SIAM J. MATRIX ANAL. APPL.
Vol. 29, No. 2, pp. 434–454

DYNAMICAL LOW-RANK APPROXIMATION*

OTHMAR KOCH† AND CHRISTIAN LUBICH†

Abstract. For the low-rank approximation of time-dependent data matrices and of solutions to matrix differential equations, an increment-based computational approach is proposed and analyzed. In this method, the derivative is projected onto the tangent space of the manifold of rank-r matrices at the current approximation. With an appropriate decomposition of rank-r matrices and their tangent matrices, this yields nonlinear differential equations that are well suited for numerical integration. The error analysis compares the result with the pointwise best approximation in the Frobenius norm. It is shown that the approach gives locally quasi-optimal low-rank approximations. Numerical experiments illustrate the theoretical results.

Key words. low-rank approximation, time-varying matrices, continuous updating, smooth decomposition, matrix differential equations

AMS subject classifications. 65F30, 15A23

DOI. 10.1137/050639703

1. Introduction. Low-rank approximation of unbearably large system matrices is a basic model reduction technique in many application areas, such as image compression, linear dynamical systems, regularization methods for ill-posed problems, and latent semantic indexing in information retrieval. In the present paper, we consider the task of computing low-rank approximations to matrices $A(t) \in \mathbb{R}^{m\times n}$ depending smoothly on a real parameter, henceforth referred to as time t. At any time t, a best approximation to $A(t)$ of rank r is a matrix $X(t)$ in the manifold $\mathcal{M}_r = \mathcal{M}_r^{m\times n}$ of rank-r matrices that satisfies

$$X(t) \in \mathcal{M}_r \quad \text{such that} \quad \|X(t) - A(t)\| = \min! \tag{1.1}$$

This is formulated for a matrix norm, which we choose as the Frobenius norm in the following. The problem is solved by a singular value decomposition (SVD) of $A(t)$, truncating all singular values after the r largest ones. When the matrix is so large that a complete SVD is not feasible, a standard approach to obtaining an approximate solution is based on the Lanczos bidiagonalization process with $A(t)$ [15].

Here, we consider instead the low-rank approximation $Y(t) \in \mathcal{M}_r$ determined from the condition that for every t the derivative $\dot{Y}(t)$, which is in the tangent space $\mathcal{T}_{Y(t)}\mathcal{M}_r$, be chosen as

$$\dot{Y}(t) \in \mathcal{T}_{Y(t)}\mathcal{M}_r \quad \text{such that} \quad \|\dot{Y}(t) - \dot{A}(t)\| = \min! \tag{1.2}$$

This is complemented with an initial condition, ideally $Y(t_0) = X(t_0)$. For given $Y(t)$, the derivative $\dot{Y}(t)$ is obtained by a *linear* projection, though onto a solution-dependent vector space. Problem (1.2) yields an initial value problem of nonlinear ordinary differential equations on $\mathcal{M}_r$, which becomes numerically efficiently accessible after choosing a suitable factorization of rank-r matrices.

*Received by the editors September 7, 2005; accepted for publication (in revised form) by H. Park September 29, 2006; published electronically April 3, 2007. This work was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 382.

http://www.siam.org/journals/simax/29-2/63970.html

†Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D–72076 Tübingen, Germany (othmar@othmar-koch.org, lubich@na.uni-tuebingen.de).

There are several independent reasons that make the approach to low-rank approximation via (1.2) attractive:

(a) Problem (1.2) and its solution algorithm depend on the increments $\dot{A}(t)$ instead of the complete data matrix $A(t)$. This appears to be an essential benefit in processes where $\dot{A}(t)$ is much *sparser* than $A(t)$, e.g., in series of moving images or in time-varying term-document matrices in information retrieval (updates are usually small compared to the whole encyclopedia).

(b) Solving the differential equations corresponding to (1.2) requires only multiplications of $\dot{A}(t)$ with matrices having few (r) columns, but *no decompositions* of matrices of the size of A, except for the low-rank approximation to the initial data $A(t_0)$.

(c) The differential equations for $Y(t)$ yield a *smooth* low-rank approximation. This is not assured when computing a pointwise best approximation, which is not unique in general.

(d) Since the problem (1.2) of determining $\dot{Y}(t)$ for given $Y(t)$ is linear, the approach extends more easily than (1.1) to *structured* low-rank approximation, where $\mathcal{M}_r$ is replaced by some submanifold.

(e) In contrast to (1.1), the approach (1.2) extends to the situation where $A(t)$ is not a given matrix but the unknown solution of a *master equation*, a matrix differential equation $\dot{A} = F(A)$. In this case, $\dot{A}(t)$ in (1.2) is simply replaced by the approximation $F(Y(t))$, so that the defect in the differential equation is minimized:

$$\dot{Y}(t) \in \mathcal{T}_{Y(t)}\mathcal{M}_r \quad \text{such that} \quad \|\dot{Y}(t) - F(Y((t))\| = \min! \tag{1.3}$$

Some comments and references to these aspects are in order: (a) and (b) are related to updating problems for low-rank approximations [3, 17], and item (c) to smooth decompositions of matrices, in particular to smooth SVD and the corresponding differential equations [1, 4, 7, 13, 16]. Item (d) refers to structured low-rank approximation as considered in [5, 6] for time-independent matrices. Item (e) and its generalization to low-rank approximation of tensors have a surprisingly long history in quantum mechanics: in 1930, Dirac [8] proposed to approximate the solution of the time-dependent Schrödinger equation, the multivariate wave function $\psi(x_1, \dots, x_d, t)$, by a rank-1 approximation, namely an (antisymmetrized) tensor product $\phi_1(x_1, t) \dots \phi_d(x_d, t)$, and derived differential equations for the functions ϕ_k from a variational principle analogous to (1.3), which is now known as the *Dirac–Frenkel time-dependent variational principle* in the chemical physics literature; see the historical references [8, 9] and, e.g., [2, 12]. Since the 1990s, the numerical approach of approximating the wave function by linear combinations of tensor products obeying differential equations derived from the Dirac–Frenkel principle (the multiconfiguration time-dependent Hartree or MCTDH method) has been used with great success for computations in quantum molecular dynamics [2]. It was, in fact, our work on variational approximations in quantum dynamics that led us to consider the dynamical low-rank matrix approximation (1.2), which does not appear to have been used or studied previously.

In the present paper we formulate the differential equations determining the solution of (1.2) and study the approximation properties of this approach, comparing the deviation from the best approximation, $Y(t) - X(t)$, with the best-approximation error $X(t) - A(t)$.

In section 2, we describe decompositions of rank-r matrices and their tangent matrices, and we derive differential equations for the factors that define the rank-r

approximation $Y(t)$. These differential equations are used for the numerical solution of the problem. In section 3 we illustrate the approach and the behavior of the dynamical low-rank approximation (1.2) by numerical experiments.

The analysis of the approximation properties of (1.2) turns out to be more demanding than the formal similarity of (1.1) and (1.2) would suggest. In section 4 we give a preparatory result on orthogonal projections onto tangent spaces of $\mathcal{M}_r$. The approximation properties of (1.2) are then studied in section 5 under the assumption that $A(t)$ is a perturbation to a matrix of rank $\leq r$. We first give near-optimality results when the effective rank of $A(t)$ is equal to r (Theorems 5.1 and 5.2), and then extend the result to the case where r in (1.2) is larger than the effective rank (Theorem 5.3). A further approximation result concerns systems without gaps in the distribution of the singular values (Theorem 5.5). Before turning to these approximation results, however, it should be noted that $Y(t)$ cannot always be expected to remain close to $X(t)$. This is already seen from the example of finding a rank-1 approximation to $\mathrm{diag}(e^{-t}, e^t)$, where starting from $t_0 < 0$ yields $X(t) = Y(t) = \mathrm{diag}(e^{-t}, 0)$ for $t < 0$, but $Y(t) = \mathrm{diag}(e^{-t}, 0)$ and $X(t) = \mathrm{diag}(0, e^t)$ for $t > 0$. The best approximation $X(t)$ here has a discontinuity at $t = 0$, caused by a crossing of singular values of which one is inside and the other outside the approximation. Our results show, however, that $Y(t)$ yields a near-optimal approximation on intervals where a good smooth approximation exists.

In section 6 we consider the following extensions of the basic approach:

- Regularization: the inverses of ill-conditioned matrices in the differential equations are replaced by regularized inverses.
- Stabilization: the differential equations are stabilized in order to drive the dynamical approximation toward the best approximation.
- Structured low-rank approximation: as an example we consider the problem of approximation by rank-r orthogonal projections.
- Matrix differential equations: we extend the method and the approximation results to the low-rank approximation (1.3) to solutions of matrix differential equations $\dot{A} = F(A)$.

The present paper deals with theoretical aspects of the dynamical low-rank approximation. Our very promising first experiences in using this technique for applications ranging from the compression of time-varying term-document matrices and of series of images to the computation of blow-up in reaction-diffusion equations are reported in [14].

Notation. Throughout the paper, $\|\cdot\|$ is the Frobenius norm,

$$\|A\| = \left(\sum_{i,j} a_{ij}^2 \right)^{1/2},$$

and $\langle \cdot, \cdot \rangle$ denotes the corresponding inner product, $\langle A, B \rangle = \mathrm{tr}\,(A^T B) = \sum_{i,j} a_{ij} b_{ij}$. We make frequent use of the inequality $\|AB\| \leq \|A\|_2 \cdot \|B\|$ and occasionally of $\|A\|_2 \leq \|A\|$, where $\|\cdot\|_2$ is the spectral norm.

2. Differential equations for low-rank approximation.

2.1. Decompositions of rank-r matrices and of their tangent matrices.

Every real rank-r matrix of dimension $m \times n$ can be written in the form

$$Y = USV^T, \tag{2.1}$$

where $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ have orthonormal columns, i.e.,

$$U^T U = I_r, \quad V^T V = I_r \tag{2.2}$$

(with the identity matrix I_r of dimension r), and $S \in \mathbb{R}^{r\times r}$ is nonsingular. The SVD yields S diagonal, but here we will not assume a special form of S. The representation (2.1) is not unique: replacing U by $\widetilde{U} = UP$ and V by $\widetilde{V} = VQ$ with orthogonal matrices $P, Q \in \mathbb{R}^{r\times r}$, and correspondingly S by $\widetilde{S} = P^T S Q$, yields the same matrix $Y = USV^T = \widetilde{U}\widetilde{S}\widetilde{V}^T$.

As a substitute for the nonuniqueness in the decomposition (2.1), we will use a *unique decomposition in the tangent space*. Let $\mathcal{V}_{m,r}$ denote the Stiefel manifold of real $m \times r$ matrices with orthonormal columns. The tangent space at $U \in \mathcal{V}_{m,r}$ is

$$\mathcal{T}_U\mathcal{V}_{m,r} = \{\delta U \in \mathbb{R}^{m\times r} : \ \delta U^T U + U^T \delta U = 0\} = \{\delta U \in \mathbb{R}^{m\times r} : \ U^T \delta U \in \mathrm{so}(r)\},$$

where $\mathrm{so}(r)$ denotes the space of skew-symmetric real $r \times r$ matrices. Consider the extended tangent map of $(S, U, V) \mapsto Y = USV^T$,

$$\begin{aligned} \mathbb{R}^{r\times r} \times \mathcal{T}_U\mathcal{V}_{m,r} \times \mathcal{T}_V\mathcal{V}_{n,r} \ &\to \ \mathcal{T}_Y\mathcal{M}_r \times \mathrm{so}(r) \times \mathrm{so}(r), \\ (\delta S, \delta U, \delta V) \ &\mapsto \ (\delta U S V^T + U \delta S V^T + U S \delta V^T, U^T \delta U, V^T \delta V). \end{aligned}$$

This linear map is an isomorphism, since it is readily seen to have zero null-space, and since the dimensions of the vector spaces on both sides agree.

Hence, every tangent matrix $\delta Y \in \mathcal{T}_Y\mathcal{M}_r$ is of the form

$$\delta Y = \delta U S V^T + U \delta S V^T + U S \delta V^T, \tag{2.3}$$

where $\delta S \in \mathbb{R}^{r\times r}$, and $\delta U \in \mathcal{T}_U\mathcal{V}_{m,r}$ and $\delta V \in \mathcal{T}_V\mathcal{V}_{n,r}$. Moreover, $\delta S, \delta U, \delta V$ are uniquely determined by δY if we impose the orthogonality constraints

$$U^T \delta U = 0, \quad V^T \delta V = 0. \tag{2.4}$$

With the identity matrices I_m, I_n of dimensions m and n, respectively, we define by

$$P_U = UU^T, \quad P_V = VV^T, \quad P_U^\perp = I_m - P_U, \quad P_V^\perp = I_n - P_V \tag{2.5}$$

the orthogonal projections onto the spaces spanned by the columns of U and V, and onto their orthogonal complements, respectively. Now, (2.3) and (2.4) yield

$$\begin{aligned} \delta S &= U^T \delta Y V, \\ \delta U &= P_U^\perp \delta Y V S^{-1}, \\ \delta V &= P_V^\perp \delta Y^T U S^{-T}. \end{aligned} \tag{2.6}$$

Formulas (2.3) and (2.6) establish an isomorphism between the subspace

$$\{(\delta S, \delta U, \delta V) \in \mathbb{R}^{r\times r} \times \mathbb{R}^{m\times r} \times \mathbb{R}^{n\times r} : U^T \delta U = 0,\ V^T \delta V = 0\}$$

and the tangent space $\mathcal{T}_Y\mathcal{M}_r$.

2.2. The differential equations for the factors. The minimization condition (1.2) on the tangent space is equivalent to an orthogonal projection: find $\dot Y \in \mathcal{T}_Y\mathcal{M}_r$ (we omit the argument t) satisfying

$$\langle \dot Y - \dot A,\, \delta Y\rangle = 0 \quad \text{for all} \quad \delta Y \in \mathcal{T}_Y\mathcal{M}_r. \tag{2.7}$$

From the viewpoint of numerical analysis, this is a Galerkin condition on the tangent space $\mathcal{T}_Y\mathcal{M}_r$. With this formulation we derive differential equations for the factors in the representation (2.1).

PROPOSITION 2.1. *For $Y = USV^T \in \mathcal{M}_r$ with nonsingular $S \in \mathbb{R}^{r\times r}$ and with $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ having orthonormal columns, condition* (1.2) *or* (2.7) *is equivalent to $\dot Y = \dot U SV^T + U\dot S V^T + US\dot V^T$, where*

$$\begin{aligned} \dot S &= U^T \dot A V, \\ \dot U &= P_U^\perp \dot A V S^{-1}, \\ \dot V &= P_V^\perp \dot A^T U S^{-T}, \end{aligned} \tag{2.8}$$

with the orthogonal projections $P_U^\perp = I_m - UU^T$ and $P_V^\perp = I_n - VV^T$.

Proof. For $u \in \mathbb{R}^m,\ v \in \mathbb{R}^n$, and $B \in \mathbb{R}^{m\times n}$, we use the identity

$$\langle uv^T, B\rangle = u^T B v.$$

In view of (2.4) we require $U^T\dot U = V^T\dot V = 0$ along the solution trajectory in order to define a unique representation of $\dot Y$. We first substitute $\delta Y = u_i v_j^T$, for $i,\ j = 1,\dots,r$, into (2.7), where u_i, v_j denote the columns of $U,\ V$, respectively. This is of the form (2.3) with $\delta U = \delta V = 0$ and one nonzero element in δS. In this way we find $\dot S = U^T\dot A V$. Similarly, choosing $\delta Y = \sum_{j=1}^r \delta u\, s_{ij} v_j^T,\ i = 1,\dots,r$, where $\delta u \in \mathbb{R}^m$ is arbitrary with $U^T\delta u = 0$, we obtain the stated differential equation for U, and likewise for $\delta Y = \sum_{j=1}^r u_j s_{ji} \delta v^T$ with $V^T\delta v = 0$ the differential equation for V. ∎

Note that with $\Lambda = U^T\dot A V$, the differential equations can be rewritten as

$$\begin{aligned} \dot S &= \Lambda, \\ \dot U S &= \dot A V - U\Lambda, \\ \dot V S^T &= \dot A^T U - V\Lambda^T. \end{aligned} \tag{2.9}$$

The matrices U and V retain orthonormal columns when the initial values have this property: since $U^T\dot U = 0$, we have $\frac{d}{dt}U^TU = \dot U^TU + U^T\dot U = 0$, and similarly for V.

The differential equations (2.8) are related to differential equations for other smooth matrix decompositions, in particular the smooth SVD; see [7, 16]. Unlike the differential equations for singular values given there, no singularities appear in (2.8) at points where singular values of $Y(t)$ coalesce. Equations (2.8) are in close relationship with the MCTDH equations [2], specialized to matrices instead of multivariate functions and stripped of the Schrödinger equation context.

In the numerical integration of (2.8), the step size control should be based on the local error in the low-rank approximation $Y = USV^T$, not on the local error in its factors (this makes a marked difference when S has small singular values). The orthogonality of the columns of U and V can be preserved in the numerical integration by the methods described, e.g., in [10, Chapter IV].

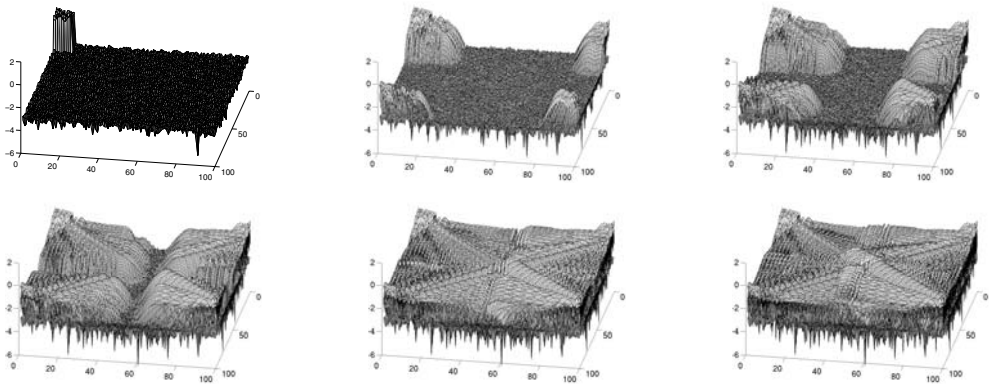

FIG. 3.1. *Size of the matrix elements for* $t = 0, 0.2, \ldots, 1$*, first example,* $\varepsilon = 1e-3$.

3. Numerical experiments. In this section we illustrate the behavior of the dynamical low-rank approximation method by three numerical examples. In all experiments, we have chosen the step sizes in the numerical integration of the differential equations (2.8) small enough that the error of the numerical integration is negligible as compared with the error of the low-rank approximation.

First example. We consider a model problem which was constructed in the following way: first, a 10×10 matrix of random numbers between 0 and 0.5 was added to the unit matrix of the same size, giving a matrix with singular values of magnitude ≈ 1. Subsequently, this matrix was added as the leading 10×10 block to a 100×100 matrix with random entries between 0 and 1 multiplied by a perturbation parameter ε, yielding a matrix A_1. Another matrix A_2 built in the same way, multiplied by $\exp(t)$, was added for $t \in [0, 1]$. Finally, to eliminate the possibility that this particular structure of the matrix might have an influence on our results, we applied a time-dependent transformation by orthogonal matrices (which does not alter the singular values, but the left and right singular vectors), which were created by solving initial value problems $\dot{Q}_i = T_i Q_i$ $(i = 1, 2)$ with skew-symmetric T_i and initial values equal to identity. To illustrate the structure of the resulting matrices,

$$A(t) = Q_1(t)(A_1 + e^t A_2)Q_2(t)^T,$$

we show, in Figure 3.1, the size of the matrix entries for values $t = 0, 0.2, 0.4, 0.6, 0.8, 1$, with the perturbation parameter $\varepsilon = 1e-3$. At $t = 0$ the large entries are located in a corner according to the construction of the test example, and afterwards the orthogonal transformations spread these large matrix elements such that the size of matrix entries is approximately the same all over.

Figure 3.2 shows the time evolution of the errors $\|Y - X\|$, $\|Y - A\|$ and the best-approximation error $\|X - A\|$ over the interval $[0, 1]$ for $\varepsilon = 1e-3$. When we vary the order of magnitude of the perturbation ε, we observe that the size of the error of the approximation Y defined by (1.2) as compared with the best approximation X from (1.1) is proportional to the error of the best approximation. In Table 3.1, the results are given for $\varepsilon = 1e-1, \ldots, 1e-5$ at time $t = 1$, where the maximum errors on the interval $[0, 1]$ occur. When the parameter ε is decreased by an order of magnitude, $\|X - A\|$ decreases proportionally, and $\|Y - X\|$ and $\|Y - A\|$ show the same behavior. We show the errors in the Frobenius norm, and additionally the norm of S^{-1}. We observe that for an approximation of rank $r = 10$, $\|S^{-1}\|$ does not increase significantly when ε decreases.

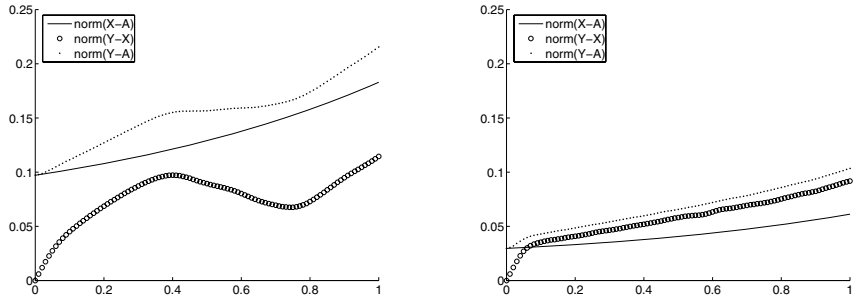

FIG. 3.2. *Errors as a function of t, for $r = 10$ (left) and $r = 20$ (right).*

TABLE 3.1
First example, $r = 10$.

ε	$\|X - A\|$	$\|Y - X\|$	$\|Y - A\|$	$\|S^{-1}\|$
1e−1	7.3762e+00	1.1808e+01	1.3478e+01	7.9878e−01
1e−2	9.3381e−01	5.1817e+00	5.2203e+00	1.4487e+00
1e−3	1.8293e−01	1.1450e−01	2.1549e−01	2.6232e+00
1e−4	1.8310e−02	1.1368e−02	2.1550e−02	2.6232e+00
1e−5	1.8312e−03	1.7596e−03	2.5395e−03	2.6230e+00

TABLE 3.2
First example, $r = 20$.

ε	$\|X - A\|$	$\|Y - X\|$	$\|Y - A\|$	$\|S^{-1}\|$
1e−1	6.0335e+00	1.2500e+01	1.3094e+01	1.5749e+00
1e−2	6.1246e−01	9.7993e−01	1.0885e+00	1.3569e+01
1e−3	6.1280e−02	9.1726e−02	1.0354e−01	1.3474e+02
1e−4	6.1282e−03	9.0416e−03	1.0298e−02	1.2940e+03

We repeat the same experiments with $r = 20$, which is larger than the effective rank 10 of the matrices that are to be approximated; see Table 3.2. We observe that $\|S^{-1}\|$ grows rapidly with ε in this case. However, the approximation quality is not negatively affected. Rather, the approximation error is smaller than in the case $r = 10$, especially for larger ε, since more singular values are taken into account. The approximation of the dominant singular values and vectors does not suffer from the bad overall conditioning introduced by the small, insignificant contributions. For small ε, a similar error behavior as in the case $r = 10$ is observed. We have not included the values for still smaller ε, because there the numerical integrator is forced to take very small step sizes. In this way the large norm of S^{-1} does have an influence on the numerical solution. This unfavorable effect can be avoided by using the regularization of section 6.

Second example. This example demonstrates a scenario which may cause a failure of the dynamical low-rank approximation because of a discontinuous best approximation. If the approximation rank r is chosen too small, then singular values which are initially small may in the course of time become larger than the singular values that are actually approximated by the algorithm. Thus, a rather large error compared to the (then discontinuous) best approximation may result when r is too small and the parameter t varies in an unfavorably large interval without a restart of the algorithm by recomputing a best approximation. Figure 3.3 shows such a situation for an exam-

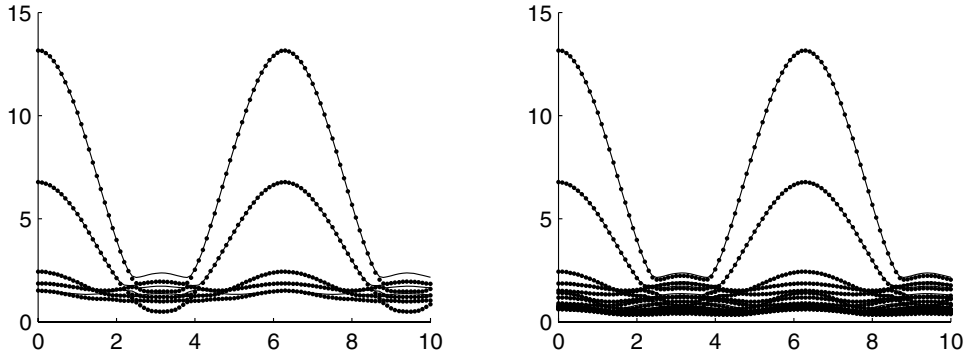

FIG. 3.3. *Second example, singular values for* $r = 5$ *and* $r = 20$.

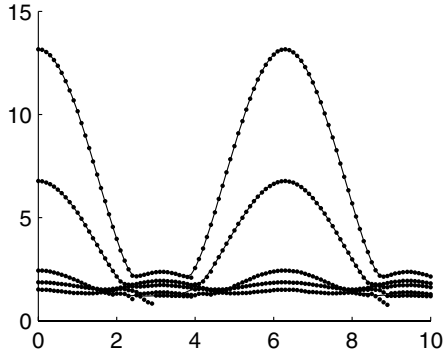

FIG. 3.4. *Second example, singular values for* $r = 5$*, algorithm restarted at* 20 *points.*

ple constructed by the same principles as for the example discussed above, yet with a time-dependence of the form $\cos(t)$ for $t \in [0, 10]$ and with $\varepsilon = 1e-1$. In both figures, we compare the r largest singular values of $A(t)$ (computed by SVD in every point considered) with the singular values of the rank-r approximation matrix $Y(t)$. In all the figures, the largest singular values computed for the exact matrix are given by a solid line, while those of the dynamical low-rank approximation (at equidistant output points) are represented by dots. If we choose $r = 5$, then the algorithm does not approximate the r largest singular values for all t. Rather, for $t \approx 3$, one of the singular values not included in the approximation becomes largest; see Figure 3.3(left). However, if we choose $r = 20$, then all the dominant singular values and vectors are included in the approximation; see Figure 3.3(right). The correct behavior is captured with $r = 5$ if the algorithm is restarted 20 times in the interval (Figure 3.4).

Third example. This example is again constructed similarly to the first example, $A(t) = Q_1(t)e^t DQ_2(t)^T \in \mathbb{R}^{100\times 100}$, where D is a diagonal matrix with entries $2^{-i/10}$, $i = 1, \ldots, 100$, in descending order. In contrast to the first example, however, there is no distinguishable gap in the set of singular values of $A(t)$. Nonetheless, the dynamical low-rank approximation yields satisfactory results. In Figure 3.5, the errors at $t = 1$ are given for a sequence of approximations of increasing rank r, where $r = 3(3)99$. We observe that $\|Y - X\|$ tends to zero at the same rate as $\|X - A\|$.

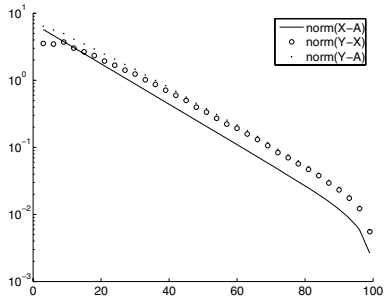

FIG. 3.5. *Third example, errors at $t = 1$ as a function of the rank r of the approximation.*

4. Tangent space projection and curvature bounds. In the following two sections we give an analysis that explains the error behavior observed in the numerical experiments. We begin with some preparation.

Condition (2.7) can be written as the differential equation on $\mathcal{M}_r$,

$$\dot{Y} = P(Y)\dot{A}, \tag{4.1}$$

where $P(Y)$ is the orthogonal projection onto the tangent space $\mathcal{T}_Y\mathcal{M}_r$. Basic properties of this projection are formulated in the following two lemmas.

LEMMA 4.1. *The orthogonal projection onto the tangent space $\mathcal{T}_Y\mathcal{M}_r$ at $Y = USV^T \in \mathcal{M}_r$ is given by*

$$P(Y) = I - P^\perp(Y) \qquad \text{with} \qquad P^\perp(Y)B = P_U^\perp B P_V^\perp \tag{4.2}$$

for $B \in \mathbb{R}^{m\times n}$.

Proof. Proposition 2.1 yields $\dot{Y} = \dot{A} - P_U^\perp \dot{A} P_V^\perp$ for $\dot{Y}$ of (2.7) or equivalently of (4.1). Since this holds for every matrix $\dot{A}$, the result follows. ∎

LEMMA 4.2. *Let the rank-r matrix $X \in \mathcal{M}_r$ be such that its smallest nonzero singular value satisfies $\sigma_r(X) \geq \rho > 0$, and let $Y \in \mathcal{M}_r$ with $\|Y - X\| \leq \frac{1}{8}\rho$. Then, the following bounds hold: for all $B \in \mathbb{R}^{m\times n}$,*

$$\|\big(P(Y) - P(X)\big)B\| \leq 8\rho^{-1}\,\|Y - X\| \cdot \|B\|_2, \tag{4.3}$$

$$\|P^\perp(Y)(Y - X)\| \leq 4\rho^{-1}\,\|Y - X\|^2. \tag{4.4}$$

Proof. (a) For $X = U_0 S_0 V_0^T \in \mathcal{M}_r$ we have the bound $\|S_0^{-1}\|_2 \leq \rho^{-1}$. Since we have, by [11, p. 448],

$$|\sigma_r(Y) - \sigma_r(X)| \leq \|Y - X\|_2 \leq \|Y - X\|,$$

we obtain for $\|Y - X\| \leq \frac{1}{8}\rho$ that

$$\sigma_r(Y) \geq \sigma_r(X) - |\sigma_r(Y) - \sigma_r(X)| \geq \tfrac{7}{8}\rho,$$

and hence $Y = U_1 S_1 V_1^T$ with $\|S_1^{-1}\|_2 \leq \frac{8}{7}\rho^{-1}$.

(b) We decompose the matrices on the straight line connecting X and Y as

$$X + \tau(Y - X) = M(\tau) + N(\tau) \qquad \text{with} \qquad M(\tau) \in \mathcal{M}_r,\ N(\tau) \perp \mathcal{T}_X\mathcal{M}_r.$$

A smooth such decomposition exists at least for small τ, but the arguments below show that it exists in fact for $0 \leq \tau \leq 1$. We denote

$$\Delta = P(X)(Y - X) \in \mathcal{T}_X\mathcal{M}_r, \qquad \text{with } \|\Delta\| \leq \delta := \|Y - X\|.$$

We then have $P(X)(M(\tau) - X) = \tau\Delta$, which yields

$$P(X)\dot{M}(\tau) = \Delta.$$

Since (4.2) gives $P(X)\dot{M} = \dot{M} - P_{U_0}^{\perp}\dot{M}P_{V_0}^{\perp}$, we obtain $P_{U_0}\dot{M} = P_{U_0}\Delta$ and $\dot{M}P_{V_0} = \Delta P_{V_0}$, which implies

$$U_0^T\dot{M}(\tau) = U_0^T\Delta, \quad \dot{M}(\tau)V_0 = \Delta V_0. \tag{4.5}$$

(c) Using Proposition 2.1 with $M(\tau) \in \mathcal{M}_r$ in the role of $A(\tau)$ and $Y(\tau)$, we get

$$M(\tau) = U(\tau)S(\tau)V(\tau)^T,$$

where S, U, V satisfy the differential equations

$$\begin{aligned} \dot{S} &= U^T\dot{M}V = U^T\Delta V + (U - U_0)^T P_{U_0}^{\perp}\dot{M}P_{V_0}^{\perp}(V - V_0), \\ \dot{U} &= P_U^{\perp}\dot{M}VS^{-1} = P_U^{\perp}\Delta V_0 S^{-1} + P_U^{\perp}\dot{M}(V - V_0)S^{-1}, \\ \dot{V} &= P_V^{\perp}\dot{M}^T US^{-T} = P_V^{\perp}\Delta^T U_0 S^{-T} + P_V^{\perp}\dot{M}^T(U - U_0)S^{-T}. \end{aligned} \tag{4.6}$$

In the second equalities we have used $\dot{M} = \Delta + P_{U_0}^{\perp}\dot{M}P_{V_0}^{\perp}$ and (4.5), and the fact that $P_{U_0}^{\perp}U_0 = 0$ and $P_{V_0}^{\perp}V_0 = 0$. In addition, we have $\dot{M} = U\dot{S}V^T + (\dot{U}S)V^T + U(\dot{V}S^T)^T$, and hence

$$\dot{M} = \Delta + P_{U_0}^{\perp}\big((U - U_0)\dot{S}(V - V_0)^T + (\dot{U}S)(V - V_0)^T + (U - U_0)(\dot{V}S^T)^T\big)P_{V_0}^{\perp}. \tag{4.7}$$

We will show that these differential equations have a solution up to $\tau = 1$. As long as $\|U - U_0\| \le \frac{1}{4}$ and $\|V - V_0\| \le \frac{1}{4}$, they give the bounds

$$\|\dot{S}\| \le \delta + \tfrac{1}{16}\|\dot{M}\|, \quad \|\dot{U}S\| \le \delta + \tfrac{1}{4}\|\dot{M}\|, \quad \|\dot{V}S^T\| \le \delta + \tfrac{1}{4}\|\dot{M}\|,$$

which inserted into the equation for $\dot{M}$ yield

$$\|\dot{M}\| \le 2\delta \quad \text{and} \quad \|\dot{S}\| \le \tfrac{9}{8}\delta, \quad \|\dot{U}S\| \le \tfrac{3}{2}\delta, \quad \|\dot{V}S^T\| \le \tfrac{3}{2}\delta. \tag{4.8}$$

The bound for $\dot{S}$ yields $\|S(\tau) - S_0\| \le \frac{9}{8}\delta$ for $\tau \le 1$. By the assumption $\delta \le \frac{1}{8}\rho$ and the argument in part (a), we thus obtain $\|S(\tau)^{-1}\|_2 \le \frac{4}{3}\rho^{-1}$ for $\tau \le 1$. From the bound for $\dot{U}S$ we then have

$$\|\dot{U}\| \le \|\dot{U}S\| \cdot \|S^{-1}\|_2 \le \tfrac{3}{2}\delta \cdot \tfrac{4}{3}\rho^{-1} = 2\rho^{-1}\delta \le \tfrac{1}{4}. \tag{4.9}$$

The same bound holds for $\dot{V}$. These bounds show that the differential equation has a solution on the whole interval $0 \le \tau \le 1$ with

$$\|S_1 - S_0\| \le \tfrac{9}{8}\,\delta, \quad \|U_1 - U_0\| \le 2\rho^{-1}\delta, \quad \|V_1 - V_0\| \le 2\rho^{-1}\delta. \tag{4.10}$$

(d) The above bounds give immediately

$$\|P_{U_1}^{\perp} - P_{U_0}^{\perp}\| \le 4\rho^{-1}\delta, \quad \|P_{V_1}^{\perp} - P_{V_0}^{\perp}\| \le 4\rho^{-1}\delta. \tag{4.11}$$

Formula (4.2) shows

$$\big(P(Y) - P(X)\big)B = P_{U_0}^{\perp}BP_{V_0}^{\perp} - P_{U_1}^{\perp}BP_{V_1}^{\perp},$$

which together with (4.11) yields the bound (4.3).

(e) With $P_U^\perp U = 0$ and $P_V^\perp V = 0$ we obtain

$$\begin{aligned} P^\perp(Y)(Y-X) &= P_{U_1}^\perp (U_1 S_1 V_1^T - U_0 S_0 V_0^T) P_{V_1}^\perp = -P_{U_1}^\perp U_0 S_0 V_0^T P_{V_1}^\perp \\ &= -P_{U_1}^\perp (U_1 - U_0) S_0 (V_1 - V_0)^T P_{V_1}^\perp. \end{aligned} \tag{4.12}$$

We write

$$(U_1 - U_0)S_0 = \int_0^1 \dot U(\tau) S_0 \, d\tau = \int_0^1 \dot U(\tau) S(\tau)\, d\tau - \int_0^1 \dot U(\tau)(S(\tau) - S_0)\, d\tau,$$

and hence (4.8) and (4.9) yield

$$\|(U_1 - U_0)S_0\| \le 2\delta.$$

Using this bound and (4.10) in (4.12) finally gives the bound for $P^\perp(Y)(Y-X)$. ∎

5. Approximation properties. We give four results that illustrate different aspects of the dynamical low-rank approximation problem.

5.1. Local quasi optimality. If the low-rank approximation problem (1.1) has a continuously differentiable best approximation $X(t) \in \mathcal{M}_r$, then the error of (1.2) can be bounded in terms of the best-approximation error $\|X(t) - A(t)\|$. The result involves a bound on $\dot A(t)$:

$$\|\dot A(t)\|_2 \le \mu \qquad \text{for } \ 0 \le t \le \bar t. \tag{5.1}$$

(For convenience we choose the initial time $t_0 = 0$.)

THEOREM 5.1. *Suppose that a continuously differentiable best approximation $X(t) \in \mathcal{M}_r$ to $A(t)$ exists for $0 \le t \le \bar t$. Let the rth singular value of $X(t)$ have the lower bound $\sigma_r(X(t)) \ge \rho > 0$, and assume that the best-approximation error is bounded by $\|X(t) - A(t)\| \le \frac{1}{16}\rho$ for $0 \le t \le \bar t$. Then, the approximation error of (1.2) with initial value $Y(0) = X(0)$ is bounded in the Frobenius norm by*

$$\|Y(t) - X(t)\| \le 2\beta\, e^{\beta t} \int_0^t \|X(s) - A(s)\|\, ds \qquad \text{with } \ \beta = 8\mu\rho^{-1}$$

for $t \le \bar t$ and as long as the right-hand side is bounded by $\frac{1}{8}\rho$.

Proof. For the best approximation it must hold that $X - A$ is orthogonal to the tangent space $\mathcal{T}_X\mathcal{M}_r$, or equivalently,

$$P(X)(X - A) = 0.$$

We differentiate this relation with respect to t and denote $(P'(X)\cdot B)\dot X = \frac{d}{dt} P(X(t))B$ to obtain

$$P(X)(\dot X - \dot A) + \big(P'(X)\cdot(X - A)\big)\dot X = 0.$$

Since $\dot X \in \mathcal{T}_X\mathcal{M}_r$, we have $P(X)\dot X = \dot X$, and the equation becomes

$$\big(I - P'(X)\cdot(X - A)\big)\dot X = P(X)\dot A. \tag{5.2}$$

Lemma 4.2 and the condition $d := \|X - A\| \le \frac{1}{16}\rho$ yield

$$\|P'(X)\cdot(X - A)\| \le 8\rho^{-1} d \le \tfrac{1}{2},$$

and hence (5.2) can be solved for $\dot{X}$ to yield

$$\dot{X} = P(X)\dot{A} + D \qquad \text{with} \quad \|D\| \le 16\rho^{-1}d\mu = 2\beta d.$$

We subtract this equation from (4.1), that is, $\dot{Y} = P(Y)\dot{A}$, and integrate from 0 to t. As long as $e := \|Y - X\| \le \frac{1}{8}\rho$, Lemma 4.2 yields

$$\|\big(P(Y) - P(X)\big)\dot{A}\| \le 8\rho^{-1}e\mu = \beta e,$$

and hence we obtain

$$e(t) \le \beta \int_0^t e(s)\,ds + 2\beta \int_0^t d(s)\,ds.$$

The result now follows with the Gronwall inequality. ∎

5.2. A farther-reaching error bound. Smaller errors over longer time intervals are obtained if not only $X - A$ but also its derivative is small. We assume that $A(t)$ is of the form

$$A(t) = X(t) + E(t), \qquad 0 \le t \le \bar{t}, \tag{5.3}$$

where $X(t) \in \mathcal{M}_r$ (now this need not necessarily be the best approximation) with

$$\|\dot{X}(t)\|_2 \le \mu, \tag{5.4}$$

and the derivative of the remainder term is bounded by

$$\|\dot{E}(t)\| \le \varepsilon \tag{5.5}$$

with a small $\varepsilon > 0$. We assume $\varepsilon \le \frac{1}{8}\mu$.

THEOREM 5.2. *In addition to the above assumptions, suppose that the rth singular value of $X(t)$ satisfies $\sigma_r(X(t)) \ge \rho > 0$. Then, the approximation error of (1.2) with initial value $Y(0) = X(0)$ is bounded in the Frobenius norm by*

$$\|Y(t) - X(t)\| \le 2t\varepsilon \qquad \text{for} \quad t \le \min\left(\bar{t}, \frac{\rho}{4\sqrt{2\mu\varepsilon}}\right).$$

Proof. We note $\dot{X} = P(X)\dot{X}$, rewrite (4.1) as $\dot{Y} = P(Y)\dot{X} + P(Y)\dot{E}$, and subtract the two equations. We observe

$$\big(P(Y) - P(X)\big)\dot{X} = -\big(P^\perp(Y) - P^\perp(X)\big)\dot{X} = -P^\perp(Y)\dot{X} = -P^\perp(Y)^2\dot{X}.$$

We take the inner product with $Y - X$ to obtain

$$\begin{aligned}\langle Y - X, \big(P(Y) - P(X)\big)\dot{X}\rangle &= -\langle Y - X, P^\perp(Y)\dot{X}\rangle = -\langle P^\perp(Y)(Y - X), P^\perp(Y)\dot{X}\rangle \\ &= \langle P^\perp(Y)(Y - X), \big(P(Y) - P(X)\big)\dot{X}\rangle.\end{aligned}$$

With Lemma 4.2 and (5.4), (5.5) this yields

$$\begin{aligned}\langle Y - X, \dot{Y} - \dot{X}\rangle &= \langle P^\perp(Y)(Y - X), \big(P(Y) - P(X)\big)\dot{X}\rangle + \langle Y - X, P(Y)\dot{E}\rangle \\ &\le 32\,\mu\rho^{-2}\|Y - X\|^3 + \|Y - X\| \cdot \varepsilon,\end{aligned}$$

and, on the other hand, we have

$$\langle Y - X, \dot{Y} - \dot{X} \rangle = \frac{1}{2} \frac{d}{dt} \|Y - X\|^2 = \|Y - X\| \, \frac{d}{dt} \|Y - X\|.$$

Taken together, we obtain for $e(t) = \|Y(t) - X(t)\|$ the differential inequality

$$\dot{e} \le \gamma e^2 + \varepsilon, \qquad e(0) = 0,$$

with $\gamma = 32\,\mu\rho^{-2}$. Hence, $e(t)$ is majorized by the solution of

$$\dot{y} = \gamma y^2 + \varepsilon, \qquad y(0) = 0,$$

which equals $y(t) = \sqrt{\varepsilon/\gamma}\,\tan(t\sqrt{\gamma\varepsilon})$ and is bounded by $2t\varepsilon$ for $t\sqrt{\gamma\varepsilon} \le 1$. Lemma 4.2 remains applicable as long as $2t\varepsilon \le \frac{1}{8}\rho$, which is satisfied on the given interval under the assumption $\varepsilon \le \frac{1}{8}\mu$. ∎

5.3. The case of overapproximation. The time interval in Theorem 5.2 becomes tiny when $\rho \le \varepsilon$. In that case, the effective rank (ε-pseudorank) of $A(t)$ is $q < r$, but the approximation is done by a rank-r matrix $Y(t)$. It is not clear a priori that $Y(t)$ preserves an effective rank q over longer times. Even if it does, the matrix $S(t)$ in (2.1) is ill-conditioned, and since its inverse appears in the differential equations (2.8), one might expect a severe adverse effect on the approximation properties. Remarkably, this does not happen, as is shown by the following result.

THEOREM 5.3. *Let* (5.3)–(5.5) *hold for* $X(t) \in \mathcal{M}_q$ *with* $q < r$. *Suppose that the* q*th singular value of* $X(t)$ *satisfies* $\sigma_q(X(t)) \ge \rho > 0$. *Let the initial value* $Y(0) \in \mathcal{M}_r$ *be* $Y(0) = X(0) + E_0$ *with* $\operatorname{Im} E_0 \perp \operatorname{Im} X(0)$ *and* $\operatorname{Im} E_0^T \perp \operatorname{Im} X(0)^T$, *and with* $\|E_0\| \le \varepsilon_0 \le \frac{1}{16}\rho\mu^{-1}\varepsilon$. *Suppose that the differential equation* (2.8) *has a solution on the interval* $0 \le t \le t^*$. *Then, the approximation error of* (1.2) *is bounded in the Frobenius norm by*

$$\|Y(t) - X(t)\| \le \varepsilon_0 + 6t\varepsilon \qquad \text{for} \quad t \le \min\left(\bar{t}, t^*, \frac{\rho}{16\mu}\right).$$

The existence of the solution of the differential equation (2.8) is not ensured over the whole interval, since $S(t)$ might become singular. The orthogonality condition on E_0 is satisfied if the best rank-r approximation is taken as initial value. However, this orthogonality condition is not essential. A similar, but less clear-cut estimate holds whenever $Y(0) \in \mathcal{M}_r$ is sufficiently close to $X(0) \in \mathcal{M}_q$.

The proof of Theorem 5.3 is based on combining the previous proof with the following two-scale lemma.

LEMMA 5.4. *Let* $Y \in \mathcal{M}_r$ *be written in the form*

$$Y = U \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix} V^T = U_1 S_1 V_1^T + U_2 S_2 V_2^T, \tag{5.6}$$

where $U = (U_1, U_2)$ *and* $V = (V_1, V_2)$ *have orthonormal columns. Assume that the smallest singular value of* $S_1 \in \mathbb{R}^{q\times q}$ *and the largest singular value of* $S_2 \in \mathbb{R}^{(r-q)\times(r-q)}$ *satisfy*

$$\sigma_{\min}(S_1) \ge \rho, \quad \sigma_{\max}(S_2) \le \delta \qquad \text{with } \rho \ge 2\delta. \tag{5.7}$$

Then, the solution of the orthogonal projection (2.7) *is given as*

$$\dot Y = \sum_{i=1}^{2} \Big(\dot U_i S_i V_i^T + U_i \dot S_i V_i^T + U_i S_i \dot V_i^T \Big), \tag{5.8}$$

where

$$\begin{aligned} &\dot S_1 = U_1^T \dot A V_1, \qquad \dot S_2 = U_2^T \dot A V_2, \\ &\dot U_1 S_1 = P_{U_1}^{\perp} \dot A V_1 + E_U, \\ &\dot V_1 S_1^T = P_{V_1}^{\perp} \dot A^T U_1 + E_V, \end{aligned} \tag{5.9}$$

with $\|E_U\| \le 2\delta\rho^{-1}\|\dot A\|$ *and* $\|E_V\| \le 2\delta\rho^{-1}\|\dot A\|$.

The point of the lemma is that for $\delta \ll \rho$ the equations for $\dot S_1, \dot U_1, \dot V_1$ are, up to the small perturbations E_U and E_V, the same as those for solving the corresponding rank-q problem for $Y_q = U_1 S_1 V_1^T \in \mathcal{M}_q$; see Proposition 2.1. Under conditions (5.3)–(5.5) with $X \in \mathcal{M}_q$, the equation for $\dot S_2$ has a small right-hand side as long as Y is close to A or X, and hence S_2 remains small. The term $U_2 S_2 V_2^T$ then gives only a small contribution to Y, no matter what the derivatives of U_2 and V_2 are. The equations for $\dot U_2$ and $\dot V_2$, which have not been stated explicitly, contain in fact S_2^{-1}, which may have an arbitrarily large norm.

Proof of Lemma 5.4. We begin by showing that $\dot S_i, \dot U_i, \dot V_i$ in (5.8) are uniquely determined if, instead of (2.4), we impose the constraints (cf. [7])

$$U^T \dot U = H, \qquad V^T \dot V = K,$$

with $r \times r$ matrices of the block form

$$H = \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix}, \qquad K = \begin{pmatrix} 0 & K_{12} \\ K_{21} & 0 \end{pmatrix},$$

which are skew-symmetric: $H_{12} = -H_{21}^T$ and $K_{12} = -K_{21}^T$. As in the proof of Proposition 2.1 we then obtain, instead of (2.8), the equation

$$U^T \dot A V = HS + \dot S + SK^T, \tag{5.10}$$

which yields $\dot S$ as the block diagonal of $\Lambda := U^T \dot A V$:

$$\dot S_1 = U_1^T \dot A V_1, \qquad \dot S_2 = U_2^T \dot A V_2.$$

We multiply (5.10) with S^{-1} from the right and take the symmetric part. Then H drops out, and we obtain

$$SK^T S^{-1} + S^{-T} K S^T = \begin{pmatrix} 0 & B_{12} S_2^{-1} \\ S_2^{-T} B_{21} & 0 \end{pmatrix},$$

where $B_{21} = B_{12}^T = \Lambda_{12}^T + S_2^T \Lambda_{21} S_1^{-1}$ is bounded by

$$\|B_{21}\| \le (1 + \delta\rho^{-1}) \, \|\dot A\| \le \tfrac{3}{2} \|\dot A\|.$$

We multiply the (2,1) block of the above equation with S_2^T from the left and with S_1^{-T} from the right to obtain

$$K_{21} - (S_2^T S_2) K_{21} (S_1^T S_1)^{-1} = B_{21} S_1^{-T}.$$

By condition (5.7), this equation can be uniquely solved for K_{21} by fixed-point iteration, and

$$\|K_{21}\| \le \tfrac{4}{3}\rho^{-1}\|B_{21}\| \le 2\rho^{-1}\|\dot A\|.$$

As in (2.8), we derive

$$\dot U S = \dot A V - U\dot S - USK^T.$$

For the first component of $\dot U = (\dot U_1, \dot U_2)$ this becomes

$$\dot U_1 S_1 = \dot A V_1 - U_1 \dot S_1 + E_U = P_{U_1}^\perp \dot A V_1 + E_U$$

with $E_U = U_2 S_2 K_{21}$, which is bounded by $\|E_U\| \le \delta \cdot 2\rho^{-1}\|\dot A\|$. The equation and estimate for $\dot V_1$ are obtained in the same way. □

In the proof of Theorem 5.3 we will actually use a variant of the above result, which is proved in the same way: if condition (5.7) is replaced by

$$\sigma_{\min}(S_1) \ge \rho, \qquad \|S_2\| \le \delta, \tag{5.11}$$

then Lemma 5.4 holds with the modified bounds

$$\|E_U\| \le 2\delta\rho^{-1}\|\dot A\|_2, \qquad \|E_V\| \le 2\delta\rho^{-1}\|\dot A\|_2. \tag{5.12}$$

Proof of Theorem 5.3. We write Y in the form (5.6) as

$$Y = Y_1 + Y_2 \equiv U_1 S_1 V_1^T + U_2 S_2 V_2^T.$$

We will estimate $e_1 = \|Y_1 - X\|$ and $e_2 = \|Y_2\|$.

(a) By Lemma 5.4 and (5.3), Y_1 satisfies the differential equation

$$\dot Y_1 = P_q(Y_1)\dot X + P_q(Y_1)\dot E + E_U V_1^T + U_1 E_V^T,$$

where $P_q(Y_1)$ denotes the orthogonal projection onto $\mathcal{T}_{Y_1}\mathcal{M}_q$. Comparing this equation with $\dot X = P_q(X)\dot X$ as in the proof of Theorem 5.2, we obtain

$$e_1(t) \le 2t\eta, \tag{5.13}$$

provided that, up to time t,

$$d := \|P_q(Y_1)\dot E + E_U V_1^T + U_1 E_V^T\| \le \eta.$$

By (5.5) and (5.12) together with (5.4) we have

$$d(t) \le \varepsilon + 5\rho^{-1}\mu\, e_2(t). \tag{5.14}$$

(The factor 5 instead of 4 takes into account that $\|\dot A\| \le \mu + \varepsilon$ may be slightly larger than μ of (5.4), and ρ in (5.11) may differ slightly from ρ in the formulation of the theorem. The bound holds as long as e_1 is sufficiently small.)

(b) We have

$$e_2(t) = \|S_2(t)\| \le \varepsilon_0 + \int_0^t \|\dot S_2(s)\|\, ds.$$

Now,

$$\begin{aligned}\|\dot S_2\| &= \|U_2^T \dot A V_2\| = \|P_{U_2} \dot A P_{V_2}\| \le \|P_{U_1}^\perp \dot A P_{V_1}^\perp\| \\ &= \|P_q^\perp(Y_1)\dot A\| \le \|P_q^\perp(Y_1)\dot X\| + \varepsilon = \|\big(P_q^\perp(Y_1) - P_q^\perp(X)\big)\dot X\| + \varepsilon \\ &= \|\big(P_q(Y_1) - P_q(X)\big)\dot X\| + \varepsilon \le 8\rho^{-1}\mu\, e_1 + \varepsilon,\end{aligned}$$

where we have used Lemma 4.2 and (5.4) in the last inequality. Hence,

$$e_2(t) \le \varepsilon_0 + 8\rho^{-1}\mu \int_0^t e_1(s)\, ds + t\varepsilon. \tag{5.15}$$

(c) With the bound (5.13) this inequality yields

$$e_2(t) \le \varepsilon_0 + 8\rho^{-1}\mu\, t^2\eta + t\varepsilon.$$

In view of (5.14), we thus need to choose η and the maximum value of t such that

$$\eta \ge \varepsilon + 5\rho^{-1}\mu\varepsilon_0 + 5(\rho^{-1}\mu t)\varepsilon + 40(\rho^{-1}\mu t)^2\eta.$$

For $\rho^{-1}\mu\varepsilon_0 \le \frac{1}{16}\varepsilon$ and $\rho^{-1}\mu t \le \frac{1}{16}$ this is satisfied for $\eta = 2\varepsilon$. This yields

$$e_1(t) \le 4t\varepsilon, \qquad e_2(t) \le \varepsilon_0 + 2t\varepsilon,$$

which implies the result. ∎

5.4. Systems without gaps between the singular values. The results of the preceding subsections give satisfactory error bounds when there is a gap in the distribution of the singular values so that essential and inessential singular values are widely separated. We now consider a situation where such a gap need not exist, as in the third numerical example. We make the assumptions of Theorem 5.2 and further that $X(t) \in \mathcal{M}_r$ with $\sigma_r(X(t)) \ge \rho > 0$ has a decomposition

$$X(t) = U_0(t)S_0(t)V_0(t)^T \qquad \text{for } \ 0 \le t \le \bar t, \tag{5.16}$$

with nonsingular $S_0(t) \in \mathbb{R}^{r\times r}$, and with $U_0(t) \in \mathbb{R}^{m\times r}$ and $V_0(t) \in \mathbb{R}^{n\times r}$ having orthogonal columns, such that the following bounds hold for $0 \le t \le \bar t$:

$$\Big\|\frac{d}{dt}S_0^{-1}(t)\Big\|_2 \le c_1\rho^{-1}, \qquad \|\dot U_0(t)\|_2 \le c_2, \quad \|\dot V_0(t)\|_2 \le c_2. \tag{5.17}$$

Under these conditions we can show an $O(\varepsilon)$ error over times $O(1)$ even with $\rho \sim \varepsilon$.

THEOREM 5.5. *Under the conditions of Theorem* 5.2 *and with* (5.16)–(5.17)*, the approximation error of* (1.2) *with initial value* $Y(0) = X(0)$ *is bounded by*

$$\|Y(t) - X(t)\| \le 2t\varepsilon \qquad \text{for} \quad t \le \min\Big(\bar t,\ \frac{1}{16c_2^{1/2}}\Big(\frac{\rho}{\varepsilon}\Big)^{1/2},\ \frac{1}{8c_1^{1/3}}\Big(\frac{\rho}{\varepsilon}\Big)^{2/3},\ \frac{1}{16}\frac{\rho}{\varepsilon}\Big).$$

Proof. From the proof of Theorem 5.2 we have the equation

$$\langle Y - X, \dot Y - \dot X\rangle = -\langle P^\perp(Y)(Y - X), P^\perp(Y)\dot X\rangle + \langle Y - X, P(Y)\dot E\rangle. \tag{5.18}$$

For $e = \|Y - X\| \le \frac{1}{8}\rho$, the proof of Lemma 4.2 shows that Y can be decomposed as $Y = U_1S_1V_1^T$ with

$$\|(U_1 - U_0)S_0\| \le 2e, \qquad \|S_0(V_1 - V_0)^T\| \le 2e.$$

By Lemma 4.1 we can write

$$\begin{aligned} P^\perp(Y)\dot X &= P^\perp_{U_1}\big(\dot U_0 S_0 V_0^T + U_0 \dot S_0 V_0^T + U_0 S_0 \dot V_0^T\big) P^\perp_{V_1} \\ &= P^\perp_{U_1}\big(-\dot U_0 S_0 (V_1 - V_0)^T + (U_1 - U_0) S_0 \cdot S_0^{-1} \dot S_0 S_0^{-1} \cdot S_0 (V_1 - V_0)^T \\ &\qquad - (U_1 - U_0) S_0 \dot V_0^T\big) P^\perp_{V_1}. \end{aligned}$$

With (4.4), (5.17), and the above estimate, (5.18) gives the differential inequality, as long as $e \le \frac{1}{8}\rho$:

$$\dot e \le 4\rho^{-1} e\big(2c_2 e + 4c_1 \rho^{-1} e^2 + 2c_2 e\big) + \varepsilon. \tag{5.19}$$

The error is bounded by $2t\varepsilon$ as long as the first term on the right-hand side is bounded by ε, which is thus satisfied for $16c_2\rho^{-1}(2t\varepsilon)^2 \le \frac{1}{2}\varepsilon$ and $16c_1\rho^{-2}(2t\varepsilon)^3 \le \frac{1}{2}\varepsilon$. This holds under the stated bounds for t. ∎

6. Extensions of the basic approach.

6.1. Regularization. Though Theorem 5.3 shows that overapproximation has no disastrous effect on the approximation properties, it *is* in fact harmful to the numerical solution of the differential equations (2.8). Near-singularity of S enforces very small step sizes in numerical integrators, and rounding errors may become important. The situation is alleviated by replacing S^{-1} by a *regularized inverse*, e.g., obtained from computing an SVD of $S \in \mathbb{R}^{r\times r}$ and replacing the ith singular value σ_i by $\sqrt{\sigma_i^2 + \epsilon^2}$. The approximation result of Theorem 5.3 remains valid (with modified constants), since transformation to the block form (5.8) yields only an $O(\epsilon)$ perturbation in (5.9).

6.2. Stabilization. In order to drive the solution toward the best approximation, we replace (2.7) by

$$\langle \dot Y - \dot A, \delta Y\rangle + \alpha\langle Y - A, \delta Y\rangle = 0 \qquad \text{for all} \quad \delta Y \in \mathcal{T}_Y\mathcal{M}_r \tag{6.1}$$

with a positive parameter α. This amounts to replacing $\dot A$ by $\dot A - \alpha(Y - A)$ in the differential equations (2.8) determining $Y = USV^T$. (This approach requires knowledge of both $\dot A$ and A and can therefore not be extended to the low-rank approximation of matrix differential equations as in section 6.4 below.)

The effect of the parameter α is easily seen in the framework of the proof of Theorem 5.1. With the notation used there, we have the differential equations

$$\begin{aligned} \dot Y &= P(Y)(\dot A - \alpha(Y - A)), \\ \dot X &= P(X)(\dot A - \alpha(X - A)) + D. \end{aligned}$$

Subtracting the equations yields

$$\begin{aligned} \dot Y - \dot X = \big(P(Y) - P(X)\big)\dot A - D \\ - \alpha(Y - X) + \alpha P^\perp(Y)(Y - X) - \alpha\big(P(Y) - P(X)\big)(X - A). \end{aligned}$$

Taking the inner product with $Y - X$ and using Lemma 4.2 yields the following differential inequality for $e = \|Y - X\|$: with $d = \|X - A\|$ and $\beta = 8\mu\rho^{-1}$,

$$\dot e \le \beta e + 2\beta d - \alpha e(1 - \tfrac{1}{2}\beta e - \beta d). \tag{6.2}$$

The last term is stabilizing, provided that d and e are small enough.

6.3. An example of structured low-rank approximation: Approximation on Grassmann manifolds. We now approximate $A(t) \in \mathbb{R}^{n\times n}$ not just by arbitrary rank-r matrices, but by orthogonal projections onto r-dimensional subspaces. We thus replace the manifold $\mathcal{M}_r^{n\times n}$ in (1.2) by the submanifold (known as a Grassmann manifold)

$$\mathcal{G} = \mathcal{G}_{n,r} = \{Y \in \mathcal{M}_r^{n\times n} : \ Y^2 = Y, \ Y^T = Y\}.$$

A projection $Y \in \mathcal{G}$ can be written, in a nonunique way, as

$$Y = UU^T \qquad \text{with} \quad U \in \mathcal{V}_{n,r};$$

that is, $U \in \mathbb{R}^{n\times r}$ has orthonormal columns. U is unique up to right-multiplication with an $r \times r$ orthogonal matrix. Tangent matrices in $\mathcal{T}_Y\mathcal{G}$ are of the form

$$\delta Y = \delta U U^T + U \delta U^T \qquad \text{with} \quad \delta U \in \mathcal{T}_U \mathcal{V}_{n,r}. \tag{6.3}$$

This representation is unique if we impose the condition $U^T \delta U = 0$, which yields $\delta U^T = U^T \delta Y$. The Galerkin condition (2.7) for the manifold $\mathcal{G}$ determines $\dot Y \in \mathcal{T}_Y\mathcal{G}$ such that

$$\langle \dot Y - \dot A, \delta Y\rangle = 0 \qquad \text{for all} \ \ \delta Y \in \mathcal{T}_Y\mathcal{G}.$$

Substituting (6.3) and using the rules $\langle A, B\rangle = \langle A^T, B^T\rangle$ and $\langle A, BC^T\rangle = \langle AC, B\rangle$, this condition becomes, with the condition $U^T\dot U = 0$,

$$\langle \dot U - \tfrac{1}{2}(\dot A + \dot A^T)U, \, \delta U\rangle = 0 \quad \text{for all } \delta U \in \mathbb{R}^{n\times r} \text{ with } U^T\delta U = 0.$$

This gives the differential equation

$$\dot U = P_U^{\perp} \tfrac{1}{2}(\dot A + \dot A^T)U. \tag{6.4}$$

With the appropriate version of Lemma 4.2 for the orthogonal projection $P(Y)$ onto the submanifold $\mathcal{G}$, the approximation estimates corresponding to Theorems 5.1 and 5.2 follow without further ado.

6.4. Minimum defect approximation of matrix differential equations. For the low-rank approximation to a solution of the matrix differential equation

$$\dot A = F(A), \tag{6.5}$$

condition (1.2) is replaced, at every time t, by

$$\dot Y \in \mathcal{T}_Y\mathcal{M}_r \quad \text{ such that } \quad \|\dot Y - F(Y)\| = \min! \tag{6.6}$$

Equivalently, condition (2.7) is replaced by the Galerkin condition

$$\langle \dot Y - F(Y), \, \delta Y\rangle = 0 \quad \text{for all} \ \ \delta Y \in \mathcal{T}_Y\mathcal{M}_r, \tag{6.7}$$

and correspondingly, the expression $\dot A$ is replaced by $F(Y)$ for $Y = USV^T$ in the differential equations (2.8) for S, U, V.

Theorems 5.1–5.3 extend to the low-rank approximation of matrix differential equations (6.5). We assume that F has a moderate bound along the approximations,

$$\|F(X(t))\| \le \mu, \qquad \|F(Y(t))\| \le \mu \qquad \text{for} \ \ 0 \le t \le \bar t, \tag{6.8}$$

and satisfies a one-sided Lipschitz condition: there is a real λ (positive or negative or zero) such that

$$\langle F(Y) - F(X), Y - X\rangle \le \lambda \|Y - X\|^2 \tag{6.9}$$

for all matrices $X, Y \in \mathcal{M}_r$. We further assume that for the best approximation $X(t)$,

$$\|F(X(t)) - F(A(t))\| \le L \|X(t) - A(t)\| \qquad \text{for } \ 0 \le t \le \bar{t}, \tag{6.10}$$

which is in particular satisfied if F is Lipschitz continuous with Lipschitz constant L. We then have the following extension of the quasi-optimality result of Theorem 5.1.

THEOREM 6.1. *Suppose that a continuously differentiable best approximation $X(t) \in \mathcal{M}_r$ to a solution $A(t)$ of (6.5) exists for $0 \le t \le \bar{t}$, and assume the bounds (6.8)–(6.10). Let the rth singular value of $X(t)$ have the lower bound $\sigma_r(X(t)) \ge \rho > 0$, and assume that the best-approximation error is bounded by $\|X(t) - A(t)\| \le \frac{1}{16}\rho$ for $0 \le t \le \bar{t}$. Then, the approximation error of (1.2) with initial value $Y(0) = X(0)$ is bounded in the Frobenius norm by*

$$\|Y(t) - X(t)\| \le (2\beta + L)\, e^{(2\beta+\lambda)t} \int_0^t \|X(s) - A(s)\|\, ds \qquad \textit{with} \ \ \beta = 8\mu\rho^{-1}$$

for $t \le \bar{t}$ and as long as the right-hand side is bounded by $\frac{1}{8}\rho$.

Proof. Equation (6.7) rewritten as in (4.1) reads

$$\dot{Y} = P(Y)F(Y). \tag{6.11}$$

As in the proof of Theorem 5.1, we have the equation

$$\dot{X} = P(X)F(A) + D \qquad \text{with} \quad \|D\| \le 2\beta d$$

for $d = \|X - A\|$. We subtract the two equations, write

$$\begin{aligned} P(Y)F(Y) - P(X)F(A) - D = {} & (P(Y) - P(X))F(X) + P(X)(F(X) - F(A)) \\ & + (F(Y) - F(X)) - P^{\perp}(Y)(F(Y) - F(X)) - D, \end{aligned}$$

and take the inner product with $Y - X$. With Lemma 4.2 we obtain

$$\begin{aligned} \langle \dot{Y} - \dot{X}, Y - X\rangle \le {} & \beta \|Y - X\|^2 + Ld \|Y - X\| \\ & + \lambda \|Y - X\|^2 + \beta \|Y - X\|^2 + 2\beta d \|Y - X\|. \end{aligned}$$

For $e = \|Y - X\|$ this gives the differential inequality

$$\dot{e} \le (2\beta + \lambda)e + (2\beta + L)d, \qquad e(0) = 0, \tag{6.12}$$

which yields the result. ∎

We refer to [12, Theorem 4.1] for a related quasi-optimality result in a situation of a linear differential equation with an unbounded operator.

In the differential equation analogue of Theorem 5.2 with the splitting (5.3), we start from the equations $\dot{Y} - \dot{X} = P(Y)F(Y) - P(X)\dot{X}$ and $\dot{X} = F(A) - \dot{E}$, yielding

$$\begin{aligned} \dot{Y} - \dot{X} = {} & (P(Y) - P(X))\dot{X} - P^{\perp}(Y)(F(Y) - F(X)) \\ & + (F(Y) - F(X)) + P(Y)(F(X) - F(A)) + P(Y)\dot{E}, \end{aligned}$$

where we now take the inner product with $Y - X$. If it is additionally assumed that F has Lipschitz constant L, then this leads to the differential inequality

$$\dot{e} \le 4\rho^{-1}(\beta + L)e^2 + \lambda e + Ld + \varepsilon, \qquad e(0) = 0. \tag{6.13}$$

With $\widehat{\gamma} = 4\rho^{-1}(\beta + L)$ and $\widehat{\varepsilon} = \varepsilon + L \max_{0 \le t \le \bar{t}} d(t)$, and with $\varphi(x) = (e^x - 1)/x$, this yields the error bound

$$\|Y(t) - X(t)\| \le 2t\,\varphi(\lambda t)\,\widehat{\varepsilon} \qquad \text{for} \quad t\varphi(\lambda t) \le \tfrac{1}{2}(\widehat{\gamma}\widehat{\varepsilon})^{-1/2} \tag{6.14}$$

and as long as $t \le \bar{t}$ and $2t\varphi(\lambda t)\widehat{\varepsilon} \le \frac{1}{8}\rho$.

Theorems 5.3 and 5.5 are extended similarly.

6.5. The special case of linear matrix differential equations. Systems

$$\dot{A} = LA + AR \tag{6.15}$$

with possibly time-dependent matrices $L(t)$ and $R(t)$ have the solution $A(t) \in \mathcal{M}_r$ for initial data $A_0 = Y_0 \in \mathcal{M}_r$. This is seen immediately from Lemma 4.1 and (6.11), which yield $\dot{Y} = LY + YR$ and hence $A(t) = Y(t) \in \mathcal{M}_r$. From the differential equations (2.8), we thus obtain a decomposition of the solution $A = USV^T$ with U and V having orthonormal columns and with the factors satisfying the differential equations

$$\begin{aligned} \dot{S} &= U^T LUS + SV^T RV, \\ \dot{U} &= P_U^{\perp} LU, \\ \dot{V} &= P_V^{\perp} R^T V. \end{aligned} \tag{6.16}$$

A different situation arises for linear problems of the type

$$\dot{A} = LA + AR + B \bullet A, \tag{6.17}$$

where $\bullet$ denotes the Hadamard (or entrywise) product of matrices. The differential equations (2.8), with $\dot{A}$ replaced by the right-hand side of (6.17) evaluated at $Y = USV^T$ instead of A, determine a low-rank approximation, but the entrywise multiplication with $B(t)$ in general requires the explicit computation of the entries of Y. The situation simplifies if B is itself a low-rank matrix $B = \sum_{j=1}^{k} \beta_j c_j d_j^T$. Writing $Y = \sum_{i=1}^{r} \sigma_i \widehat{u}_i \widehat{v}_i^T$ (obtained from an SVD of the matrix S of $Y = USV^T$), we can use $B \bullet Y$ in the differential equations for S, U, V in the decomposed form

$$B \bullet Y = \sum_{j=1}^{k} \sum_{i=1}^{r} \beta_j \sigma_i (c_j \bullet \widehat{u}_i)(d_j \bullet \widehat{v}_i)^T;$$

cf. [2] for an analogous observation for the potential in the Schrödinger equation. The dynamical low-rank approximation $Y(t)$ to $A(t)$ of (6.17) can thus be computed inexpensively if $B(t)$ is of low rank or otherwise approximated by a matrix of low rank, and in the present paper we have seen how this can be computed.

7. Conclusions and outlook. The dynamical low-rank approximation (1.2), or equivalently (2.7) or (4.1), becomes an attractive computational approach via the differential equations (2.8) that determine the factors in the representation (2.1) of the approximation. The method yields a near-optimal smooth low-rank approximation, as

is shown in Theorems 5.1, 5.2, 5.3, and 5.5 and observed in numerical experiments. A direct but very noteworthy extension is the minimum-defect low-rank approximation (1.3) to solutions of matrix differential equations. Our first numerical experience in compressing time-varying term-document matrices and series of images, and in approximating time-dependent PDEs whose solutions are essentially of low rank (e.g., smooth with the exception of a few pulses or spikes, as in blow-up problems in reaction-diffusion equations), is very promising, as reported in [14]. It will be interesting to see the dynamical low-rank approximation used for large-scale problems in applications, well beyond the already important area of quantum dynamics, where basic ideas for this approach originated 75 years ago as a physical model reduction technique.

REFERENCES

[1] M. Baumann and U. Helmke, *Singular value decomposition of time-varying matrices*, Future Generation Computer Systems, 19 (2003), pp. 353–361.

[2] M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer, *The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets*, Phys. Rep., 324 (2000), pp. 1–105.

[3] M. W. Berry, S. T. Dumais, and G. W. O'Brien, *Using linear algebra for intelligent information retrieval*, SIAM Rev., 37 (1995), pp. 573–595.

[4] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nichols, *Numerical computation of an analytic singular value decomposition of a matrix valued function*, Numer. Math., 60 (1991), pp. 1–39.

[5] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, *Structured low rank approximation*, Linear Algebra Appl., 366 (2003), pp. 157–172.

[6] M. Chu, N. Del Buono, L. Lopez, and T. Politi, *On the low-rank approximation of data on the unit sphere*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 46–60.

[7] L. Dieci and T. Eirola, *On smooth decompositions of matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 800–819.

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

[9] J. Frenkel, *Wave Mechanics, Advanced General Theory*, Clarendon Press, Oxford, 1934.

[10] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed., Springer, Berlin, 2006.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, London, 1985.

[12] C. Lubich, *On variational approximations in quantum molecular dynamics*, Math. Comp., 74 (2005), pp. 765–779.

[13] V. Mehrmann and W. Rath, *Numerical methods for the computation of analytic singular value decompositions*, Electron. Trans. Numer. Anal., 1 (1993), pp. 72–88.

[14] A. Nonnenmacher and C. Lubich, *Dynamical Low-Rank Approximation: Applications and Numerical Experiments*, technical report, Mathematics Institut, University of Tübingen, 2006; available online at http://na.uni-tuebingen.de/preprints.shtml.

[15] H. D. Simon and H. Zha, *Low-rank matrix approximation using the Lanczos bidiagonalization process with applications*, SIAM J. Sci. Comput., 21 (2000), pp. 2257–2274.

[16] K. Wright, *Differential equations for the analytic singular value decomposition of a matrix*, Numer. Math., 63 (1992), pp. 283–295.

[17] H. Zha and H. D. Simon, *On updating problems in latent semantic indexing*, SIAM J. Sci. Comput., 21 (1999), pp. 782–791.