulty with small singular values is already seen at this abstract level: the local Lipschitz constant of the tangent space projection P at Y, or in other words the curvature of the manifold $\mathcal{M}_r$ at Y, is proportional to the inverse of the smallest nonzero singular value of Y; see [11, Lemma 4.2].

The rank-r matrix $Y(t)$ is not computed via its $m \times n$ entries but is considered in factorized form as

$$Y(t) = U(t)S(t)V(t)^*$$

with $U(t) \in \mathbb{C}^{m\times r}$ and $V(t) \in \mathbb{C}^{n\times r}$ having orthonormal columns and with invertible $S(t) \in \mathbb{C}^{r\times r}$. This nonunique decomposition is similar to the singular value decomposition, except that $S(t)$ is not assumed diagonal. When $r \ll m, n$, this representation offers a significant reduction in memory requirements, compared to storing the full matrix $Y(t)$. The factors are determined from (1.2) as the solution of the following system of differential equations [11]:

$$
\begin{aligned}
\dot{U}(t) &= (I - U(t)U(t)^*)F(t, Y(t))V(t)S(t)^{-1}, \\
\dot{V}(t) &= (I - V(t)V(t)^*)F(t, Y(t))^*U(t)S(t)^{*-1}, \\
\dot{S}(t) &= U(t)^*F(t, Y(t))V(t).
\end{aligned} \tag{1.3}
$$

The nonzero singular values of $Y(t)$ are those of the $r \times r$ matrix $S(t)$, whose inverse appears in the first two differential equations. The presence of small singular values therefore leads to severe problems when these differential equations are integrated numerically by standard methods such as explicit or implicit Runge–Kutta methods.

In contrast, the projector-splitting integrator of [16] does not deteriorate in the case of an ill-conditioning of $S(t)$, although it also computes the factors U, S, V. In [16] this was observed numerically and shown analytically in the special case of a distinct gap in the distribution of the singular values of given time-dependent matrices $A(t)$. Moreover, if the given matrices $A(t)$ are all of rank at most r, then the splitting integrator was shown to reproduce $A(t)$ exactly. This remarkable exactness property will be an important tool in the present paper.

In section 2 we give an approximation result for the projector-splitting integrator applied with stepsize h to the differential equation (1.2) with Lipschitz-continuous functions $F(t, Y)$ that map onto the tangent space of $\mathcal{M}_r$ at Y up to a small remainder of size ε. We prove an $O(\varepsilon + h)$ error bound when the initial value is chosen of rank r. The constants in this error bound are *independent of the singular values* of $A(t)$ or $Y(t)$ or $S(t)$. Such a robust error bound cannot be obtained for standard integrators applied to (1.3).

A limitation of our theoretical result is that it requires a (local) Lipschitz condition on F and is applicable to stiff differential equations such as discretized partial differential equations only under a severe CFL condition $hL \ll 1$, where h is the stepsize and L is the Lipschitz constant. Such a restriction is not observed to be necessary in numerical experiments, and it would thus be of interest to improve the results of this paper beyond such a limitation in future work.

As shown in section 3, the robust error bound extends to the projector-splitting integrator of [17] for approximations of time-dependent tensors in the tensor-train format. This is a data-sparse tensor format introduced in [20] in the mathematical literature. It has previously been used in physics under the name of matrix product states; see, e.g., [22, 25] and, in a time-dependent context, [6, 7].

Numerical experiments presented in section 4 conclude the paper. They corroborate our theoretical results and further illustrate the potential of the projector-splitting method beyond the numerical experiments presented in [16, 17].

2. The projector-splitting integrator for matrix differential equations.

In this section, after briefly presenting the projector-splitting integrator of [16], we state and prove local and global error bounds for the integrator that do not deteriorate in the presence of small singular values in the exact solution or its low-rank approximation. But first, we use a small numerical example to illustrate how small singular values pose a problem to a standard integrator, while the projector-splitting integrator performs well also in this situation.

2.1. A motivating numerical example.

We present the results of numerical experiments with two explicit numerical methods for (1.2): Using the classical fourth order Runge–Kutta method to solve (1.3) and the (first order) Lie–Trotter projector-splitting integrator of [16]. In this example $A(t) \in \mathbb{R}^{100\times 100}$ is given explicitly by constructing two 100×100 skew-symmetric matrices W_1, W_2 and a diagonal matrix D of the same dimension with exponentially decreasing diagonal elements $d_j = 2^{-j}$, $j = 1, \ldots, 100$. By means of those matrices, we generate

$$A(t) = e^{tW_1} e^t D \big(e^{tW_2}\big)^T$$

on the time interval $0 \le t \le 1$. The singular values of $A(t)$ are $\sigma_j(t) = e^t d_j$, $j = 1, \ldots, 100$.

Figure 1 shows the approximation errors, that is, the norm of the difference between the given matrix $A(t)$ and the numerical solution $Y_n = U_n S_n V_n^T$ of rank r obtained with n steps of stepsize h for $t = nh$, at $t = 1$. The errors versus the stepsize are shown for both methods and for different ranks.

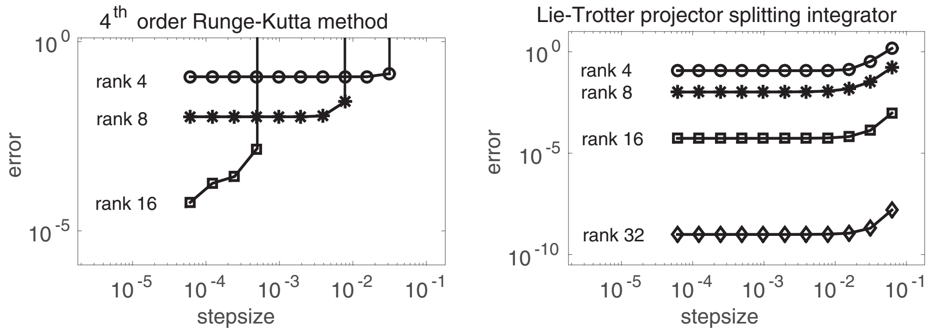

FIG. 1. *Comparing the Runge–Kutta method (left) and the Lie–Trotter integrator (right) for different approximation ranks and stepsizes.*

The Runge–Kutta method with approximation rank r turns out to be stable only for small stepsizes and small approximation ranks. For larger approximation ranks r the Runge–Kutta method demands very small step sizes, proportional to σ_r. This restriction is due to the small singular values of $S(t)$ in (1.3).

In contrast, we are able to choose large time steps for the Lie–Trotter projector-splitting integrator independently of the chosen rank. The error decays linearly with h as $h \to 0$ and decreases with increasing rank, which indicates that the method is accurate and robust with respect to small singular values.

It is the objective of this paper to explain this favorable error behavior.

2.2. The practical integration algorithm. A step of the integrator for (1.3) given in [16] proceeds as follows, starting from the factorized rank-r matrix $Y_0 = U_0 S_0 V_0^*$ and computing the factors of the approximation $Y_1 = U_1 S_1 V_1^*$ at time $t_1 = t_0 + h$:

1. Solve the differential equation on $\mathbb{C}^{m \times r}$

$$\dot{K}(t) = F(t, K(t)V_0^*)V_0, \qquad K(t_0) = U_0 S_0,$$

and orthonormalize the columns of $K(t_1)$ (by QR factorization),

$$U_1 \widehat{S}_1 = K(t_1),$$

where $U_1 \in \mathbb{C}^{m \times r}$ has orthonormal columns and $\widehat{S}_1 \in \mathbb{C}^{r \times r}$.

2. Solve the differential equation on $\mathbb{C}^{r \times r}$

$$\dot{S}(t) = -U_1^* F(t, U_1 S(t) V_0^*)V_0, \qquad S(t_0) = \widehat{S}_1,$$

and set $\widetilde{S}_0 = S(t_1)$.

3. Solve the differential equation on $\mathbb{C}^{n\times r}$

$$\dot{L}(t) = F(t, U_1 L(t)^*)^* U_1, \qquad L(t_0) = V_0 \widetilde{S}_0^*,$$

and orthonormalize the columns of $L(t_1)$ (by QR factorization),

$$V_1 S_1^* = L(t_1),$$

where $V_1 \in \mathbb{C}^{n\times r}$ has orthonormal columns and $S_1 \in \mathbb{C}^{r\times r}$.

The algorithm computes a factorization of the rank-r matrix

$$Y_1 = U_1 S_1 V_1^*,$$

which is taken as an approximation to $Y(t_1)$.

We note that the differential equations in the substeps can be solved exactly when $F(t,Y) = \dot{A}(t)$ for given matrices $A(t)$. In the first substep we then have $K(t_1) = K(t_0) + \Delta A V_0$ with the increment $\Delta A = A(t_1) - A(t_0)$, and similar formulas are obtained also for the second and third substeps. The algorithm just uses the increment ΔA but does not require the time derivative $\dot{A}(t)$.

In the general case, the differential equations need to be solved approximately, e.g., using a step of a Runge–Kutta method or, when F is independent of t and linear in Y, by Krylov methods for computing the action of a matrix exponential [9, 21].

2.3. The integrator as a projector-splitting scheme. The above algorithm and its time-symmetrized variant can be interpreted as splitting integrators based on splitting the tangent space projection $P(Y)$ in (1.2). With $Y = USV^*$, the projection $P(Y)$ can be decomposed as (cf. [11])

$$P(Y) = P_1^+(Y) - P_1^-(Y) + P_2^+(Y) \tag{2.1}$$

with

$$P_1^+(Y)Z = ZVV^*, \quad P_1^-(Y)Z = UU^*ZVV^*, \quad P_2^+(Y)Z = UU^*Z.$$

UU^* and VV^* are the orthogonal projections to the ranges of Y and Y^*, respectively, and do not depend on how we decompose $Y = USV^*$. The notation $P_i^{\pm}$, which might seem odd at this point, is chosen to be consistent with the projections for the tensor case.

We introduce the notation

$$F_i^{\pm}(t,Y) = \pm P_i^{\pm}(Y) F(t,Y)$$

for the right-hand sides of the subproblems. The first order Lie–Trotter projector-splitting scheme is the consecutive solution of the partial problems $\dot{Y} = F_1^+(t,Y)$, $\dot{Y} = F_1^-(t,Y)$, and $\dot{Y} = F_2^+(t,Y)$. We denote the solution operator of $\dot{A} = F(t,A)$, $A(t_0) = A_0$ by

$$A(t) = \Phi_F(t, t_0, A_0),$$

and similarly

$$Y_i^{\pm}(t) = \Phi_{F_i^{\pm}}(t, t_0, Y_i^{\pm}(t_0))$$

for $F_i^{\pm}$ instead of F. We can then write one step of the splitting scheme from t_0 to $t_1 = t_0 + h$ as

$$Y_1 = \mathcal{S}(t_1, t_0, Y_0) = \Phi_{F_2^+}(t_1, t_0, \Phi_{F_1^-}(t_1, t_0, \Phi_{F_1^+}(t_1, t_0, Y_0))).$$

This matrix Y_1 is the same as that obtained by the algorithm of the previous subsection; see [16]. By Y_n we denote the solution after n time steps with the splitting method. The adjoint scheme is

$$\mathcal{S}^*(t_1, t_0, Y_0) = \Phi_{F_1^+}(t_1, t_0, \Phi_{F_1^-}(t_1, t_0, \Phi_{F_2^+}(t_1, t_0, Y_0))).$$

The Strang splitting scheme is formed by concatenating a half-step of the Lie–Trotter scheme with a half-step of its adjoint,

$$Y_1 = \mathcal{S}^{(S)}(t_1, t_0, Y_0) = \mathcal{S}^*(t_0 + h, t_0 + h/2, \mathcal{S}(t_0 + h/2, t_0, Y_0)).$$

2.4. Error bounds. We assume that F is Lipschitz continuous and bounded,

$$\begin{aligned} &\|F(t,Y) - F(t,\widetilde{Y})\| \le L\|Y - \widetilde{Y}\| \quad \text{for all } Y, \widetilde{Y} \in \mathbb{C}^{m\times n}, \\ &\|F(t,Y)\| \le B \quad \text{for all } Y \in \mathbb{C}^{m\times n}. \end{aligned} \tag{2.2}$$

Here and in the following, the chosen norm $\|\cdot\|$ is the Frobenius norm. As usual in the numerical analysis of ordinary differential equations, this could be weakened to a local Lipschitz condition and local bound in a neighborhood of the exact solution $A(t) = \Phi_F(t, t_0, A_0)$, but for convenience we will work with the global Lipschitz condition and bound.

We further assume that $F(t, Y)$ is in the tangent space $\mathcal{T}_Y\mathcal{M}_r$ up to a small remainder, in the sense that

$$F(t, Y) = M(t, Y) + R(t, Y),$$

where M maps to the tangent bundle of $\mathcal{M}_r$ and the remainder R is small on $\mathcal{M}_r$:

$$M(t,Y) \in \mathcal{T}_Y\mathcal{M}_r \quad \text{and} \quad \|R(t,Y)\| \le \varepsilon \quad \text{for all } Y \in \mathcal{M}_r \text{ and all } t. \tag{2.3}$$

This implies that the flow of M preserves the rank for initial data $Y_0 \in \mathcal{M}_r$,

$$Y_0 \in \mathcal{M}_r \quad \Rightarrow \quad \Phi_M(t, t_0, Y_0) \in \mathcal{M}_r \quad \text{for all } t.$$

The assumption (2.3) is needed along the trajectory $\{Y(t) : 0 \le t \le T\} \subset \mathcal{M}_r$ in order to obtain an approximation error $Y(t) - A(t) = \mathcal{O}(\varepsilon)$ for the time-continuous dynamical low-rank approximation $Y(t) \in \mathcal{M}_r$. It is reasonable to make this assumption in a neighborhood on $\mathcal{M}_r$ of the trajectory. For convenience only, this assumption is made here for *all* $Y \in \mathcal{M}_r$, but we would obtain the same result if we impose the assumption only in a small neighborhood on $\mathcal{M}_r$ of the trajectory.

The obvious choice for the decomposition of F is $M(t, Y) = P(Y)F(t, Y) \in \mathcal{T}_Y\mathcal{M}_r$, where again $P(Y)$ denotes the orthogonal projection onto the tangent space. We will not use any Lipschitz bound for M, since this would involve a local Lipschitz constant of $P(Y)$, which is inversely proportional to the smallest nonzero singular value of Y and can thus become arbitrarily large. The objective in the following is to avoid invoking local Lipschitz bounds for the projections P and $P_i^\pm$ in the error analysis, so that the error bounds do not deteriorate in the presence of small singular values.

By (2.2) and (2.3), M is bounded by $B + \varepsilon$. For convenience we assume that M is also bounded by the same bound B as F, that is, $\|M(t, Y)\| \le B$ for all $Y \in \mathcal{M}_r$ and all t.

For the initial value $A_0 \in \mathbb{C}^{m\times n}$ we denote again by $A(t)$ the solution of the original problem (1.1). We assume that the initial value A_0 and the starting value $Y_0 \in \mathcal{M}_r$ of the numerical method are δ-close:

$$\|Y_0 - A_0\| \leq \delta.$$

We are now in position to state the error estimate of the projector-splitting integrators. Remarkably, this error bound is independent of the singular values of $A(t)$ and Y_n.

THEOREM 2.1. *Under the above assumptions, the errors of the Lie–Trotter and Strang splitting schemes at $t_n = t_0 + nh$ are bounded by*

$$\|Y_n - A(t_n)\| \leq c_0\delta + c_1\varepsilon + c_2 h \qquad \textit{for} \;\; t_n \leq T,$$

where c_i depend only on L, B, and T.

Section 2.5 is devoted to the proof of this theorem. There we will also obtain explicit expressions for c_i.

2.5. Proof of Theorem 2.1. The difficulty in the proof lies in the fact that we analyze the numerical integrator for the differential equation $\dot{Y} = P(Y)F(t,Y)$ in a situation in which the local Lipschitz constant of the projection $P(\cdot)$ can become arbitrarily large in any neighborhood of the solution $Y(t)$, inversely proportional to the smallest nonzero singular value of $Y(t)$ [11]. Similarly, also the local Lipschitz constants of the projections $P_i^{\pm}(\cdot)$ that are used in the splitting integrator can become arbitrarily large near matrices with small singular values. We therefore need to avoid using the Lipschitz continuity of these projections. What comes to our rescue are two ingredients:

- The exactness result of [16, Theorem 4.1], which states that when the integrator is applied to $\dot{Y}(t) = P(Y(t))\dot{X}(t)$ with $X(t) \in \mathcal{M}_r$, then it yields the exact solution $Y(t) = X(t)$ at the time gridpoints $t = t_n$.
- Range and co-range preservation under the split flows (cf. [16, Lemma 3.1]): Under the solution operators $\Phi_{F_i^{\pm}}(t,s,Y)$, the range of Y^* for $i=1$ and $+$, the range of both Y and Y^* for $-$, and the range of Y for $i=2$ and $+$, respectively, do not change, and so we have

$$P_i^{\pm}(\Phi_{F_i^{\pm}}(t,s,Y)) = P_i^{\pm}(Y) \quad \text{for all } Y \in \mathcal{M}_r. \tag{2.4}$$

Moreover, $P_i^{\pm}(Y)$ is invariant under adding a multiple of $P_i^{\pm}(Y)Z$ to the argument Y:

$$P_i^{\pm}(Y + P_i^{\pm}(Y)Z) = P_i^{\pm}(Y) \quad \text{for all } Y \in \mathcal{M}_r,\; Z \in \mathbb{C}^{m\times n}. \tag{2.5}$$

With these tools we will prove the following local error bound. Here we denote by

$$X(t) = \Phi_M(t,t_0,X_0) \in \mathcal{M}_r \tag{2.6}$$

the solution to the problem with $R=0$, i.e., $\dot{X}(t) = M(t,X(t)),\; X(t_0) = X_0 \in \mathcal{M}_r$ for an appropriately chosen initial value X_0.

LEMMA 2.2. *In the situation of Theorem 2.1, there exists an initial value $X_0 \in \mathcal{M}_r$ with $\|X_0 - Y_0\| \leq h(4BLh + 2\varepsilon)$ such that for $X(t)$ of (2.6), there is the local error bound*

$$\|Y_1 - X(t_0+h)\| \leq h(9BLh + 4\varepsilon). \tag{2.7}$$

Proof. We concentrate on the simpler Lie–Trotter scheme and only briefly mention at the end how the arguments extend to the Strang scheme. We rewrite the differential equation (1.2) as

$$\begin{aligned}\dot{Y}(t) &= P(Y(t))F(t,Y(t)) = M(t,Y(t)) + P(Y(t))R(t,Y(t))\\ &= \dot{X}(t) - M(t,X(t)) + M(t,Y(t)) + P(Y(t))R(t,Y(t))\\ &= \dot{X}(t) - F(t,X(t)) + F(t,Y(t))\\ &\qquad + \ R(t,X(t)) - R(t,Y(t)) + P(Y(t))R(t,Y(t))\end{aligned}$$

so that, with the perturbation term

$$\Delta(t,Y) = F(t,Y) - F(t,X(t)) - (I - P(Y))R(t,Y) + R(t,X(t)),$$

we have the differential equation

$$\dot{Y}(t) = \dot{X}(t) + \Delta(t,Y(t)).$$

By the Lipschitz condition on F and the bound of R, the perturbation term is bounded by

$$\|\Delta(t,Y)\| \le L\|Y - X(t)\| + 2\varepsilon. \tag{2.8}$$

Since $P_i^{\pm}(Y)P(Y) = P_i^{\pm}(Y)$, the differential equations solved in the substeps of the splitting integrator can be written as

$$\dot{Y}_i^{\pm}(t) = \pm P_i^{\pm}(Y_i^{\pm}(t))\dot{X}(t) \pm P_i^{\pm}(Y_i^{\pm}(t))\Delta(t,Y_i^{\pm}(t)), \qquad t_0 \le t \le t_0 + h,$$

which on setting $G_i^{\pm}(t,Y) = \pm P_i^{\pm}(Y)\dot{X}(t)$ and $\Delta_i^{\pm}(t,Y) = \pm P_i^{\pm}(Y)\Delta(t,Y)$ becomes

$$\dot{Y}_i^{\pm}(t) = F_i^{\pm}(t,Y_i^{\pm}(t)) = G_i^{\pm}(t,Y_i^{\pm}(t)) + \Delta_i^{\pm}(t,Y_i^{\pm}(t)), \qquad t_0 \le t \le t_0 + h.$$

We will use the fact that without the perturbation term, the initial value problem $\dot{Y}(t) = P(Y(t))\dot{X}(t)$, $Y(t_0) = X(t_0)$ with the solution $Y(t) = X(t)$ is solved exactly by the splitting integrator according to [16, Theorem 4.1]:

$$X(t_1) = \Phi_{G_2^+}(t_1,t_0,\Phi_{G_1^-}(t_1,t_0,\Phi_{G_1^+}(t_1,t_0,X(t_0)))).$$

The challenge is to bound the effect of the perturbation $\Delta(t,Y)$ without invoking Lipschitz constants of the projections P and $P_i^{\pm}$.

For each substep, we split the contributions $G_i^{\pm}$ and $\Delta_i^{\pm}$ using the Gröbner–Alekseev lemma [8, Theorem I.14.5]: with $\partial\Phi_{G_i^{\pm}}(t_1,t,Y) = (\partial/\partial Y)\Phi_{G_i^{\pm}}(t_1,t,Y)$ and $Y_i^{\pm}(t) = \Phi_{F_i^{\pm}}(t,t_0,Y_i^{\pm}(t_0))$ we have at $t_1 = t_0 + h$

$$\begin{aligned}\Phi_{F_i^{\pm}}(t_1,t_0,Y_i^{\pm}(t_0)) &= \Phi_{G_i^{\pm}}(t_1,t_0,Y_i^{\pm}(t_0))\\ &\qquad + \int_{t_0}^{t_1} \partial\Phi_{G_i^{\pm}}(t_1,t,Y_i^{\pm}(t))\Delta_i^{\pm}(t,Y_i^{\pm}(t))\,dt.\end{aligned} \tag{2.9}$$

We want to bound the integrand but cannot bound $\partial\Phi_{G_i^{\pm}}(t_1,t,Y)$ directly in the operator norm, since that would give an undesired dependence on the singular values of Y. Instead we consider the directional derivative explicitly. To simplify notation, we fix $Y \in \mathcal{M}_r$ and $Z \in \mathbb{C}^{m\times n}$ and consider expressions of the form

$$K_i^{\pm}(\tau) = \partial\Phi_{G_i^{\pm}}(t_1,\tau,Y)P_i^{\pm}(Y)Z.$$

The integrand is of this form with $Y = Y_i^\pm(t)$ and $Z = \pm\Delta(t, Y_i^\pm(t))$. At $\tau = t_1$ we get

$$K_i^\pm(t_1) = \partial\Phi_{G_i^\pm}(t_1, t_1, Y) P_i^\pm(Y) Z = P_i^\pm(Y) Z$$

since $\partial\Phi_{G_i^\pm}(t_1, t_1, Y)$ is the identity matrix.

We now show that $K_i^\pm(\tau)$ is actually independent of τ. By (2.4) and (2.5) (for $G_i^\pm$ instead of $F_i^\pm$) we have

$$P_i^\pm(\Phi_{G_i^\pm}(t_1, \tau, Y + \theta\, P_i^\pm(Y)Z)) = P_i^\pm(Y) = P_i^\pm(\Phi_{F_i^\pm}(t_1, \tau, Y)) \quad \text{for} \quad \theta \in \mathbb{R}.$$

To compute the derivative of $K_i^\pm$, we express the directional derivative with respect to the initial data by explicitly taking the limit

$$K_i^\pm(\tau) = \lim_{\theta\to 0} \frac{1}{\theta}\Big(\Phi_{G_i^\pm}(t_1, \tau, Y + \theta P_i^\pm(Y)Z) - \Phi_{G_i^\pm}(t_1, \tau, Y)\Big).$$

Then, by differentiation with respect to τ,

$$\begin{aligned}
\dot K_i^\pm(\tau) &= -\lim_{\theta\to 0} \frac{1}{\theta}\Big(G_i^\pm(\tau, \Phi_{G_i^\pm}(t_1, \tau, Y + \theta P_i^\pm(Y)Z)) - G_i^\pm(\tau, \Phi_{G_i^\pm}(t_1, \tau, Y))\Big) \\
&= \mp \lim_{\theta\to 0} \frac{1}{\theta}\Big(P_i^\pm(\Phi_{G_i^\pm}(t_1, \tau, Y + \theta P_i^\pm(Y)Z))\dot X(\tau) \\
&\qquad\qquad - P_i^\pm(\Phi_{G_i^\pm}(t_1, \tau, Y))\dot X(\tau)\Big) \\
&= \mp \lim_{\theta\to 0} \frac{1}{\theta}\Big(P_i^\pm(Y)\dot X(\tau) - P_i^\pm(Y)\dot X(\tau)\Big) = 0.
\end{aligned}$$

Hence we have

$$K_i^\pm(\tau) = P_i^\pm(Y)Z, \qquad t_0 \le \tau \le t_1.$$

We can thus determine the perturbation in (2.9) as

$$Y_i^\pm(t_1) = \Phi_{F_i^\pm}(t_1, t_0, Y_i^\pm(t_0)) = \Phi_{G_i^\pm}(t_1, t_0, Y_i^\pm(t_0)) + hE_i^\pm$$

with

$$hE_i^\pm = \pm\int_{t_0}^{t_1} P_i^\pm(Y_i^\pm(t))\Delta(t, Y_i^\pm(t))\, dt = \pm P_i^\pm(Y_i^\pm(t_0)) \int_{t_0}^{t_1} \Delta(t, Y_i^\pm(t))\, dt.$$

We obtain the result Y_1 of one step of the splitting method as

$$\begin{aligned}
Y_1 &= \Phi_{F_2^+}(t_1, t_0, \Phi_{F_1^-}(t_1, t_0, \Phi_{F_1^+}(t_1, t_0, Y_0))) \\
&= \Phi_{G_2^+}(t_1, t_0, \Phi_{G_1^-}(t_1, t_0, \Phi_{G_1^+}(t_1, t_0, Y_0) + hE_1^+) + hE_1^-) + hE_2^+.
\end{aligned}$$

The error term E_2^+ is directly of the form required by (2.7); next we move also E_1^+ and E_1^- out of the splitting scheme. In the following we use for the splitting scheme the notation

$$\begin{aligned}
Y_1^+ &= \Phi_{F_1^+}(t_1, t_0, Y_0) = \Phi_{G_1^+}(t_1, t_0, Y_0) + hE_1^+, \\
Y_1^- &= \Phi_{F_1^-}(t_1, t_0, Y_1^+) = \Phi_{G_1^-}(t_1, t_0, Y_1^+) + hE_1^-, \\
Y_1 &= \Phi_{F_2^+}(t_1, t_0, Y_1^-) = \Phi_{G_2^+}(t_1, t_0, Y_1^-) + hE_2^+.
\end{aligned}$$

Since the solution operator $\Phi_{F_1^-}(t, t_0, Y_1^+)$ preserves both the range and co-range of Y_1^+, and since $P_2^+(Y)$ is the projection onto the range of Y, we have

$$P_2^+(Y_1^-) = P_2^+(Y_1^+).$$

Since also the solution operator $\Phi_{G_2^+}$ preserves the range, this yields

$$\begin{aligned}
&\frac{d}{dt}\Big(\Phi_{G_2^+}(t, t_0, Y_1^-) - \Phi_{G_2^+}(t, t_0, \Phi_{G_1^-}(t_1, t_0, Y_1^+))\Big)\\
&= G_2^+(t, \Phi_{G_2^+}(t, t_0, Y_1^-)) - G_2^+(t, \Phi_{G_2^+}(t, t_0, \Phi_{G_1^-}(t_1, t_0, Y_1^+)))\\
&= P_2^+(\Phi_{G_2^+}(t, t_0, Y_1^-))\dot{X}(t) - P_2^+(\Phi_{G_2^+}(t, t_0, \Phi_{G_1^-}(t_1, t_0, Y_1^+)))\dot{X}(t)\\
&= P_2^+(Y_1^-)\dot{X}(t) - P_2^+(\Phi_{G_1^-}(t_1, t_0, Y_1^+))\dot{X}(t)\\
&= P_2^+(Y_1^+)\dot{X}(t) - P_2^+(Y_1^+)\dot{X}(t)\\
&= 0.
\end{aligned}$$

It follows that

$$\Phi_{G_2^+}(t_1, t_0, Y_1^-) - \Phi_{G_2^+}(t_1, t_0, \Phi_{G_1^-}(t_1, t_0, Y_1^+)) = Y_1^- - \Phi_{G_1^-}(t_1, t_0, Y_1^+) = hE_1^-,$$

and hence

$$\begin{aligned}
Y_1 &= \Phi_{G_2^+}(t_1, t_0, Y_1^-) + hE_2^+\\
&= \Phi_{G_2^+}(t_1, t_0, \Phi_{G_1^-}(t_1, t_0, Y_1^+)) + hE_1^- + hE_2^+\\
&= \Phi_{G_2^+}(t_1, t_0, \Phi_{G_1^-}(t_1, t_0, \Phi_{G_1^+}(t_1, t_0, Y_0) + hE_1^+)) + hE_1^- + hE_2^+.
\end{aligned}$$

To expel the final error term hE_1^+ from the arguments on the right-hand side, we note that since $P_1^+(Y_0)E_1^+ = E_1^+$ we get as above, this time using the conservation of the co-range,

$$\frac{d}{dt}\Big(\Phi_{G_1^+}(t, t_0, Y_0 + hE_1^+) - \Phi_{G_1^+}(t, t_0, Y_0)\Big) = 0$$

and hence

$$\Phi_{G_1^+}(t_1, t_0, Y_0 + hE_1^+) - \Phi_{G_1^+}(t_1, t_0, Y_0) = (Y_0 + hE_1^+) - Y_0 = hE_1^+, \tag{2.10}$$

so that

$$Y_1 = \Phi_{G_2^+}(t_1, t_0, \Phi_{G_1^-}(t_1, t_0, \Phi_{G_1^+}(t_1, t_0, Y_0 + hE_1^+))) + hE_1^- + hE_2^+.$$

By the exactness result of [16, Theorem 4.1] we obtain on choosing $X_0 = Y_0 + hE_1^+$

$$\Phi_{G_2^+}(t_1, t_0, \Phi_{G_1^-}(t_1, t_0, \Phi_{G_1^+}(t_1, t_0, Y_0 + hE_1^+))) = X(t_1)$$

so that

$$Y_1 = X(t_1) + hE_1^- + hE_2^+.$$

Finally, we bound $E_i^\pm$ using (2.8). With $X_0 = Y_0 + hE_1^+$, (2.10) gives

$$Y_1^+(t_1) = \Phi_{G_1^+}(t_1, t_0, Y_0) + hE_1^+ = \Phi_{G_1^+}(t_1, t_0, X_0).$$

Then, by the bound B of F and M, we have

$$\|Y_1^+(t_1) - X(t_1)\| \le \|\Phi_{G_1^+}(t_1, t_0, X_0) - X_0\| + \|X(t_1) - X_0\| \le 2Bh,$$

and for $t_0 \le t \le t_1 = t_0 + h$

$$\|Y_1^+(t) - X(t)\| \le \|Y_1^+(t) - Y_1^+(t_1)\| + \|Y_1^+(t_1) - X(t_1)\| + \|X(t_1) - X(t)\| \le 4Bh.$$

Since $Y_1^-(t_0) = Y_1^+(t_1)$, we have further

$$\|Y_1^-(t) - X(t)\| \le \|Y_1^-(t) - Y_1^-(t_0)\| + \|Y_1^+(t_1) - X(t_1)\| + \|X(t_1) - X(t)\| \le 4Bh,$$

and similarly

$$\|Y_2^+(t) - X(t)\| \le 5Bh.$$

Hence we obtain from (2.8)

$$\|E_1^+\| \le 4BLh + 2\varepsilon, \quad \|E_1^-\| \le 4BLh + 2\varepsilon, \quad \|E_2^+\| \le 5BLh + 2\varepsilon.$$

This concludes the proof for the Lie–Trotter scheme. The same result holds for the adjoint scheme and hence also for the Strang splitting scheme. ◻

Proof of Theorem 2.1. Using the bound of R and the Lipschitz continuity of F, we obtain by Grönwall's inequality for $X(t) = \Phi_{F-R}(t, t_0, X_0)$ at $t_1 = t_0 + h$

$$\|\Phi_F(t_1, t_0, Y_0) - X(t_1)\| \le e^{Lh}(h(4BLh + 2\varepsilon) + h\varepsilon),$$

which together with Lemma 2.2 yields an estimate of the local error $Y_1 - \Phi_F(t_1, t_0, Y_0)$. Since the solution operator $\Phi_F(t, s, \cdot)$ satisfies, by the Lipschitz continuity of F and the Grönwall inequality,

$$(2.11) \quad \|\Phi_F(t, s, A) - \Phi_F(t, s, \widetilde{A})\| \le e^{L(t-s)}\|A - \widetilde{A}\| \quad \text{for all } A, \widetilde{A} \in \mathbb{C}^{m\times n}, \ t > s,$$

the result of Theorem 2.1 is obtained from Lemma 2.2 with the standard argument of Lady Windermere's fan [8, II.3] with error propagation by Φ_F. We obtain the stated bound with $c_0 = e^{L(T-t_0)}$, with $c_1 = (4 + 3e^{Lh_0})(e^{L(T-t_0)} - 1)/L$, where h_0 is an upper bound of the stepsize h, and with $c_2 = (9 + 4e^{Lh_0})B(e^{L(T-t_0)} - 1)$. ◻

2.6. Remarks and extensions. We discuss a special case and two modifications of Theorem 2.1.

2.6.1. Low-rank approximation of given time-dependent matrices. Consider the case where $A(t)$ are given time-dependent matrices to which approximations of rank r are sought. In this case the integrator of section 2.2, with $F(t, Y) = \dot{A}(t)$, just uses the increments $A(t_{n+1}) - A(t_n)$. If

$$A(t) = X(t) + R(t) \quad \text{with } X(t) \text{ of rank } r \text{ and } \|R(t_0)\| \le \delta, \ \|\dot{R}(t)\| \le \varepsilon,$$

then we are in the situation of Theorem 2.1, where $F(t, Y) = \dot{A}(t)$ is independent of Y, so that the Lipschitz constant is $L = 0$. As a consequence, we get $c_2 = 0$ and hence the error bound becomes independent of the stepsize h,

$$\|Y_n - A(t_n)\| \le \delta + 7(t_n - t_0)\varepsilon, \qquad t_0 \le t_n \le T.$$

2.6.2. Functions with a one-sided Lipschitz condition. Suppose that with respect to the Frobenius inner product $\langle \cdot, \cdot \rangle$ we have the one-sided Lipschitz bound

$$\langle F(t,Y) - F(t,\widetilde{Y}), Y - \widetilde{Y} \rangle \le \ell \|Y - \widetilde{Y}\|^2 \quad \text{for all } Y, \widetilde{Y} \in \mathbb{C}^{m \times n}$$

with $\ell \le L$ and possibly $\ell \ll L$. In this case the error propagation improves to

$$\|\Phi_F(t,s,A) - \Phi_F(t,s,\widetilde{A})\| \le e^{\ell(t-s)} \|A - \widetilde{A}\| \quad \text{for all } \; A, \widetilde{A} \in \mathbb{C}^{m \times n}, \; t > s,$$

where the factor $e^{L(t-s)}$ from (2.11) is replaced by the smaller factor $e^{\ell(t-s)}$. The above proof then yields an error bound as in Theorem 2.1 with improved constants

$$c_0 = e^{\ell(T-t_0)}, \quad c_1 = (4 + 3e^{\ell h_0})(e^{\ell(T-t_0)} - 1)/\ell, \quad c_2 = (9 + 4e^{\ell h_0}) BL (e^{\ell(T-t_0)} - 1)/\ell.$$

However, we are not able to avoid the linear dependence on the Lipschitz constant L in c_2, which stems from (2.8).

2.6.3. Inexact solution of the differential equations in the substeps of the splitting scheme. Suppose that instead of the exact value $Y_i^{\pm}(t_1)$ only an approximate value

$$\widetilde{Y}_i^{\pm}(t_1) = \Phi_{F_i^{\pm}}(t_1, t_0, \widetilde{Y}_i^{\pm}(t_0)) + h\widetilde{E}_i^{\pm}$$

is computed, so that in one full step of the method, instead of Y_1 one actually computes

$$\widetilde{Y}_1 = \Phi_{F_2^+}(t_1, t_0, \Phi_{F_1^-}(t_1, t_0, \Phi_{F_1^+}(t_1, t_0, Y_0) + h\widetilde{E}_1^+) + h\widetilde{E}_1^-) + h\widetilde{E}_2^+.$$

Suppose now that the errors satisfy

$$\|\widetilde{E}_i^{\pm}\| \le \eta, \qquad P_i^{\pm}(\widetilde{Y}_i^{\pm}(t_0))\widetilde{E}_i^{\pm} = \widetilde{E}_i^{\pm}.$$

The latter equation is a natural condition from the way the differential equations for the factors U, S, V of $Y = USV^*$ are actually solved in the algorithm of section 2.2. In fact, if some numerical integrator is applied in the first substep of the algorithm, then instead of $K(t_1)$ a perturbed value $K(t_1) + hE_K(t_1)$ is computed. We then have $\widetilde{E}_1^+ = E_K(t_1)V_0^* = E_K(t_1)V_0^* V_0 V_0^* = P(Y_1^+(t_0))\widetilde{E}_1^+$, and similarly for the second and third substeps.

In this situation the error bound of Theorem 2.1 changes, with the same proof, to

$$\|\widetilde{Y}_n - A(t_n)\| \le c_0 \delta + c_1 \varepsilon + c_2 h + c_3 \eta,$$

where c_0, c_1, c_2 are as before, and $c_3 = (2 + e^{\ell h_0})(e^{\ell(T-t_0)} - 1)/\ell$.

3. Time integration of tensor trains. We next extend the results to tensor differential equations and their low-rank approximation in tensor-train format. After introducing the necessary notation below, we prove a tensor analogue to Theorem 2.1 in section 3.2.

3.1. The splitting scheme for TT-tensors. A tensor $Y \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ is in tensor train format if there exist core tensors $C_i \in \mathbb{C}^{r_{i-1} \times n_i \times r_i}$ of full multilinear rank, such that every element of Y can be written as

$$Y(l_1, ..., l_d) = \sum_{j_1=1}^{r_1} \cdots \sum_{j_{d-1}=1}^{r_{d-1}} C_1(1, l_1, j_1) \cdots C_i(j_{i-1}, l_i, j_i) \cdots C_d(j_{d-1}, l_d, 1),$$

where $l_i = 1, \dots, n_i$, $i = 1, \dots, d$. The TT-rank of Y is defined as the vector $\mathbf{r} = (r_0, r_1, \dots, r_d) \in \mathbb{N}^{d+1}$ with $r_0 = r_d = 1$. The set of all $n_1 \times \cdots \times n_d$-tensors of a given TT-rank $\mathbf{r}$ is a manifold [10, 24], which we denote by $\mathcal{N}_{\mathbf{r}}$.

As a direct generalization of the matrix case, consider the dth order tensor $A(t) \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ which solves the tensor differential equation

$$\dot{A}(t) = F(t, A(t)), \qquad A(t_0) = A_0 \in \mathbb{C}^{n_1 \times \cdots \times n_d}.$$

We seek to approximate $A(t)$ by a tensor $Y(t)$ in the manifold $\mathcal{N}_{\mathbf{r}}$ of $n_1 \times \cdots \times n_d$-tensors of TT-rank $\mathbf{r}$. To find the approximate solution $Y(t) \in \mathcal{N}_{\mathbf{r}}$ we consider—in the same manner as for the matrix differential equation—the evolution equation

$$\dot{Y}(t) = P(Y(t))F(t, Y(t)), \qquad Y(t_0) = Y_0 \in \mathcal{N}_{\mathbf{r}}, \tag{3.1}$$

where $P(Y(t))$ is the orthogonal projection onto the tangent space $\mathcal{T}_{Y(t)}\mathcal{N}_{\mathbf{r}}$ of $\mathcal{N}_{\mathbf{r}}$ at $Y(t)$. $P(Y)$ can be decomposed as [17]

$$P(Y) = \sum_{i=1}^{d-1} \left(P_{\le i-1} P_{\ge i+1} - P_{\le i} P_{\ge i+1} \right) + P_{\le d-1} P_{\ge d+1}.$$

Also $P_{\le i}$ and $P_{\ge i}$ are projections. Note that $P_{\le i} = P_{\le i}(Y)$ and $P_{\ge i} = P_{\ge i}(Y)$ depend on Y. To simplify the notation, we will not denote their Y-dependence explicitly when this is possible without causing confusion. $P_{\le i}$ and $P_{\ge i}$ are constructed using a generalization of singular vectors to higher-dimensional tensors. For a more detailed definition, we refer to [17]. $P_{\le i}$ operates on the core tensors C_k with $k = 1, \dots, i$ and for $i \le j$, $P_{\le i}$ is a "subprojection" of $P_{\le j}$, i.e., $P_{\le i} P_{\le j} Z = P_{\le j} Z$ for all $Z \in \mathbb{C}^{n_1 \times \cdots \times n_d}$. The situation for $P_{\ge i}$, which operates on C_k with $k = i, \dots, d$, is analogous. $P_{\le 0}$ and $P_{\ge d+1}$ are both the identity operator. When $i < j$, $P_{\le i}$ and $P_{\ge j}$ commute. These properties will be important in our proof of the error estimate below. We also introduce the abbreviated notation

$$P_i^+ Z = P_{\le i-1} P_{\ge i+1} Z, \qquad P_i^- Z = P_{\le i} P_{\ge i+1} Z.$$

Hence we can write the projection onto the tangent space as

$$P(Y) = \sum_{i=1}^{d-1} \left(P_i^+ - P_i^- \right) + P_d^+. \tag{3.2}$$

Note the similarity with the projection decomposition (2.1) from the matrix case.

The Lie–Trotter splitting scheme amounts to splitting the right-hand side in (3.1) according to the decomposition (3.2) of the projection and solving the resulting $2d-1$ subproblems sequentially. The subproblems to be solved read

$$\begin{aligned}
\dot{Y}_1^+(t) &= P_1^+(Y_1^+(t))F(t, Y_1^+(t)), & Y_1^+(t_0) &= Y_0, & \\
\dot{Y}_i^-(t) &= -P_i^-(Y_i^-(t))F(t, Y_i^-(t)), & Y_i^-(t_0) &= Y_i^+(t_0 + h) & &\text{for } i = 1, \dots, d-1, \\
\dot{Y}_i^+(t) &= P_i^+(Y_i^+(t))F(t, Y_i^+(t)), & Y_i^+(t_0) &= Y_{i-1}^-(t_0 + h) & &\text{for } i = 2, \dots, d.
\end{aligned}$$

We refer to [17] for the practical algorithm, which for each i works with small 3-tensors and combines them in a forward sweep from 1 to d. The adjoint method makes a backward sweep from d down to 1.

As in the matrix case, the projections $P_i^{\pm}$ are constant during the solution of the corresponding subproblems. The proof of this is a simple adaptation of [17, Theorem 4.1]. Using a similar argument one can show that along the solution of

$$\dot{Y}(t) = P_{\le i}(Y(t))P_{\ge j}(Y(t))F(t, Y(t)), \quad Y(t_0) = Y_0 \in \mathcal{N}_{\mathbf{r}},$$

the projections $P_{\le i}$ and $P_{\ge j}$ are preserved when $i < j$,

$$P_{\le i}(Y(t)) = P_{\le i}(Y_0), \quad P_{\ge j}(Y(t)) = P_{\ge j}(Y_0) \quad \text{for all } t \ge t_0.$$

This property will be used in the proof of Theorem 3.1, which is the generalization of Theorem 2.1 to tensor trains.

With $F_i^{\pm}(t, Y) = \pm P_i^{\pm}(Y)F(t, Y)$, we denote a step with the splitting scheme as

$$Y_1 = \mathcal{S}(t_1, t_0, Y_0) = \Phi_{F_d^+}(t_1, t_0, \Phi_{F_{d-1}^-}(t_1, t_0, \Phi_{F_{d-1}^+}(\cdots \Phi_{F_1^+}(t_1, t_0, Y_0) \cdots))).$$

The Strang splitting scheme is, as previously, formed by concatenating the Lie–Trotter scheme with its adjoint.

3.2. Error bounds. We can prove an error estimate similar to Theorem 2.1 also for the tensor case, under similar assumptions. We assume that $F(Y)$ is Lipschitz continuous and bounded,

$$\begin{aligned}&\|F(t, Y) - F(t, \widetilde{Y})\| \le L\|Y - \widetilde{Y}\| \quad \text{for all } Y, \widetilde{Y} \in \mathbb{C}^{n_1 \times \cdots \times n_d},\\ &\|F(t, Y)\| \le B \quad \text{for all } Y \in \mathbb{C}^{n_1 \times \cdots \times n_d},\end{aligned}$$

and that it can be subdivided as $F(t, Y) = M(t, Y) + R(t, Y)$ with

$$M(t, Y) \in \mathcal{T}_Y \mathcal{N}_{\mathbf{r}} \quad \text{and} \quad \|R(t, Y)\| \le \varepsilon \quad \text{for all } Y \in \mathcal{N}_{\mathbf{r}} \text{ and all } t.$$

We assume for convenience that also M is bounded by B. We also assume that the perturbation in the initial value is bounded by

$$\|Y_0 - A_0\| \le \delta.$$

We can then prove the following error estimate.

THEOREM 3.1. *Under the above assumptions, the Lie–Trotter and Strang splitting schemes satisfy the error estimate*

$$\|Y_n - A(t_n)\| \le c_0\delta + c_1\varepsilon + c_2 h \quad \text{for } t_n \le T, \tag{3.3}$$

where c_i only depend on L, B, T, and the dimension d.

The same remarks and extensions as for Theorem 2.1 (see section 2.6) apply also to Theorem 3.1.

Proof. The result is obtained by induction on the dimension d and using the result for the matrix case, which corresponds to $d = 2$. For a d-dimensional tensor train $Y \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ with ranks $r_1, \ldots, r_d$, the mode-1 matricization $Y^{\langle 1 \rangle} \in \mathbb{C}^{n_1 \times (n_2 \ldots n_d)}$ has a factorization

$$Y^{\langle 1 \rangle} = USV^*,$$

where $U \in \mathbb{C}^{n_1 \times r_1}$ and $V \in \mathbb{C}^{(n_2 \ldots n_d) \times r_1}$ have orthogonal columns, and $S \in \mathbb{C}^{r_1 \times r_1}$. The projection onto the tangent space of the manifold of (re-tensorized) rank-r_1 matrices at $Y^{\langle 1 \rangle}$ is then

$$P_1^+(Y) - P_1^-(Y) + P_{\ge 2}^+(Y),$$

where $P_1^{\pm}(Y)$ are the projections appearing in (3.2), and $P_{\geq 2}^{+}(Y)$ is the orthogonal projection onto the range of $Y^{\langle 1 \rangle}$: for $Z \in \mathbb{C}^{n_1 \times \cdots \times n_d}$,

$$\Big(P_1^{+}(Y)Z\Big)^{\langle 1 \rangle} = Z^{\langle 1 \rangle} V V^*, \qquad \Big(P_1^{-}(Y)Z\Big)^{\langle 1 \rangle} = U U^* Z^{\langle 1 \rangle} V V^*,$$
$$\Big(P_{\geq 2}^{+}(Y)Z\Big)^{\langle 1 \rangle} = U U^* Z^{\langle 1 \rangle}.$$

The first two substeps of the tensor-train projector-splitting integrator for the approximation of $\dot{Y} = F(t, Y)$ are the same as in the above matrix projector-splitting algorithm. The third substep of the matrix projector-splitting integrator solves the differential equation

$$\dot{Y}_{\geq 2} = P_{\geq 2}^{+}(Y_{\geq 2})\, F(t, Y_{\geq 2}), \tag{3.4}$$

where $P_{\geq 2}^{+}(Y_{\geq 2}(t))$ is the orthogonal projection onto a fixed subspace independent of t, since $\bar{U}$ is not changed any more.

In the further substeps of the tensor-train projector-splitting integrator, the differential equation (3.4) is solved inexactly by the $(d-1)$-dimensional tensor-train projector-splitting integrator, splitting the projection

$$\sum_{i=2}^{d-1} \big(P_i^{+}(Y) - P_i^{-}(Y)\big) + P_d^{+}(Y).$$

This sum is the $(d-1)$-dimensional tensor-train tangent space projection at Y in the fixed subspace defined by the range of $Y^{\langle 1 \rangle}$. We note that $P_i^{\pm}(Y) P_{\geq 2}^{+}(Y) = P_i^{\pm}(Y)$ for $i \geq 2$. This yields that the substeps of the $(d-1)$-dimensional tensor-train projector-splitting integrator for (3.4) are identical to those of the d-dimensional projector-splitting integrator applied to $\dot{Y} = F(t, Y)$ from the third substep onward.

By the induction hypothesis, the error of the result $(Y_{\geq 2})_1$ obtained after one time step of the $(d-1)$-dimensional tensor-train projector-splitting integrator is bounded by

$$\|(Y_{\geq 2})_1 - Y_{\geq 2}(h)\| \leq C h \eta \quad \text{with} \quad \eta = \varepsilon + h.$$

We are thus in the situation of the matrix projector-splitting integrator with inexact solution of the substeps up to a local error $O(h\eta)$. By the result in section 2.6.3, we obtain for the local error of the integrator

$$Y_1 - Y(h) = O(h(\delta + \varepsilon + h + \eta)) = O(h(\delta + \varepsilon + h))$$

and for the global error

$$Y_n - Y(t_n) = O(\delta + \varepsilon + h) \qquad \text{for} \;\; t_n \leq T,$$

where the constants symbolized by the O-notation depend only on L, B, T, and d. □

4. Numerical experiments.

We present three numerical examples which corroborate the theoretical results. The final experiment additionally indicates that the method is robust to stiff problems.

4.1. A discrete nonlinear Schrödinger equation. We consider a discrete nonlinear Schrödinger equation, modeling a Bose–Einstein condensate in an optical lattice [23]. The problem reads

$$\begin{aligned} i\dot{A}(t) &= -\frac{1}{2}TA(t) - \frac{1}{2}A(t)T - \varepsilon|A(t)|^2 \bullet A(t), \\ A_{jk}(0) &= \exp(-(j-j_1)^2/\sigma^2 - (k-k_1)^2/\sigma^2) \\ &\quad - \exp(-(j-j_2)^2/\sigma^2 - (k-k_2)^2/\sigma^2), \qquad j,k = 1,\dots,n, \end{aligned} \tag{4.1}$$

where $T = \mathrm{tridiag}\,(1,0,1)$, the squared modulus is taken elementwise, and $\bullet$ denotes the elementwise product. We use $n = 100$, $\sigma = 10$, $(j_1,k_1) = (60,50)$, and $(j_2,k_2) = (50,40)$. Note that T is *not* a discretized derivative but a bounded operator modeling the coupling between nodes in the lattice. Since the Frobenius norm of the exact solution is conserved, the right-hand side of (4.1) is bounded and Lipschitz continuous in a neighborhood around the exact solution.

We let $Y(t)$ denote an approximation to $A(t)$ on the low-rank manifold $\mathcal{M}_r$ with rank $r = 10$. The linear terms in (4.1) map onto the tangent space $\mathcal{T}_{Y(t)}\mathcal{M}_r$, while the nonlinear term does not. This makes the dependence of the error on ε explicit. In Table 1 we study the effect of varying ε and the time step h. We show the error in Frobenius norm after solving the problem up to $t = 5$ with different ε and h, using the Lie–Trotter projector-splitting scheme. Each subproblem is solved using the fourth order Runge–Kutta method with time step $h = 0.001$. The approximate solution is compared to a full rank reference solution, computed with fourth order Runge–Kutta using the time step $h = 0.0005$. We see how the error decays with ε as predicted. We also see convergence with respect to h in the bottom rows of the table, albeit not of a clear order. The unclear convergence rate with respect to h may be an indication that the error estimate is not quite sharp. The results in section 4.3 for a discretized partial differential equation, which look good despite violating the assumption of F being Lipschitz continuous, also indicate that there is more to learn about the projector-splitting integrator.

Table 1

Error in Frobenius norm after solving (4.1) *using the Lie–Trotter splitting scheme with different* ε *and* h.

$\varepsilon \setminus h$	1	10^{-1}	10^{-2}	10^{-3}
1	9.83e-2	9.73e-2	9.73e-2	9.73e-2
10^{-1}	1.32e-4	8.63e-5	8.63e-5	8.63e-5
10^{-2}	3.13e-6	3.51e-7	3.44e-7	3.44e-7
10^{-3}	2.47e-7	3.44e-9	1.26e-9	1.26e-9
10^{-4}	2.19e-8	2.58e-10	4.09e-11	4.00e-11

The results for the Strang splitting scheme look similar. If we use the double time step for Strang, such that the same total number of stages is used, only a few of the numbers in Table 1 change in the third digit. When there are no small nonzero singular values, the standard error estimates for splitting methods are valid and the Strang splitting scheme converges toward $Y(kh)$ at second order in h. In the presence of small singular values, however, this does not seem to be the case, and due to the h-dependence in the bounds of the remainder term $\|\Delta(t,Y)\|$ this is also not promised by our analysis.

4.2. Addition of matrices and tensors. We consider the addition of two tensors, $C = N + A$, where $N \in \mathcal{N}_{\mathbf{r}}$ is an $n_1 \times \cdots \times n_d$-tensor of TT-rank $\mathbf{r}$, and

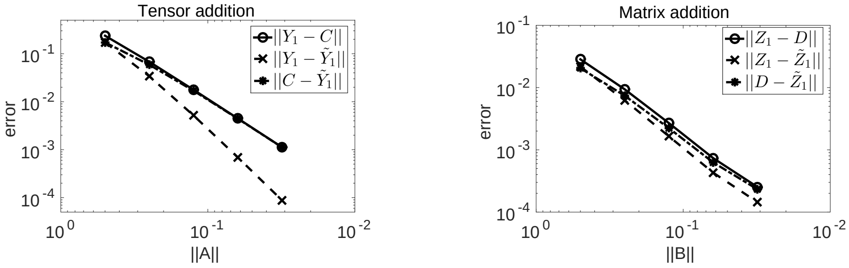

FIG. 2. *Error for tensor (left) and matrix (right) addition using the projector-splitting method for tangential increments A and B of decreasing norms.*

$A \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ is an increment. The sum is to be computed approximately with a result in $\mathcal{N}_{\mathbf{r}}$. Such truncated (or retracted) additions are required in iterative methods on low-rank tensor manifolds, in particular in optimization problems; see, e.g., [1].

A standard approach is to first compute $N + A$, which is in $\mathcal{N}_{2\mathbf{r}}$ in the case that $A \in \mathcal{T}_N \mathcal{N}_{\mathbf{r}}$, and then to project the result to $\mathcal{N}_{\mathbf{r}}$ using a TT-SVD [20]. Note that the TT-SVD gives a quasi-optimal approximation [20, Corollary 2.4] on the manifold but not the best approximation, as is the case for the SVD of matrices.

Alternatively, as proposed in [1, 17], we can perform approximate addition on the low-rank manifold using the projector-splitting integrator. We then solve

$$\dot{Y}(t) = P(Y)A, \qquad Y(0) = N \tag{4.2}$$

up to $t = 1$ using one step of the Lie–Trotter projector-splitting scheme. As before, $P(Y)$ denotes the orthogonal projection onto the tangent space $\mathcal{T}_Y \mathcal{N}_{\mathbf{r}}$ at $Y \in \mathcal{N}_{\mathbf{r}}$. We then get an approximation $Y_1 \in \mathcal{N}_{\mathbf{r}}$ for $C = N + A$. Note that we never leave the low-rank manifold $\mathcal{N}_{\mathbf{r}}$ when using the splitting method.

In our numerical example we illustrate this procedure for an example with small singular values. We construct $N \in \mathbb{C}^{100 \times 100 \times 100 \times 100}$ of TT-rank $\mathbf{r} = (1, 10, 10, 10, 1)$ with orthogonalized cores and with the singular values of its matricizations decreasing exponentially as $\sigma_j = e^{-j}$, $j = 1, \ldots, 10$. We let A be a random tensor in $\mathcal{T}_N \mathcal{N}_{\mathbf{r}}$. We also consider the similar matrix addition $D = M + B$, where $M \in \mathbb{C}^{100 \times 100}$ is of rank $s = 10$ and has decreasing singular values as in the tensor case. B is a random tensor in $\mathcal{T}_M \mathcal{M}_s$. We add $C = N + A$ and $D = M + B$ directly to get the full-rank solutions and compare this with solving (4.2) using the Lie–Trotter splitting integrator, which gives us the low-rank solutions Y_1 and Z_1, respectively. We compare with the projections $\tilde{Y}_1$ and $\tilde{Z}_1$ obtained with TT-SVD and SVD, respectively. These comparisons are illustrated in Figure 2. We see how the error of the splitting method decays as the norm of the increments A and B is reduced. Note also how close the solution given by the splitting method is to the (TT-)SVD approximation, at reduced computational cost.

4.3. The time-dependent Schrödinger equation. Since the results in this paper rely on boundedness and Lipschitz continuity of F, they do not transfer directly to stiff problems such as spatially discretized partial differential equations. Numerical evidence, however, suggests that the projector-splitting scheme is robust and accurate also in this case. We conclude the paper with such an example. We consider the time-

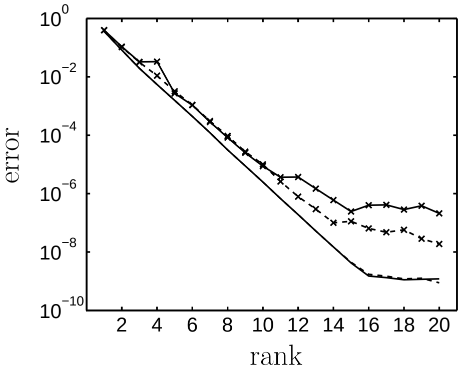

FIG. 3. *Error at different approximation ranks when solving the Schrödinger equation on an $n \times n$ spatial grid. We use $n = 64$ (dashed) and $n = 128$ (solid) grid points per dimension, and the time steps $h = 0.02$ ($\times$) and $h = 0.01$ (plain).*

dependent Schrödinger equation in two dimensions with a harmonic potential,

$$\begin{aligned} iu_t(x,t) &= -\frac{1}{2}\Delta u(x,t) + \frac{1}{2}x^T Ax\, u(x,t), \qquad x \in \mathbb{R}^2,\ t > 0,\\ u(x,0) &= \pi^{-1/2}\exp\left(\frac{1}{2}x_1^2 + \frac{1}{2}(x_2-1)^2\right)\\ \text{with} \qquad A &= \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}. \end{aligned}$$

As the right-hand side contains a second order differential operator, its Lipschitz constant scales as Δx^{-2}, where Δx is the spatial stepsize. While the initial data is of rank 1, the nondiagonal potential will increase the effective rank of the solution during time evolution. We discretize the problem using Fourier collocation with $n \times n$ grid points on $\Omega = [-7.5, 7.5]^2$. The spatially localized solution is essentially supported within Ω. The approximate solution at the respective grid points is arranged in an $n \times n$ matrix. We solve low-rank approximations to the problem with ranks $r = 1, 2, \ldots, 20$ using the Lie–Trotter splitting scheme, integrating up to the time $t = 5$. We use $n = 64$ and $n = 128$, and time steps of length $h = 0.02$ and $h = 0.01$. The subproblems are solved to high accuracy by approximating the action of the matrix exponential in a Krylov subspace generated by the Arnoldi process. We compare the low-rank approximation to a full-rank reference solution computed by standard Fourier collocation and Arnoldi time stepping with $n = 128$, $h = 0.01$. The error, depicted in Figure 3, is measured in the Frobenius norm, scaled such that it approximates the continuous $L^2(\Omega)$-norm. The error decreases exponentially with the rank, which indicates that the method is robust with respect to small singular values also for stiff problems. For the time step $h = 0.02$ we see how the error at high approximation ranks is slightly larger for the finer spatial grid, suggesting a dependence on the Lipschitz constant. The dependence is, however, mild and the method much more robust with respect to stiffness than explained by the theory presented in this paper.

REFERENCES

[1] P.-A. Absil and I. V. Oseledets, *Low-rank retractions: A survey and new results*, Comput. Optim. Appl., 62 (2015), pp. 5–29.

[2] A. Arnold and T. Jahnke, *On the approximation of high-dimensional differential equations in the hierarchical Tucker format*, BIT, 54 (2014), pp. 305–341.

[3] P. A. M. Dirac, *Note on exchange phenomena in the Thomas atom*, Math. Proc. Cambridge Philos. Soc., 26 (1930), pp. 376–385.

[4] L. Grasedyck, D. Kressner, and C. Tobler, *A literature survey of low-rank tensor approximation techniques*, GAMM-Mitt., 36 (2013), pp. 53–78.

[5] W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer, Berlin, 2012.

[6] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, *Unifying Time Evolution and Optimization with Matrix Product States*, preprint, arXiv:1408.5056, 2014.

[7] J. Haegeman, T. J. Osborne, and F. Verstraete, *Post-matrix product state methods: To tangent space and beyond*, Phys. Rev. B, 88 (2013), 075133.

[8] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I. *Nonstiff Problems*, 2nd ed., Springer, Berlin, 1993.

[9] M. Hochbruck and C. Lubich, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.

[10] S. Holtz, T. Rohwedder, and R. Schneider, *On manifolds of tensors of fixed TT-rank*, Numer. Math., 120 (2012), pp. 701–731.

[11] O. Koch and C. Lubich, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 434–454.

[12] O. Koch and C. Lubich, *Dynamical tensor approximation*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2360–2375.

[13] P. Kramer and M. Saraceno, *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, Lecture Notes in Phys. 140, Springer, Berlin, 1981.

[14] C. Lubich, *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*, European Mathematical Society, Zürich, 2008.

[15] C. Lubich, *Low-rank dynamics*, in Extraction of Quantifiable Information from Complex Systems, S. Dahlke, ed., Lect. Notes Comput. Sci. Eng., 102, Springer, Berlin, 2014, pp. 381–396.

[16] C. Lubich and I. V. Oseledets, *A projector-splitting integrator for dynamical low-rank approximation*, BIT, 54 (2014), pp. 171–188.

[17] C. Lubich, I. V. Oseledets, and B. Vandereycken, *Time integration of tensor trains*, SIAM J. Numer. Anal., 53 (2015), pp. 917–941.

[18] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken, *Dynamical approximation by hierarchical Tucker and tensor-train tensors*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 470–494.

[19] A. Nonnenmacher and C. Lubich, *Dynamical low-rank approximation: applications and numerical experiments*, Math. Comput. Simulation, 79 (2008), pp. 1346–1357.

[20] I. V. Oseledets, *Tensor-train decomposition*, SIAM J. Sci. Comput., 33 (2011), pp. 2295–2317.

[21] Y. Saad, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), pp. 209–228.

[22] U. Schollwöck, *The density-matrix renormalization group in the age of matrix product states*, Ann. Physics, 326 (2011), pp. 96–192.

[23] A. Trombettoni and A. Smerzi, *Discrete solitons and breathers with dilute Bose–Einstein condensates*, Phys. Rev. Lett., 86 (2001), pp. 2353–2356.

[24] A. Uschmajew and B. Vandereycken, *The geometry of algorithms using hierarchical tensors*, Linear Algebra Appl., 439 (2013), pp. 133–166.

[25] F. Verstraete, V. Murg, and J. I. Cirac, *Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems*, Adv. Phys., 57 (2008), pp. 143–224.